Road Funds, User Charges, and Taxes
Ken Gwilliam and Zmarak Shalizi
159

Between the State and the Market:
Can Informal Insurance Patch the Safety Net?
Jonathan Morduch
187

Using Household Surveys to Build Analytic Capacity
Lorraine Blank and Margaret Grosh
209

Managing Government Exposure to Private Infrastructure Risks
Timothy Irwin, Michael Klein, Guillermo E. Perry,
and Mateen Thobani
229

The Evolution of the World Bank’s Land Policy
Klaus Deininger and Hans Binswanger
247

Climate Change, Agriculture, and Developing Countries:
Does Adaptation Matter?
Robert Mendelsohn and Ariel Dinar
277

What Does Climate Change Mean for Agriculture
in Developing Countries? A Comment on Mendelsohn and Dinar
John Reilly
295
Insufficient or uncertain budgetary allocations to road maintenance have resulted in road deterioration that has significantly increased production and transport costs in many countries. To avoid this problem, highway professionals advocate the establishment of dedicated road funds, managed by independent road boards made up of user representatives. The road boards would have the power to determine both the level of charges for road use and the level of expenditure on road maintenance. By contrast, macroeconomists and public finance specialists have tended to resist the establishment of dedicated road funds. They argue that road funds reduce fiscal flexibility, do not adequately address problems associated with the provision of public goods or the internalization of externalities, and often are not well managed.

In general, there are two long-term institutional options for reconciling fiscal prudence with asset maintenance: a road agency that is operated commercially (subject to the normal oversight of behavior accorded to privatized monopolies), or a reformed and well-functioning budget process. This article argues that road funds must be viewed as a provisional, case-specific intermediate step in the direction of one of the long-term solutions. The role and nature of road funds should be assessed not on general principles but on a case-by-case basis through the analysis of likely micro- and macroeconomic effects. The article recommends indicators for use in specific cases to determine whether a road fund should be introduced, continued, or abolished.

Spending on roads accounts for between 0.5 and 10 percent of public expenditure and between 10 and 20 percent of the development budget in many countries. Expenditures on road construction and rehabilitation, which are usually classified as capital expenditures, financed through borrowing and subsequent debt servicing, are not discussed in this paper. This article focuses on expenditures for periodic maintenance, routine maintenance, and operational management of roads. These costs are usually treated as current expenditures financed through annually budgeted allocations from the central treasury.
Road maintenance expenditures account for 30 to 60 percent of total road expenditures and as much as 0.5 percent of gross national product, or GNP (table A-1). Some weak evidence indicates that as road network densities and trade and traffic volumes increase with economic growth, road maintenance expenditures increase with GNP per capita (figure 1). In general, however, road maintenance expenditures depend not just on the size of the network but also on the nature of the terrain (higher costs in mountainous countries such as Bolivia and Nepal), the severity of the weather (higher costs in cold or wet climates such as Norway or Bangladesh), and construction standards. Insufficient or uncertain budgetary allocations undermine orderly planning and execution of road maintenance, resulting in road deterioration, which significantly increases transport costs. This article addresses the fiscal and institutional means for avoiding this outcome.

Deferring road maintenance increases not only total costs but also the present value of the future cost stream at any reasonable rate of discount. If roads deteriorate to the point at which they need reconstruction, restoring them to the original level of service costs three to five times more than timely and effective maintenance (Harral and Faiz 1988). Gyamfi (1992) estimates that the ratio of restoration costs to maintenance costs lies between 1.0 and 2.5 in Chile and Costa Rica. In addition, for every dollar that road agencies withhold by underfunding maintenance, road users must spend about three dollars in additional vehicle operating costs. Therefore, the economy as a whole (although not necessarily the road agency itself) benefits from timely road maintenance at any reasonable discount rate.

Figure 1. Road Maintenance Expenditure Patterns in Selected Countries, Annual Averages for the Period 1988–95

Source: Table A-1.
The World Bank’s Operations Evaluation Department (OED) database supports the assessment that road maintenance is a highly productive expenditure. The OED database covers 341 road projects evaluated between 1961 and 1988. The average estimated economic internal rate of return for this class of projects was 38.6 percent. This rate compares with an average of 26 percent for all transport projects and 21 percent for all World Bank investment projects (Heggie and Vickers 1998).

In the 1960s and 1970s many countries in Africa, Asia, and Latin America established road funds as an extrabudgetary arrangement. Earmarking of a stream of revenues (often from fuel taxes) was introduced for use by the road department or agency, with the intention of insulating maintenance expenditures from erratic and frequently delayed general budget procedures and reviews (Johansen 1989). (For a definition of earmarking and other terms used throughout the text, see the glossary.) Often, the World Bank requested that these funds be set up to protect its investments from economically inefficient asset deterioration. Some recently established road funds in Eastern Europe (in Russia and Georgia) are still based on this model.

From a macroeconomic viewpoint, analysts have criticized earmarking because it infringes on the policymaking powers of state executives and legislatures and reduces the leverage and flexibility of central governments (Deran 1965). Therefore the World Bank and the International Monetary Fund (IMF) have discouraged first-generation road funds in practice on general macroeconomic grounds (World Bank 1986). Even within the transport sector, some analysts have argued that the preferential access of road funds to lucrative revenue sources, such as gasoline taxes, hinders development of a more balanced, multimodal transport system.

The performance of state road agencies continues to be a matter of concern. Therefore a new generation of road funds is being established in the 1990s as part of an agenda to commercialize the road sector. This process runs parallel with (and primarily for the same reasons as) the privatization of state enterprises. Second-generation road funds are funded by levies or surcharges designated as user charges and identified separately from general taxation. Revenues are paid directly into a fund managed by a road board whose membership is chosen to represent users. The road board determines both the level of charges and the allocation of expenditures. Second-generation road funds have been established in several African countries and in countries as diverse as El Salvador, Guatemala, Jordan, Lebanon, and Pakistan (Heggie 1995).

Off-budget financing, or earmarking, is not universally necessary or effective. Many European countries with good governance maintain road systems through traditional public sector budgetary arrangements. Poor governance or governments’ lack of self-discipline may make it impossible to maintain roads even with the existence of a statutory road fund. Governments may not be able to guarantee the security of assigned revenue streams or the designated allocation of expenditures (Potter 1997a). Most countries fall between the extremes of good and bad governance. Assigning
responsibility for tax revenues and expenditures to a representative road board to manage a specific road fund (similar to specialized taxing districts) may make a real difference.

The next section sets out the analytical framework for assessing the desirability of road funds, which has much in common with the framework for fiscal decentralization. An important caveat is in order. Selective establishment of a road fund is justified only where quasi-commercial forms of organization might improve the allocation of resources through the implementation of a system of direct charges for road infrastructure use and a system of fund governance by users. Road maintenance does not inherently have a higher priority than other uses of public resources; thus, even though the discussion focuses on road funds, the general principles developed in the article apply equally to other sectors where quasi user charges and user governance can be implemented.

Analytical Framework for Considering Road Funds

The establishment of a road fund may affect the efficient working of the economy through three main channels. The first channel, fiscal control, influences the efficiency with which resources are collected and allocated among activities to maximize total community welfare. The second channel, management incentives, partly determines the efficiency with which the agents of production use the resources allocated to them. The third channel, rent-seeking behavior, can adversely affect both fiscal control and management incentives. Rent-seeking behavior occurs when individuals or agencies attempt to secure their own specific advantage at society's expense (Teja 1988). The relative importance of and balance among these channels critically affects the assessment of the efficacy of road funds.

Fiscal Control and Allocational Efficiency

Public finance economists argue against earmarking because the optimal charge on road users would be unlikely to generate the revenue stream required to finance the optimal stream of road maintenance expenditures in terms of scale or timing. In any given time period, any earmarked stream of revenue is likely to generate either insufficient or excessive funding. Too little funding causes the road authority to require continued recourse to the general budget. Too much funding creates the potential for financing lower-priority expenditures. In the former case, ensuring independence from poor budgetary processes and allocations requires rate-setting capabilities; in the latter case, avoidance of wasteful allocations requires public accountability through strict monitoring and auditing mechanisms. Often, the earmarking mechanism lacks these features.
A related argument against earmarking arises from the fact that budgets in developing countries are often very fragmented. The development budget is typically a set of separate budgets, ringfenced by sponsoring donors. The recurrent budget is also fragmented by the prior call of debt servicing and other statutory expenditures and by the large share of wage expenditures that are difficult to cut. The proliferation of earmarked funds by sector or objective would make it difficult to rationalize the allocation of resources as expenditure priorities change. In short, a problem of macro-economic control arises if extensive earmarking eliminates fiscal flexibility. This situation occurred in Colombia in the 1980s (Premchand 1983) and in Ecuador in the late 1990s.

The fiscal argument against earmarking in general and against road funds in particular is based on the assumption that general-purpose governments are better informed than special-purpose governments about the costs and benefits of alternative expenditure possibilities. The argument also assumes that general-purpose governments are committed to optimizing social welfare; that is, they are informed and benevolent. These assumptions may be true in some cases, where governance is representative and transparent. If the government system lacks these qualities, however, one of the main arguments against earmarking fails.

A related point is that current political pressures or the electoral cycle may result in myopic decisions. Vehicle operating costs, which do not enter into the road agency accounts, make up a high proportion of total transport costs (75–90 percent); these costs increase progressively as road conditions deteriorate (Harral and Faiz 1988). Unfortunately, road deterioration reveals its symptoms late. Expenditures on timely maintenance do not yield such obvious improvements in system performance as do expenditures on new investment. Yet long-term investment funded at the expense of optimum maintenance actually leads to a long-term decline in total available road system quantity and quality. The introduction of explicit road user charges, directed to a road fund in lieu of allocations from the general revenue budget, would contribute to allocative efficiency.

The introduction of explicit road user charges, however, would not automatically eliminate the need to address tradeoffs. In the absence of complete independence between specific road user charges and general taxes, securing funding for roads would entail an opportunity cost in other sectors. For example, in developing countries with low taxable capacity, fuel taxes represent a fairly secure tax source, accounting for 7 to 30 percent of total tax revenues and 1 to 3.5 percent of gross domestic product, or GDP (Gupta and Mahler 1995). The loss of control over this source of revenue may particularly damage the central government’s economic management abilities. Introducing an indirect road user charge, in the form of a surcharge on fuel taxes, would limit the government’s ability to increase taxes on fuel for general tax purposes. This limitation could lead to increased instability in the use of remaining tax revenue for social expenditures, such as health and education.
The independence of general taxing capacity from the level of road user charges is likely to be greatest when a group of beneficiaries is well defined and the payment of user charges is directly linked to the receipt of services. The benefit rationale for earmarking aims to reveal taxpayer preferences for public services and to send a demand signal to the public sector about how much of the public service to supply (Bird 1997). With an efficient and fair charging mechanism, no one receives a service without paying for it or pays without receiving the service. Second-generation road funds attempt to honor this principle by generating the bulk of revenues through vehicle license fees, axle loading or distance fees, and, in a few cases, toll revenues. To a lesser extent, they generate revenue through the separation of the pure tax element from the ex ante explicit user charge element of public revenues collected from fuel taxes.

Management Incentives and Operational Efficiency

The life of a highway investment and the benefits accruing from it depend on the maintenance of the facility. Most appraisals assume optimal maintenance, although they may not explicitly address what this implies. Failure to provide the required maintenance effort means that the return on the initial investment will be lower. If normal budgetary practices do not provide the necessary funding for optimal maintenance, then project designers and evaluators should reduce the likely benefit stream (and therefore the expected rate of return) or introduce complementary institutional mechanisms to ensure appropriate maintenance practices. In the first case, fewer investment projects would meet the criteria for selection. In the second case, establishment of a road fund to ensure funding for road maintenance from road user charges (quasi prices) may be the logical corollary of accepting projects with attractive rates of return, but in contexts where budget practices are poor.

The introduction of road user charges payable directly to a road fund can improve managerial incentives by increasing autonomy from unwarranted political interference. In many countries, wrangling over the budget delays its approval and disbursement. Studies in Latin America show that uncertainty or untimely availability of funding to maintain regular work schedules and to buy fuel and supplies explains in part the low equipment utilization rates and low number of kilometers maintained per employee (Gyamfi 1992). Even if the total level of road funding is open to competition from other demands, a road fund may enable the executing agency to perform more efficiently by guaranteeing the availability of a secure core of funding. In Ghana the establishment of a road fund has substantially reduced the problems of disruption of the planning and execution of maintenance work. These problems are caused by delays in approving the budget or releasing the budget allocations and lack of synchronization between the budget year (the calendar year) and the construction season (September to May). These delays necessitated the awarding of small continuation contracts to contractors who already had commitments with the adminis-
tration. The establishment of road funds has eliminated payment delays, given a significant boost to contractor cash flow, and reduced unit costs by 15 to 20 percent (Pankaj 1989).

The guarantee of a core of finance may also allow road agencies to extend and improve contracting arrangements with the private sector. The same studies for countries in Latin America and the Caribbean suggest that maintenance by force account (that is, staff on government payrolls) is little more than half as efficient as maintenance that is contracted out to the private sector. In Ghana the greater certainty of funding associated with earmarking allowed the introduction of effective competitive bidding. In general, more reliable financial arrangements lead to better use of resources.

Operational efficiency may also increase if users willingly pay for maintenance because the road authority channels payments more directly to the provision of a service of value to the users. (The availability of these additional resources, which might not be forthcoming otherwise, can also improve the government’s ability to manage macroeconomic imbalances.) Some countries, including many in Sub-Saharan Africa, have experienced a severe crisis in the maintenance of their main road networks. Heavy users, such as truckers and other operators of commercial vehicles, have demonstrated a willingness to levy an additional charge on their “own” use of fuel to finance a road fund with responsibility for maintaining a core network. There is no mystery to this behavior. Users more than recoup the surcharge if it is dedicated to fund better road maintenance, which in turn reduces vehicle operating costs. For example, Heggie (1995) estimates that vehicle operating costs decrease by two to four dollars for every additional dollar spent on road maintenance.

Rent-Seeking Behavior and the Distribution of Welfare

At the heart of the problem of traditional road funds was the failure of the associated earmarking arrangements to address incentive and governance issues. Unlike marketable commodities, including deregulated rail and airline services, the typical traditional road fund had no link between the tax rate (or the amount of taxes earmarked) and spending priorities (in light of the level of road use). Road fund managers had incentives to maximize their discretionary expenditures (including investment in low-priority roads or ancillary activities) rather than to optimize the level of road maintenance. The combination of public scrutiny and periodic monitoring by a competent central bureaucracy may provide some defense against this problem in industrial countries, but that combination is less likely in developing countries with less developed institutional capabilities.

Public choice theorists express skepticism even about institutional capabilities in industrial countries, including the political process that translates citizens’ preferences into public action. Essentially they believe that diverse citizens’ preferences do not permit aggregation into a well-defined community preference function. They
also believe that monitoring costs and informational asymmetries may enable public officials (regardless of whether they respond to organized pressure groups) to project their personal interests onto their function of allocating resources. And they believe that budget choices do not depend solely on the inherent costs and benefits of services but also on the ability of one set of taxpayers to transfer the costs of programs that benefit them to others.

Where individual preferences for public goods differ, separate earmarked funds could potentially increase general welfare if the payments to those funds reflect individuals’ relative marginal utilities for different public goods (Johansen 1963). Despite the ingenuity devoted to designing ways of getting consumers of public goods to reveal their marginal utilities truthfully, this analysis remains difficult to apply practically. Quasi prices, or user charges, may have welfare advantages because they can be levied approximately in proportion to the demonstrated benefit of consumption.

Public choice theorists have pointed out a fundamental flaw in general fund budgeting. That is, heavy consumers of a service that is financed through general taxes would benefit from lobbying for larger expenditures on that service (thereby transferring welfare to themselves). At the same time, nonconsumers would argue for lower expenditures. The outcome depends on the respective political power of the parties, rather than the aggregate value attached to each individual service. The road authority can eliminate this bias by setting prices, such as tolls and vehicle duties, for the beneficiaries of a specific service. Using a fuel surcharge as a quasi price for road use (with appropriate corrections for agricultural vehicles and for fuel not used for road vehicles) is analogous to establishing a special taxing district. These districts are common in the provision of some facilities, such as water, and could be consistent with the government’s pursuit of redistribution objectives through its policies on general taxation and the allocation of merit goods.

The argument for earmarking as a way of separating allocation and distribution issues may also be applied spatially. The government could implement a program of regional financing for services consumed regionally. This program would avoid overprovision in some regions at the expense of others as the regions compete to maximize their share of the national budget. Of course, the road authority may justify some regional disparities in provision, particularly of road investment, on both efficiency and equity grounds. The government would require operational criteria for the spatial allocation of resources, as have recently been developed for second-generation road funds, regardless of whether the funds are earmarked.

Empirical Evidence

The influences discussed above do not militate in the same direction and may demonstrate variable quantitative significance in different circumstances. Therefore, the
evaluation of second-generation road funds cannot be resolved solely by reference to a priori theorizing but instead must be approached on a case-by-case basis in light of national circumstances. Several issues affect the balance between the micro- and macroeconomic considerations, including the demonstrated effectiveness of road funds in meeting their public service objectives, the magnitude of the economic cost of undermaintenance, the extent of the backlog of road maintenance work, the demonstrated capability of existing fiscal arrangements to finance efficient maintenance, and the degree of current fiscal dependence on taxation of road users.

**Experience with the Operation of Road Funds**

The empirical evidence suggests a positive but statistically weak ability of traditional forms of earmarking to alter resource allocation significantly (Eklund 1967). Wherever the government retains control over the level of the user charges or over the allocation of complementary funds, the total level of funding may be just as vulnerable with a road fund as without one. For example, the earmarked funds for the Colombian National Road Fund grew at the same rate as GDP in 1979–87, but road expenditures per kilometer fell by 25 percent (Dick 1989). The fall in expenditures occurred in part because funds were preempted to cover increased debt service payments.

McCleary (1991) reviews the experience in the Central African Republic, Colombia, Ghana, Mali, and Zaire (now the Democratic Republic of Congo) in the late 1980s. He concludes that the road fund was successful only in Ghana, where the government was strongly committed to increased expenditure on roads in any case. In general, the governments found it difficult to set appropriate levels for earmarked taxes. Allocations tended to continue to depend on the general budget situation, and the adequacy of overall road fund resources provided no assurance of an appropriate balance among maintenance, rehabilitation, and new investment. A more recent review of 10 African road funds confirms many of these failings (box 1). It also emphasizes the need to examine more closely why the funds failed. In contrast, second-generation road funds overcame many traditional defects when they had clear objectives, an independent revenue source, and efficient management and accounting arrangements (Balcerac de Richencour and Heggie 1995). Some newer funds (for example, in Ghana and Zambia) ensured the automaticity of payment by separating the fuel levy from general taxes.

**Experience with Existing Fiscal Arrangements**

A survey of the views of country road agencies and task managers in nine countries shows the ratio of actual maintenance expenditures to those considered necessary to maintain constant road conditions. The ratio varies between 29 percent for Jordan and 89 percent for Argentina, with most countries falling in the 50–70 percent range.
Continued inadequate maintenance of the road network

**Insufficient revenue base.** The revenue base of the road fund is large enough to finance only part of the qualifying expenditures, with the balance financed through the government's general budget. Once road agencies obtain some money from the road fund, they may find it even more difficult to get funds through the normal budgetary process.

- **Ghana.** The road fund's original revenue base could finance only 60 percent of periodic maintenance.
- **Zaire.** In 1985 a higher fuel tax increased the revenue base of the existing road fund, but it was insufficient to cover all of the qualifying expenditures. In 1986 the new budget contribution dropped to zero, resulting in a drop in total road funding. The situation was repeated in 1987.

**Funds raided by the general budget.** The revenues are collected by the customs and excise department and channeled through the ministry of finance, which almost invariably withholds some of these revenues.

- **Sierra Leone.** In February 1990 the central customs authority reduced the amount paid into the road fund and in June 1991 suspended payments altogether, even though the road users continued to pay the same fuel levy.
- **Similar situations occurred in Ghana, Mozambique, and Tanzania.**

**Poor governance, limited operational efficiency, and misappropriation of funds.** Without satisfactory audit procedures, the government cannot track how monies paid into the road fund are spent. Nor can the government ensure that only approved expenditures are covered and that these expenditures meet required specifications.

- **Mozambique.** Audit reports for 1993 and 1994 could not verify the flow of money between Petromac (the oil company that imports fuel), the Ministry of Finance, and the road fund. Nor could the audit determine whether the road fund had received all the revenue due it from the levies on gasoline and diesel fuel.
- **Rwanda.** The 1991 audit report was unable to certify accounts because of a general lack of financial information and a lack of specific information on the revenue side.
- **Sierra Leone.** The audit for September 1989 to December 1991 showed that about $200,000 was used to purchase vehicles that were never delivered.
- **Tanzania.** The audit report for fiscal year 1992/93 showed that about $1.5 million in payments was made without any supporting vouchers or other documents.
- **Zambia.** $760,000 of the revenue collected for roads between May 1993 and November 1994 never reached the road fund. The audit also showed that about $500,000 was paid for materials that were never supplied, about $6 million was paid to contractors without any reliable records to show that work had been done, and about $70,000 was paid for the purchase of vehicles that were never delivered.

**Misallocation of resources**

**Diverting funds to low-priority and nonrelated activities to avoid showing a surplus.** When the road fund finances roads managed by different agencies without any transparent and equitable mechanism for allocating revenues among them, allocations are often subject to political whim rather than economic criteria. A weak or nonexistent road fund board and ambiguous legislation add to the problem.

- **Central African Republic.** The 1993–94 audit report showed that the central government took an irregular loan of $340 million from the road fund to pay civil service salaries.
• *Sierra Leone.* From September 1989 to December 1991, there were no guidelines regarding use of funds. The presumption was that the funds were intended for road maintenance. Substantial sums were spent on refurbishing offices, purchasing 1,800 yards of carpet, and carrying out repairs at the State House and the Parliament building.

• *Tanzania.* The audit report for fiscal year 1992/93 found that the road fund was used to finance recurrent expense items not covered by the directives issued by the Ministry of Finance, including leave payments, ferry operators, electric bills, and gratuities. Of the funds allocated for urban and rural roads, 75 percent went only to urban areas because they were better able to prepare plans—not because their needs were inherently greater.

• *Zambia.* Between May 1973 and November 1994, $4.7 million was released through provincial accounting control units for road rehabilitation. No guidelines for disbursing funds were provided, no expenditure returns were submitted, and no expenditure returns were requested before further funds were released. About $600,000 was paid for items that should have been covered through normal budgetary provisions.

*High revenue flow into the road fund.* A persistent buildup of a surplus that is not accessible by the general budget can severely limit the financial resources of the central government in a fiscal crisis.

• *Colombia.* In 1986 more than 35 percent of central government tax revenue was earmarked, with the road fund alone accounting for nearly 7 percent.

*Poor governance because the governing agency is captured by those who shift allocative priorities.*

• *South Africa.* Its initial road fund board, which primarily consisted of provincial representatives, found it difficult to get the members to act in the national interest.


(1998). Table 1 shows even more pessimistic conclusions from assessments of maintenance provisions based on the World Bank’s Highway Design and Maintenance Standards model for World Bank projects.

**Backlog of Road Maintenance Work**

In a study of road deterioration in developing countries, Harral and Faiz (1988) estimate the annual maintenance expenditures required to prevent deterioration. For 1986–90, expenditures varied from 0.2 percent of GDP, on average, for countries in East Asia and the Pacific to 1.0 percent for countries in West Africa. Harral and Faiz estimate that the backlog of maintenance work varied from 1.6 percent of GDP in East Asia and the Pacific to 3.5 percent in South Asia. Although these calculations have not been updated recently, anecdotal evidence suggests that the pattern persists. These estimates in table 1 do not suggest that lack of maintenance in the road sector was worse than the underfunding of maintenance expenditures in the provision of other public goods and services. And they do not suggest that the recommended levels of expenditure were socially optimal in economic terms. The estimates in table 1 show only the extent to which funding was insufficient to maintain constant service levels for an economically critical asset.

Ken Guwilliam and Zmarak Shalizi
Table 1. Inadequacy of Annual Road Maintenance Expenditures, Selected Countries, 1980s–1990s

<table>
<thead>
<tr>
<th>Time period</th>
<th>Country</th>
<th>Network</th>
<th>Amount spent</th>
<th>Amount recommended</th>
<th>Ratio of amount spent to amount recommended (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>Bangladesh</td>
<td>Roads and Highways</td>
<td>24.6</td>
<td>42.4</td>
<td>58</td>
</tr>
<tr>
<td>1991–92</td>
<td>Cameroon</td>
<td>All</td>
<td>28.1</td>
<td>44.9</td>
<td>63</td>
</tr>
<tr>
<td>1988–91</td>
<td>Honduras</td>
<td>All</td>
<td>11</td>
<td>45</td>
<td>24</td>
</tr>
<tr>
<td>1991–92</td>
<td>Madagascar</td>
<td>All</td>
<td>8</td>
<td>27.9</td>
<td>29</td>
</tr>
<tr>
<td>1992</td>
<td>Nepal</td>
<td>All</td>
<td>2.6</td>
<td>18.1</td>
<td>15</td>
</tr>
<tr>
<td>1988</td>
<td>Nigeria</td>
<td>Federal highways</td>
<td>112.3</td>
<td>248.5</td>
<td>45</td>
</tr>
<tr>
<td>1991–92</td>
<td>Rwanda</td>
<td>All</td>
<td>4.1</td>
<td>8.6</td>
<td>48</td>
</tr>
<tr>
<td>1991–92</td>
<td>Uganda</td>
<td>All</td>
<td>4.1</td>
<td>12.7</td>
<td>32</td>
</tr>
<tr>
<td>1992</td>
<td>Zambia</td>
<td>All</td>
<td>6.1</td>
<td>32.7</td>
<td>19</td>
</tr>
<tr>
<td>1990–91</td>
<td>Zimbabwe</td>
<td>All</td>
<td>25.9</td>
<td>35.7</td>
<td>73</td>
</tr>
</tbody>
</table>

a. The recommendation was based on maintaining at least the existing network at a level that would generate a constant flow of services (except for Nigeria, where the recommendation was intended to eliminate a five-year backlog).


Strategic Options and Interim Arrangements

The previous sections show that there is some evidence of a backlog in road maintenance work and that savings in operating costs could fund increased payments for road use without reducing general taxable capacity. The critical question is whether road funds constitute the most effective way of achieving an efficient allocation of resources for implementation of road maintenance expenditures. The decision to introduce or eliminate road funds must be based on a practical and systematic assessment of the context and of the available options for reconciling the micro- and macroeconomic objectives. Two broad options are available.

Option 1: A Commercial Road Agency

In the first main option, the government would move toward commercialization of the road sector. This would involve creating an independent, regulated road authority (similar to a monopolistic public utility). The road authority would have responsibility for the entire road network and for revenues derived from direct user charges (quasi prices) rather than from taxation. A first step in this direction might involve creating an independent road board to manage a road fund financed through various
instruments, including a surcharge on fuel taxes. In the longer term, a commercial road agency would replace the road fund. Tolls charged on a fee-for-service basis or charges more directly related to the costs imposed by road use (for example, charges for trucks based on axle weight and distance traveled) would replace the fuel surcharge. New Zealand is already moving in this direction.

Option 2: A Reformed Budget Process

In the second main option, the government would rebuild the capability of, and confidence in, its budgetary processes. In some cases, commercialization is not politically or practically feasible, and the public is unwilling to pay additional general taxes. Earmarking some special-purpose taxes and creating a road board–cum–road fund may provide the most practical interim means for generating additional funds for a priority economic activity and rebuilding public confidence in government. Even weak governments may accept some transient impediments to raiding activities with high benefits but low profiles, thereby changing the balance of expediency. In the longer term, the government would phase out the road fund and return all revenue and expenditure responsibilities to the budget. Countries in the European Union commonly rely on the budget rather than on road funds.

Interim Arrangements

The decision whether to establish a road fund as an interim step in option 1 or option 2 is a complex issue and requires case-by-case analysis (box 2). Three principles guide the decisionmaker about whether to introduce (or retain) a road fund:

- Will it improve resource allocation—for example, by ensuring better funding of activities with economically high returns but politically low profiles (bearing in mind that other services such as primary schooling and basic health clinics may also fall in the same category)?
- Will it improve operational efficiency—for example, through the introduction of better incentives for managing resources?
- Will it reduce rent-seeking—for example, by strengthening the link between benefits and payments?

Expenditure and Revenue Assignments and Governance

Assuming that conditions justify the creation or continuation of a road fund, its charter should explicitly address three broad issues. First, what road expenditure line items should the road fund protect, and what is the primary purpose of the fund? Second, what revenue streams or revenue authority should finance the chosen expenditure items
Box 2. Conditions for Introducing a Road Fund

Introduction of road funds may be justified if all the following conditions apply.

Maintenance is poor because of:

- **Insufficient funds.** Poor setting of budget priorities with bias in favor of new investments, often donor driven.
- **Unreliable timing of funds.** Poor budgetary processes that cannot ensure credible commitments, disbursements, or both.
- **Inefficient implementation of works.** Absence of incentives to use resources efficiently in the agency.

There is political commitment to increase maintenance expenditures on roads

There is a political commitment to establish long-term reliable mechanisms for improved allocation and accountability for the core network.

**Potential indicators**

- Asset condition of core network predicted to decline over the next 10-year period (increasing percentage in poor condition by road class).
- User costs predicted to increase over 10 years (increasing ratio of vehicle operating costs/vehicle/year by vehicle class).
- Substantial maintenance forgone (with economic rate of return greater than 20 percent).
- Net present value of near-optimal program more than 1.5 times higher than that of current program.
- Total costs per mile of road maintained in the core expected to drop by 25 percent throughout the life of the fund (relative to current expenditures or future benchmarks).
- Cabinet-level commitment (acts, regulations, gazetting) to increase road maintenance expenditures.
- Cabinet-level commitment to permit direct user charges (or surcharges on fuel taxation) to generate funds.
- Cabinet-level commitment not to reduce parallel funding.
- Principles accepted for major allocation decisions.
- Representatives of key user groups included on the road board.
- Economic criteria accepted as key to setting priorities.

**Note:** The size of the core network will vary over time as unused and lightly used routes are dropped and emerging heavily used routes are included.

in the road fund? And third, how will the proposed institutional structure ensure responsible governance and appropriate incentives to reconcile the conflicting micro- and macroeconomic management objectives? In particular, what features will ensure that unrepresentative or unaccountable interests will not capture the funds?

**Expenditure Assignments**

Should the government or the road authority decide about the allocation of resources between investment and maintenance, among regions or road types, and between administration and implementation activities?
ALLOCATION OF RESOURCES BETWEEN INVESTMENT AND MAINTENANCE. The most commonly identified problem, systematic bias against maintenance, occurs in fiscal regimes that fund both investment and maintenance through the same channels (with or without road funds). The regime tends to favor investment because large schemes are politically attractive and have easier access to external financing. In addition, the benefits of investment are more apparent ex ante than the benefits of maintenance. As a result, decisionmakers behave as if their discount rate were greater than the technical rate. This problem commonly occurs in developing countries as well as in transition economies, such as in Eastern Europe.

The creation of a road fund with both investment and maintenance responsibilities does not automatically ensure against such a systematic bias. In Mali the road fund provided the national counterpart funding for foreign lending for road construction and financed the servicing of debt on earlier investments. These became the first calls on the road fund, and maintenance suffered accordingly. This suggests that road funds be dedicated to maintenance, thereby providing a needed counterbalance to a systematic bias.

Experience provides less clear guidance on the treatment of investment. Several countries (including Japan, the Republic of Korea, South Africa, and the United States) introduced road funds to facilitate crash investment programs. The governments considered the investment programs too large for the general budget, thus justifying special treatment, including extra special-purpose taxation. But such arrangements can create a temptation to misallocate funds to lower-priority investments if they continue to generate large amounts of revenue after the real need that stimulated their creation has been satisfied. This problem, combined with the systematic bias in favor of investment, would appear to provide sufficient reason to exclude investment from any ringfenced allocation in an interim arrangement.

Unlinking closely related investment and maintenance requirements, however, is not appropriate over the long run. Freed from the responsibility of funding the operating and maintenance consequences of their investment decisions, governments and donors could continue to indulge in excessive road investments, leading to debt-servicing and maintenance-funding burdens that, combined, are unsupportable. This concern would argue in favor of a commercialized road agency or road fund, recouping from users the costs of investment and maintenance. For the road agency or road fund to influence investment expenditures, however, it would need a broad membership that extended beyond direct road users (for example, to groups displaced or inconvenienced by new road construction). It would be politically difficult to delegate such right-of-way and eminent domain functions to an entity such as a road fund.

ALLOCATION OF RESOURCES AMONG REGIONS OR ROAD TYPES. Another reason for excluding investment is that road boards dominated by user representatives may not
allocate investment resources optimally beyond the narrow confines of road maintenance. Roads perform social and strategic, as well as economic functions, which may justify cross-subsidies. For example, governments keep some rural roads in existence and repair even though their users cannot pay sufficient sums to maintain them. Ecological or aesthetic reasons in environmentally sensitive areas could require additional expenditures on roads beyond the amounts users would willingly pay. The same is true for strategic functions. There are several ways to approach this issue, but all of them introduce complications in the governance of road funds.

First, the road authority could introduce a system equivalent to the public service obligations of transit operators. Such a system would use a contracted payment from the state to compensate the commercial road fund for meeting an explicit public obligation. This arrangement reduces the neatness of the separation between the road fund administration and the political process and reintroduces substantial scope for negotiations about financial responsibility.

Second, the government could expand the road board to include representatives of noncommercial, environmental, and local interests in the management of the road fund. The expanded road board would address issues of acquisition of rights of way, resettlement, and other problems associated with the expansion of the road network. This solution, too, would reintroduce an element of politicization.

Third, the government could create multiple agencies, each concentrating on a more restricted set of roads with representative management. This arrangement, however, could restrict legitimate transfers across regions or types of roads. Box 3 summarizes the issues created by multiple road funds.

<table>
<thead>
<tr>
<th>Box 3. Single or Multiple Road Funds</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Single road fund</strong></td>
</tr>
<tr>
<td>National</td>
</tr>
<tr>
<td>Benefits</td>
</tr>
<tr>
<td>• Potential to cross-subsidize between high-use and low-use links, particularly if the fuel surcharge applies to all road users irrespective of the specific links they use.</td>
</tr>
<tr>
<td>Issues</td>
</tr>
<tr>
<td>• How to design a transparent spatial or functional allocative formula, for example, by population, length of road network, type of network, priority.</td>
</tr>
<tr>
<td>• How to ensure the representativeness of the board, especially for the purpose of setting priorities and allocating resources.</td>
</tr>
<tr>
<td><strong>Multiple road funds</strong></td>
</tr>
<tr>
<td>By type of road (rural, interurban, urban)</td>
</tr>
<tr>
<td>By administrative unit (district, province, city)</td>
</tr>
<tr>
<td>• Potential for charge discrimination based on relatively distinct user segments.</td>
</tr>
<tr>
<td>• Whether each fund should have its own revenue base, allocative criteria, and governance structure.</td>
</tr>
<tr>
<td>• How to deal with components of subnational roads that are part of the national network.</td>
</tr>
</tbody>
</table>
Allocation of resources between administration and implementation. Staff complements for road agencies are often adequate and well protected. But salaries for road agency staff are usually well below the market rate, with the result that the agency tends to be weak and dysfunctional. A commercialized road agency (as well as a reformed civil service–cum–budgetary arrangement) is more likely to do a better job of determining the quantity, level, and remuneration of staff, thereby leading to a more efficient allocation of resources between administration and implementation. The interim road fund arrangement might still have a strong inducement for management efficiency if user charges (or earmarked taxes) finance operational expenditures. Core management staffing would remain funded through the central budget, provided that the availability of this core budget is linked to some indicators of performance that are public and transparent.

Revenue Assignments

Determining the level and sources of funding is not straightforward. Defining the level in terms of expenditure categories (for example, rehabilitation and periodic and routine maintenance) invites the padding of these expenditures and the substitution of these items for other categories of expenditure. Defining the level in terms of the whole, or a predetermined proportion, of a particular tax could be inappropriate because tax yields and expenditure requirements change relative to each other over time. Making the level of funding subject to regular review returns the whole issue to the political arena. One advantage of a road board representing user interests is that it would most likely exert a strong downward pressure on spending.

Defining the sources of funding is somewhat easier. Where direct charges, such as tolls or weight-distance charges, are feasible, there is little conflict with fiscal objectives. The road authority could collect a weight-distance charge from truckers and channel the revenue to a road fund. Iceland, New Zealand, Norway, and Sweden charge weight-distance fees to diesel vehicles according to axle configuration and gross vehicle weight. In practice, the system of fees is difficult to administer and vulnerable to evasion. In New Zealand administrative work such as collection and enforcement absorbs 5 percent of gross revenues; evasion is estimated to be between 10 and 20 percent (Heggie 1995). Annual vehicle taxes, even if weight related, are weaker proxies because they do not reflect distance traveled and hence embody perverse incentives to intensify vehicle use.

Similar problems arise with the use of a fuel tax surcharge as a proxy user charge (figure 2). Road use generates several externalities, of which road damage—requiring maintenance—is only one. A surcharge on fuel use provides a reasonable proxy for the road damage externality caused by automobiles and implies a modest tax on gasoline.

Ken Gwilliam and Zmarak Shalizi
Figure 2. Gas or Diesel Fuel as a Base for Road Maintenance Charges

<table>
<thead>
<tr>
<th>Component of retail fuel price</th>
<th>Basis for charge</th>
<th>Appropriateness of fuel tax to basis for charge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Road user charge</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pollution</td>
<td>Medium/high</td>
<td></td>
</tr>
<tr>
<td>Congestion</td>
<td>High</td>
<td></td>
</tr>
<tr>
<td>Accidents</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>General tax element (VAT, sales tax, excise duty)</td>
<td>Sumptuary taxation</td>
<td>High</td>
</tr>
<tr>
<td>Border price of fuel (including distribution markup)</td>
<td>Resource cost of fuel</td>
<td>Not applicable</td>
</tr>
<tr>
<td>Exchequer/treasury (for reallocation to appropriate agencies; some relating to pollution, congestion, and accidents might be passed to a road agency)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Authors' calculations.

Fuel use, however, is not as good a proxy for road damage by trucks, which account for the bulk of road damage. The impact of a heavy truck is 10,000 times as great as that of a medium auto (Newbery 1987). It also varies nonlinearly with axle weight. Truck use would require a very high surcharge on diesel fuel, thereby creating problems for the use of diesel outside the transport sector (as in power generation or agriculture). Otherwise, the charge is too low to collect the necessary revenue or to affect truck use.

A surcharge on fuel may be appropriate for other externalities, such as air pollution, but not for other externalities such as congestion (which varies by time and location) or safety. Given the multiple claims on the surcharge, the road authority need not direct all the revenue generated by the surcharge to a road administration,
a road fund, or a commercialized road agency. Some of the road user charges will have to be transferred to other budgets (see figure 2).

If a fuel charge is used, the road authority should base the design on a transparent and supportable predetermined formula that is directly linked to the externalities or services consumed. The introduction of a fuel tax surcharge as a quasi price for road use should be approached particularly cautiously if the following circumstances apply:

- Taxation on fuel is a high proportion of total tax revenue (say, more than 10 percent), and any surcharge is likely to detract from general taxing capability. This situation occurs in many small developing countries. For example, the proportion is 19 percent in Costa Rica, 16 percent in Guatemala, and 13 percent in Nepal.
- The general fiscal situation is weak and many sectors are making similar claims for special fiscal treatment (as in many countries with internal deficits in excess of 5 percent of GDP or tax-to-GDP ratios of less than 15 percent).
- Diesel fuel is used extensively in power generation and in agriculture (as in many countries in South and East Asia).

The road authority should anticipate and avoid the danger of resistance to supplementary funding from the general budget, once the road fund has been set up. Road maintenance may still require funding from the general budget. Insofar as the efficiency arguments in favor of a partially protected budget relate to the ability to plan and phase a core maintenance work program, the benefits of a road fund do not depend on complete separation of the road budget from the rest of the budgetary process. Funding for links outside the core network or to cope with unforeseen circumstances (for example, natural disasters) would still require appropriations from the general budget.

**Governance**

The crux of the argument in favor of second-generation road funds is the separation of road user charges from general taxes and improved arrangements for governance of these funds. Many traditional road funds failed because funds accumulated or were misallocated. Better governance is essential to ensure that budget constraints are hard and that expenditure decisions are responsive to users. The following four institutional components should therefore be included in the package.

1. The package should have a strong legal basis. Road funds should be established by law to ensure clear terms of reference and some minimum protection from arbitrary political interference. The legal instrument not only needs to guarantee the source of funds but also must ensure that the funds are automatically channeled to
Where the setting of user charges and taxes overlaps between the road board and the government, the instrument should establish clear procedures enabling the executing agency to have the greatest possible security of a base level of funding and the greatest possible notice of changes in that base. It is unreasonable, however, to expect legal commitment to any particular level of tax or tax yield.

2. The package should have an independent executive authority. The need for efficient maintenance suggests that the executing agency should be accountable in a clear and transparent framework. It should have the primary role of formulating maintenance policy, marshaling and allocating funds, and securing effective implementation. Where these functions are clearly stated and well publicized as the responsibilities of a quasi-independent executive, as in Brazil and Chile, they establish proper management incentives and facilitate effective performance.

3. The package should have a third-party monitoring system. Given the problems of securing representative governance, a monopoly supplier of road services (the operator of the road fund or commercialized road authority) should be subject to regulatory supervision. The fund or authority should be required to inform the public and supervisory authorities of its activities; its accounts should be externally audited and periodically reviewed by an independent review body; and, if corporatized, its net income should be subject to corporate income tax.

4. The package should delineate administrative competence and proper criteria for expenditure. Assigning expenditure responsibilities to the road fund does not per se ensure efficient allocation within the ringfenced area. As necessary conditions for assigning revenues, the executing agency should have well-established procedures for allocating funds efficiently and the necessary administrative competence to administer the allocation and to monitor and report on performance. Mozambique and New Zealand have developed effective procedures for allocation of funds (Heggie and Vickers 1998).

Where the road fund is established as a temporary arrangement until general budgetary procedures are improved, termination criteria and a sunset clause must be in place to determine what should occur when effective budgetary procedures are judged to have been reestablished. Box 4 summarizes suggested criteria for the termination of earmarking taxes to a road fund based on an independent review.

Where the road fund is established as a stepping-stone to a commercial entity because the government wants to disengage from the direct production of goods and services, a review should be commissioned. If the review determines that the performance criteria set out in box 2 have been satisfied, it should then formalize the road authority as a commercialized public utility with user charges accruing directly to it rather than passing through the government treasury. Subsequently, the road authority should be subject to the same general form of public scrutiny as other privatized monopolies. The response to an unfavorable review should be, as
### Box 4. Sunset Provisions for Road Funds

**Closure of road funds may be justified if the following apply.**

**Potential indicators**

- Ratio of revenue flow to the fund to the required (nearly optimal) maintenance expenditure flow less than 0.7.
- Ratio of actual expenditures to required maintenance expenditures less than that of the revenue flow to required expenditure, and the actual expenditure flow also less than that before institution of the fund.
- Increase over time of predicted real vehicle operating costs for the roads maintained.
- Actual maintenance cost per kilometer of road greater than comparable benchmark.
- Percentages of roads in fair and poor condition increasing.

**Misallocation of resources:**

- Wasteful spending to avoid showing a surplus.
- As a result of high revenue flow into the fund, a persistent buildup of a surplus that is not accessible by the general budget.

**Poor governance due to capture that shifts allocative priorities.**

- Inadequate representation of stakeholder groups in board management structure.
- Inadequate public reporting of plans, expenditures, and road conditions.
- Maintenance expenditures on links not in proportion to the corresponding economic rates of return from maintaining the links.

---

in most countries, a matter for political and legislative rather than simple administrative action.

### Conclusions and Recommendations

Road funds are a continuing subject of controversy. Highway specialists regard them as a boon to facilitating the provision and maintenance of a highly productive asset by means entirely consistent with the general shift away from direct government production of goods and services. Macroeconomists and public finance specialists have tended to regard them as a bane because they reduce fiscal flexibility, do not
adequately address problems associated with the provision of public goods or the internalization of externalities, and are often not well managed.

This article has argued that the issue is not one to be resolved on general principles but on a case-by-case basis through the analysis of likely micro- and macroeconomic effects. In general, two radically different long-term options reconcile fiscal prudence with asset maintenance. The first is a fully commercially operated road agency, subject to the normal oversight of behavior accorded to privatized monopolies. The second is a reformed and well-functioning budget process. Neither option exists at present in most developing and transition economies. Thus, any recommendations on the role and nature of road funds must be viewed as a provisional, case-specific intermediate step in the direction of one of the long-term solutions (Potter 1997b).

This article lists a few indicators that can be used. In addition, the article argues that the interim arrangement requires four conditions. First, the road fund’s expenditure responsibilities should be limited to maintenance in order to correct a systematic bias against maintenance despite the link between investment and maintenance. Second, the road fund’s revenues should come only from direct charges on road users, except in the case of fuel surcharges (which need to be separated from general taxes in agreement with the treasury). Where feasible, weight-distance charges should be introduced for trucks. Third, the road fund should have professional management, under the direction of a user representative board and subject to strict oversight and auditing arrangements by third parties. Fourth, the road fund requires explicit transition arrangements as it moves toward a long-term solution.

As recent experience in Africa shows, when there is a crisis and main road systems fail, commercial interests can be mobilized to pay a surcharge on the existing fuel tax, so long as the user charge is devoted to improving the quality of the road infrastructure. Experience with second-generation road funds is very limited, however. It remains to be seen whether road fund management will be sufficiently immune to the kind of political interference that currently disturbs the flow of funds into road maintenance. And road authorities need to overcome the immediate, most extreme problems of deficient maintenance.

Therefore, analysts and policymakers should monitor carefully the recently introduced road funds in Sub-Saharan Africa and elsewhere. Sector programs and investment projects establishing these funds should include appropriate and comparable monitoring and evaluation components. Monitoring and evaluation of second-generation road funds will help to determine the utility and applicability of such interim arrangements and facilitate the amendment of initial designs on the basis of experience.

Notes

Ken Gwilliam is a transport economist with the World Bank’s Transport, Water, and Urban Development Department, and Zmarak Shalizi is a research manager with the Development Economics Division, World Bank.
Research Group, World Bank. A longer and more technical version of this article appeared as World Bank TWU Discussion Paper 26. The authors are grateful for the comments of Chris Hoban, Ian Heggie, Gunnar Eskeland, Vinaya Swaroop, and colleagues at the IMF (particularly Barry Potter). Debu Talukdar and Elysa Coles provided research assistance.

1. For this result, a heavy truck is defined as a load of about 10 tons (21,000 pounds) per axle and a medium car as a load of about 1 ton (2,000 pounds) per axle. The road damage of a vehicle axle load, $L$, (in tons) is calculated using the fourth-power formula: $[L/8.2]^{4}$. It is measured in equivalent standard axle loads, or ESALS; one ESALS is defined as 18,000 pounds (Hau 1992). For a more technical analysis, see Paterson (1988).

2. The World Bank policy review paper Sustainable Transport: Priorities for Policy Reform (1996) argues that fuel taxation may be an appropriate mechanism for covering many of the environmental impact costs, as well as the infrastructure use costs associated with road traffic, but only if there is no better, more direct charge available for each of the different externalities.

3. The difficulties of ensuring efficient channeling of revenues were exemplified in Mali in the late 1970s and early 1980s. Almost all of the revenue of the road fund was received through a postal checking service, the illiquidity of which prevented effective and timely finance of the routine road maintenance program. Most second-generation road funds are set up as a special account under an existing finance act. Money collected under the general taxing powers of the government is first paid into a consolidated fund and then transferred to the road fund. This arrangement works as long as it has the continuing support of the ministry of finance. Legislation under preparation in Ghana, Malawi, and Zambia and the arrangement already in operation in Yemen enable charges collected from users to be paid directly into the road fund.

4. To evaluate effectiveness, a case-specific set of indicators needs to be created to track changes for at least five years before and after the introduction of road funds. To facilitate comparison between countries, there should be a subset of indicators that are tracked for all road funds. The indicators should include the impact of road funds on general fiscal management (overall balances as well as allocations to other sectors) compared with their impact on revenues generated and spent on road maintenance. Some of the indicators should highlight the positive or negative impacts on incentives and accountability created by the governance structure (including the ability to resolve conflicts and withstand political pressures).

Table A-1. Road Expenditure Patterns in Selected Countries, 1988–95
(average annual values)

<table>
<thead>
<tr>
<th>Country</th>
<th>GNP per capita (1995 U.S. dollars)</th>
<th>Total government expenditure (% of GNP)</th>
<th>Total road expenditure (% of total government expenditures)</th>
<th>Road maintenance expenditure % of total</th>
<th>% of GNP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Countries with per capita income less than $8,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tanzania</td>
<td>174</td>
<td>43.24</td>
<td>8.67</td>
<td>13.73</td>
<td>0.51</td>
</tr>
<tr>
<td>Yemen, Rep.</td>
<td>292</td>
<td>27.16</td>
<td>16.36</td>
<td>4.54</td>
<td>0.20</td>
</tr>
<tr>
<td>Kenya</td>
<td>392</td>
<td>23.32</td>
<td>3.43</td>
<td>29.21</td>
<td>0.23</td>
</tr>
<tr>
<td>Pakistan</td>
<td>481</td>
<td>29.11</td>
<td>3.45</td>
<td>6.54</td>
<td>0.07</td>
</tr>
<tr>
<td>Sri Lanka</td>
<td>615</td>
<td>26.15</td>
<td>19.95</td>
<td>11.83</td>
<td>0.62</td>
</tr>
<tr>
<td>Bolivia</td>
<td>853</td>
<td>21.65</td>
<td>14.71</td>
<td>14.09</td>
<td>0.33</td>
</tr>
<tr>
<td>Ecuador</td>
<td>1,315</td>
<td>16.72</td>
<td>6.30</td>
<td>18.12</td>
<td>0.19</td>
</tr>
<tr>
<td>Thailand</td>
<td>2,084</td>
<td>27.08</td>
<td>2.77</td>
<td>28.73</td>
<td>0.22</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>2,326</td>
<td>27.08</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Table continues on the following page.)
<table>
<thead>
<tr>
<th>Country</th>
<th>GNP per capita (1995 U.S. dollars)</th>
<th>Total government expenditure (% of GNP)</th>
<th>Total road expenditure (% of total government expenditure)</th>
<th>Road maintenance expenditure</th>
<th>% of total</th>
<th>% of GNP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turkey</td>
<td>2,832</td>
<td>20.28</td>
<td>3.26</td>
<td>34.63</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>Brazil</td>
<td>3,152</td>
<td>51.70</td>
<td>0.38</td>
<td>30.05</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>South Africa</td>
<td>3,195</td>
<td>45.24</td>
<td>1.73</td>
<td>32.37</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>Hungary</td>
<td>3,686</td>
<td>70.93</td>
<td>0.89</td>
<td>29.07</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>Korea, Rep.</td>
<td>7,629</td>
<td>16.68</td>
<td>12.00</td>
<td>9.70</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>Average&lt;sup&gt;3&lt;/sup&gt;</td>
<td>2,160</td>
<td>31.80</td>
<td>7.22</td>
<td>21.03</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>Countries with per capita income $8,000 or greater</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Portugal</td>
<td>8,386</td>
<td>44.07</td>
<td>2.40</td>
<td>6.09</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>Ireland</td>
<td>13,429</td>
<td>61.98</td>
<td>2.10</td>
<td>34.60</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td>Spain</td>
<td>13,771</td>
<td>50.61</td>
<td>2.17</td>
<td>19.81</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td>United Kingdom</td>
<td>19,205</td>
<td>52.26</td>
<td>1.56</td>
<td>40.82</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>19,484</td>
<td>46.47</td>
<td>2.85</td>
<td>29.68</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>20,853</td>
<td>54.93</td>
<td>2.60</td>
<td>33.66</td>
<td>0.48</td>
<td></td>
</tr>
<tr>
<td>Netherlands</td>
<td>22,479</td>
<td>69.45</td>
<td>0.70</td>
<td>57.20</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>24,209</td>
<td>54.41</td>
<td>1.04</td>
<td>11.11</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>Austria</td>
<td>24,427</td>
<td>57.77</td>
<td>2.96</td>
<td>32.06</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>Finland</td>
<td>24,858</td>
<td>62.69</td>
<td>3.06</td>
<td>52.53</td>
<td>1.01</td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>26,543</td>
<td>55.80</td>
<td>1.79</td>
<td>11.51</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td>27,047</td>
<td>42.94</td>
<td>2.89</td>
<td>32.25</td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>27,659</td>
<td>72.00</td>
<td>1.31</td>
<td>66.28</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>28,236</td>
<td>77.90</td>
<td>1.16</td>
<td>52.59</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td>Norway</td>
<td>31,067</td>
<td>62.37</td>
<td>2.05</td>
<td>32.69</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td>Japan</td>
<td>33,761</td>
<td>19.01</td>
<td>10.54</td>
<td>14.14</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>Switzerland</td>
<td>39,760</td>
<td>44.45</td>
<td>1.82</td>
<td>18.45</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>Average&lt;sup&gt;4&lt;/sup&gt;</td>
<td>23,833</td>
<td>54.65</td>
<td>2.53</td>
<td>32.09</td>
<td>0.37</td>
<td></td>
</tr>
</tbody>
</table>

—Not available.

Note: Gross national product (GNP) per capita is based on the market exchange rate. Calculation of GNP per capita in terms of purchasing power parity shows that while values change for all the countries selected, resulting in a change in internal rankings, there is no shift in countries between the two groups.

a. Total government expenditure includes all expenditures at the central, state, and local government levels. Original data in local currencies are converted to dollars using the exchange rate data from the World Development Indicators database of the World Bank.

b. Total road expenditure includes all capital and maintenance expenditures on roads at the central, state, and local government levels. Original data in SDR (Special Drawing Rights) are converted to dollars using the exchange rate data from IMF, International Financial Statistics Yearbook 1997.

c. Road maintenance expenditure includes all maintenance expenditures on roads at the central, state, and local government levels converted to dollars using the exchange rate data from IMF, International Financial Statistics Yearbook 1997.

d. The average for each variable is computed as the sample mean for the group of selected countries.

Glossary

Earmarking refers to the precommitment of taxes to support or fully fund prespecified expenditure items. These revenues may be channeled through the general treasury or may be paid directly to a dedicated fund.

First-generation road funds were established in the 1960s and 1970s as extrabudgetary arrangements through which an earmarked stream of tax revenues was put at the disposal of a road department or agency.

Ringfences are theoretical enclosures established by tax legislation around certain transactions in order to isolate them. This mechanism is used to separate the income or loss from one project from the income or loss from other projects. For example, the United Kingdom uses such a provision to prevent oil companies from using tax losses and relief from activities on the mainland to reduce their taxable profits from North Sea oilfields.

Second-generation road funds moved away from using earmarked tax revenues in the 1990s. Instead, they are funded by levies or surcharges designated as “user charges” and identified separately from general taxation.

Special taxing district refers to limited special-purpose forms of government to which taxing powers are devolved. In contrast to earmarking a part of tax revenues, the allocation of taxing powers includes the ability to set tax rates in a regulated framework.

Taxes are public charges that are generally unrelated to the costs of production of a particular transaction. These charges generate revenue that is collected by one set of government departments (treasury, internal affairs, customs and excise, or energy) and distributed through the budgetary process to another set of government departments (transport, public works, local government) for spending purposes.

Tolls are direct charges for public services that function like prices (analogous to public utility tariffs) and that are retained by the collecting entity.

User charges (or quasi prices) are indirect charges for infrastructure services that are often levied as fees on proxy transactions. The choice of proxy varies with the type of infrastructure. The more indirect the relationship between the transaction subject to charges and the behavior to be influenced, whether of suppliers or consumers, the more the user charge functions as a tax rather than a price.

References

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


——. 1997b. IMF Summary of the Meeting on Road Funds, memo, January. Washington, D.C.


Between the State and the Market: Can Informal Insurance Patch the Safety Net?

Jonathan Morduch

Most households in low-income countries deal with economic hardships through informal insurance arrangements between individuals and communities rather than through publicly managed programs or market-provided insurance schemes. Households may, for example, draw on savings, sell physical assets, rely on reciprocal gift exchanges, or diversify into alternative income-generating activities. These mechanisms can be highly effective in the right circumstances, but most recent studies show that informal insurance arrangements are often weak. Poor households, in particular, have substantial difficulties coping with even local, idiosyncratic risks. Public policy can help reduce vulnerability by encouraging private, flexible coping mechanisms while discouraging those that are fragile or that hinder economic and social mobility. Promising policies include creating self-regulating workfare programs and providing a supportive setting for institutions working to improve access to credit, crop and health insurance, and safe and convenient saving opportunities.

Many low-income countries, from Sub-Saharan Africa to Southeast Asia, have suffered major natural disasters and political upheavals in the 1990s. These events remind observers of a reality hidden in official poverty statistics: that the condition of poverty is linked closely to vulnerability. Many poor households are exposed regularly to risks from illness, harsh weather, political instability, and economic mismanagement. Concern with vulnerability may be both intrinsic and tied to implications for income generation as well as to the longer-term consequences for the health and education of children (Jacob and Skoufias 1997; Hoddinott and Kinsey 1998; Rose 1999). Fear of risk can lead poor households to forgo potentially valuable new technologies and profitable production choices. Rosenzweig and Binswanger (1993), for example, use data from rural South India to show that an increase in risk (as measured by an increase of one standard deviation in the coefficient of variation of the

© 1999 The International Bank for Reconstruction and Development / THE WORLD BANK

187
date of the onset of the monsoon) leads to a 35 percent reduction in farm profits for the poorest quarter of households but has no effect on the wealthiest farmers. Vulnerable households may also spiral downward into poverty following adverse economic or climatic shocks as productive assets are depleted to protect consumption levels. Addressing risk can thus be an important complement to redistributive efforts and antipoverty strategies focused on increasing economic growth and employment.

Yet even with holes in both public safety nets and private insurance markets, poor households are not completely exposed to risk. Most have developed coping strategies to deal with the harshest blows. Most of these mechanisms are provided neither by the market nor by the state but instead are private "informal insurance" arrangements. They include individual and community actions, such as drawing down savings, selling of physical assets, reciprocal exchanges of gifts and loans, diversifying crops, and expanding income-generating activities. (For recent surveys, see Alderman and Paxson 1994; Besley 1995; Morduch 1995; and Haddad and Zeller 1996.) Some, like ritualized gift giving, have roots going back generations or even centuries, while others are newer responses to difficult situations (Mauss 1967).

Recent studies warn that some public policies may do little more than crowd out these informal mechanisms, but most evidence shows that crowding out is unlikely to be a substantial problem. Most informal insurance mechanisms are typically weak and often provide only inadequate protection to poor households. Studies from regions as diverse as rural India, China, and Sub-Saharan Africa suggest that despite informal insurance arrangements, households are exposed to considerable risk from adverse shocks—even idiosyncratic shocks that do not simultaneously affect their neighbors. Moreover, private informal mechanisms that are effective in reducing vulnerability can retard economic growth and social mobility. Thus, even where informal insurance is well developed, public actions that displace informal mechanisms can yield net benefits.

The emerging evidence suggests that policymakers need to be concerned about more than providing disaster relief in the wake of large, aggregate shocks such as floods, earthquakes, droughts, and other natural disasters. It is equally important to consider the needs of households that are facing losses due to adverse personal, economic, or other crises such as illness, poor (local) harvests, and temporary unemployment. Policy options include creating a supportive environment for institutions that offer safe and reliable means for poor households to borrow and, in particular, to save. Recent experience shows that it is also possible to offer limited life insurance and protection against other basic exigencies in a simple, low-cost manner. Much more speculatively, it may be possible to improve on existing insurance arrangements for poor households by drawing lessons from the emerging microfinance movement and by relying on nongovernmental organizations and profit-making commercial enterprises to take key roles. Public workfare programs that offer temporary employment at wages that are too low to attract those who already have work,
such as India’s Employment Guarantee Scheme, can also provide households with a flexible means for self-insurance in times of particular need.

Evidence on Risk Sharing

New tests of informal insurance mechanisms relate the variability of total household consumption to income variability. If households can use coping mechanisms to “smooth” consumption somewhat, income should be more variable than consumption. Sharper testable implications can be drawn with respect to perfect consumption-smoothing arrangements. If communities perfectly pool their incomes to share risks (and any given household’s income is a small part of the total), the consumption level of a given household relative to its neighbors should be a function only of total community income and the household’s assigned share of the total. The household’s own income realization should then not affect consumption patterns, and all idiosyncratic risk (relative to village shocks) should be eliminated.²

Townsend (1994) first tested the idea using data from the International Crop Research Institute for the Semi-Arid Tropics (ICRISAT) studies of villages in rural South India. He finds that the evidence does not fully support the proposition with respect to perfect risk pooling but that it comes surprisingly close: having controlled for community resources, Townsend finds that the marginal propensity to consume out of a household’s own income is nowhere greater than 0.14, while the theory of perfect risk sharing predicts it should be zero. Morduch (1991), Ligon, Thomas, and Worall (1997), and Ravallion and Chaudhuri (1997) find weaker evidence using the same data, however, with Ravallion and Chaudhuri’s estimates of the marginal propensity to consume falling between 0.12 and 0.46. These results suggest that informal insurance exists but that it is not nearly perfect. More critically, the studies do not reveal the specific mechanisms that drive the results: they are consistent with both gift exchange within communities (but not with perfect risk sharing) and with self-insurance activities such as borrowing and saving (but not with the perfect ability to smooth consumption). In practice, borrowing and saving are typically far more important coping mechanisms than the exchange of transfers (see, for example, Lim and Townsend 1998).

Similar studies from other countries also find evidence of highly imperfect informal insurance. Deaton (1997), for example, finds little evidence of strong risk sharing in samples from Côte d’Ivoire, and Townsend (1995a, b) reports a mixed record of risk sharing in a sample of Thai villages. Jalan and Ravallion (1997) note that the poorest 10 percent of households in rural China can protect themselves from just 60 percent of an adverse income shock—although the richest 10 percent can cope with 90 percent on average. This measure echoes my evidence from rural South India that households with large landholdings have little difficulty coping with idiosyncratic

Jonathan Morduch
income shocks but that the consumption levels of landless households and smallholders decline sharply as income falls (Morduch 1993).

Gertler and Gruber (1997) take a different cut on tests of risk sharing. In evaluating the ability to cope with the costs of illness in Indonesia, they find that households insure adequately against about 70 percent of common health shocks but can protect their consumption levels against only about 30 percent of the illnesses that seriously impair long-term performance. Their findings mirror Cochrane’s (1991) evidence using data from the United States and Lund and Fafchamps’s (1997) data from the rural Philippines. Lund and Fafchamps, in particular, find that informal insurance arrangements are effective only in the case of young adults who are acutely ill; older adults who fall ill are far less likely to be helped. Lund and Fafchamps’s data show that informal insurance also helps with funerals but not with crop failures, mild illnesses, or unemployment (other than that of the head of the family and his spouse). The household’s social network also matters: households with more friends (especially richer friends) have a greater ability to use informal insurance. Households that are not so well connected fare much worse.

These results, based on disaggregations by class and type of shock, reveal weaknesses in informal insurance that would not otherwise be evident. And they imply that there is ample scope for potentially beneficial interventions that go beyond disaster relief.

**Informal Insurance Mechanisms**

Reciprocal gift giving is a common way to solidify social and economic relationships—and one potentially important form of informal insurance. In North China, for example, gifts are given to mark births, deaths, and weddings, as well as to help the elderly, the ill, and women who have just given birth (Kipnis 1997). Anthropologists have tended to downplay gift giving as a product of a rational calculus associated with informal insurance systems, instead highlighting its role in securing social status and signaling commitment to the community (Malinowski 1922). But, given the sense of obligation engendered by gift exchange and the great potential for gift exchange to address risk, it is only natural that economists have taken the gift-insurance relationship seriously. Economists have thus tended to view gifts as they do other transfers such as public aid (Cox and Jimenez 1991).

Private transfers of cash, food, and clothing are large and frequent in some countries. For example, 40 percent of black South Africans reported either receiving or giving cash transfers (Cox and Jimenez 1997). Cash transfers reach large numbers of urban residents in Colombia (Cox and Jimenez 1998), Thailand (Paulson 1995), and the Philippines, where 82 percent of urban and 89 percent of rural households report receiving transfers (Cox and Jimenez 1995).
Remittances from migrants—whether migrants to another country or to the city from the country—can also be substantial (see Lucas and Stark 1985 on Botswana and Paulson 1995 on Thailand). Roughly two-thirds of all transfer inflows in Pakistan between 1985 and 1988 came from migrants who sent money home to their families (Foster and Rosenzweig 1999). In the Philippines, 26 percent of urban and 13 percent of rural households received remittances from migrant parents or children (Cox and Jimenez 1995).

But elsewhere, especially where migration is limited, reported transfers are of minor consequence. In a data set from India fewer than 400 of 4,000 households reported receiving net transfers in 1968–71 (Foster and Rosenzweig 1999). The lack of transfers is especially notable in contexts in which they are expected to be most valuable. In the smaller ICRISAT survey of poor villages in rural South India, Rosenzweig (1988) finds that transfers respond to risk but that they cover less than 10 percent of the typical shortfalls in income. Studies of exchange in Sub-Saharan Africa by Reardon, Matlon, and Delgado (1988) and Czukas, Fafchamps, and Udry (1998) reinforce this picture. Data from rural China are only slightly more optimistic. In a four-year study of 16 villages in the north of the country, Morduch and Sicular (1999) find that no more than a quarter of households report receipts of transfers from their neighbors; 10 percent report receiving gifts from outside their village. The transfers were modest, moreover, averaging about 10–20 percent of average household income.

Lim and Townsend (1998), who have examined ICRISAT files, conclude that households bridge the gap between income and desired consumption levels through saving, borrowing, and the use of buffer stocks of grain. The importance of these mechanisms is evident in most low-income economies (Deaton 1992; Alderman 1996). Households might also take actions to smooth income after an economic shock by working longer hours, for example, or taking an extra job (Kochar 1999). And they may take precautions beforehand to reduce the probability or extent of loss (Morduch 1995). These mechanisms can be relatively effective in the right circumstances, leading to concerns that publicly provided resources would simply replace—or crowd out—these informal activities.

Reexamining the Costs of “Crowding Out”

To the extent that informal mechanisms are limited, the concern about crowding out should be small. Still, some evidence is disturbing. Cox and Jimenez (1995) use household-level data from the urban Philippines to give a particularly striking example. They estimate that the receipt of net transfers from other households is particularly sensitive to whether the recipient is unemployed. The magnitude of the sensitivity is such that they conclude that if the government were to institute a simple
unemployment insurance scheme, net private transfers to the unemployed would fall by 92 pesos for every 100 pesos offered by the public program. In the end, the average unemployed worker would be better off by only 8 pesos.

Exercises based on extrapolations like this yield provocative—but not definitive—results. First, the research is not based on the effects of an actual program. Second, researchers typically have data from just one side of any given transfer: who makes it or who gets it—but not both. Without complete information, it is difficult to tease out the exact reasons for the transfers. A transfer that looks like informal insurance against a bout of illness, for example, might instead reflect a correlation arising for other reasons, such as a gift from a child to provide a parent with old age support. In the evidence from the Philippines, Cox and Jiminez (1995) concluded that the critical retirement income variables reflected the desire of givers to help retirees. Is that in fact the case? Or do the variables instead signal a type of household (one with little or no retirement income) that is more likely to include migrants—and thus to receive remittances—rather than one without migrants? In the latter case, the receipt of net transfers could look like informal social insurance but instead could simply reflect steps taken to maximize household income. Disentangling the explanations requires richer data.

Some of the sharpest evidence on crowding out is from South Africa. When the apartheid system was falling apart, the government extended basic pension benefits to black South Africans on terms similar to those that had been available to whites who had no private pensions. The program, which was fully implemented by early 1993, provided a state pension equal to about $3 a day to all women over age 60 and to all men over age 65, subject to a means test (Case and Deaton 1998). The means test excluded nearly all whites and only the richest blacks. In the past, blacks had had to rely mainly on their own means to cope with aging and with economic downturns, and for the most part the new benefits were not expected. How were traditional private mechanisms affected by the postapartheid public pension reform? Jensen (1998) estimates that for those households receiving private transfers, every publicly provided rand led to a reduction of 0.2–0.4 rand in private transfers to the elderly. Migration by children was also reduced slightly.

A similar degree of crowding out is predicted on the basis of studies of transfers to urban residents in Peru and the Philippines: a displacement of 17 percent and 37 percent, respectively, for each unit transferred as a retirement benefit (Cox and Jimenez 1995, 1998). Without the chance to evaluate the introduction of an actual new program (as in Jensen 1998), the predictions derive from estimates of the sensitivity of net transfer receipts to existing social security or private pension arrangements. Cox and Jimenez then use coefficients from this exercise to predict the consequences of the introduction of a broad state pension system.

Even with displacement rates as high as 20–40 percent, are the social losses proportionately high? Not necessarily, because leakage does not imply pure wastage.
The degree to which displacement undercuts policy objectives depends on the specific objectives of the program in question: Is it poverty reduction? Old-age support? Enhancing economic efficiency? Reducing vulnerability?

In the case of South Africa, older citizens received less in total than supposed, but others gained—and those gains were socially valuable. In richer contexts, private transfers tend to be from the old to the young. In a sample of white South Africans, for example, the net recipients were five-and-a-half years younger than the net givers (Cox and Jimenez 1997). But in poorer contexts, the reverse is most often true, largely because parents invest in children with the expectation that the children will support them in their old age. In line with that hypothesis, net recipients among black South Africans are on average eight years older than net givers. The displaced transfers thus tended to return to young households, many of which are as poor as older households, yielding little leakage as far as poverty reduction is concerned. In addition, keeping the funds in the hands of younger households is more likely to encourage investment in human capital accumulation and other productive activities. Second, public transfer systems may be more efficiently delivered than private transfers, yielding a net gain to society through displacement. For example, public transfer schemes may be able to pool resources more efficiently than local private arrangements (Cox and Jimenez 1997). Third, some displacement, even if it constitutes an unwanted leakage, may be a required cost of strengthening and widening the safety net to include particularly vulnerable households. In South Africa, for example, just under half of pension recipients do not receive private transfers at all (Jensen 1998). Putting the other arguments aside, tolerating some crowding out of the transfers received by half of the elderly black population can be seen as a cost of extending the safety net to the other half.

Tensions in Informal Insurance

Why is evidence of risk sharing so weak in places where it is expected to be strong? For example, consider the highly risk-prone semiarid tropics of South India, where two out of every ten years on average bring drought. Despite the importance of bad weather, most of the variation in measured household incomes over time is idiosyncratic to particular households. Morduch (1991) shows that 75–96 percent of the variance of the logarithm of household income remains after removing variation due to changes in average village income and average household income over the study period (1976–82). Some of this idiosyncratic, residual variation is surely measurement error, but even if half of it is error, substantial idiosyncratic variation still remains. As a result, within-village gift exchanges, designed so that no net redistribution takes place over the period, could in principle reduce the variability of household after-gift income by as much as 90 percent in one of the villages.
under study. In practice, reported transfers are not nearly big enough to do so (Rosenzweig 1988).

Tensions in Systems of Reciprocal Transfers

To explain why systems of reciprocal transfers are weak, researchers point first to problems in enforcing understandings. Household A will help household B today, with the expectation that B will eventually reciprocate. But what will keep B from reneging? If A and B are related by blood or marriage, altruism may hold them together. But without altruism or enforceable contracts, self-interest is needed to keep incentives in line. The repeated nature of the interaction over time allows for self-interested reciprocity. If A can credibly commit to end all future insurance relationships with B in the event that B reneges, B may well see fit to fulfill obligations. B’s decision will depend on whether the gain from reneging today is smaller than the flow of future benefits from continued participation.

As Coate and Ravallion (1993) suggest, however, the degree of effective insurance that is provided will adjust so as not to tip the balance toward reneging. In practice, tensions are heightened when both parties are down on their luck (during a drought, say) or when a partner’s luck is particularly bad. When an individual is pushed close to the subsistence constraint, holding onto whatever one has may be especially tempting, despite the agreement to share with others. As a result, reciprocal exchange tends to fall apart (or to offer less of a return) when insurance is most needed. In general, it works best when participants have a cushion against poverty. Consistent with the evidence from China, the Philippines, and South India, theory suggests that systems of reciprocal transfers will be more effective for slightly richer households and those in less dire contexts (Coate and Ravallion 1993; Ligon, Thomas, and Worall 1997; Kletzer and Wright 1998).

Moral hazard is also likely to limit group-based informal insurance, just as it undermines standard insurance markets. When insurers cannot adequately observe and enforce that insurees are taking all due precautions, incentives can be enhanced by providing only partial insurance coverage. This is one reason that informal insurance is most prevalent among relatives or neighbors in similar professions, that is, those with good flows of information.

Transfer-based systems can also run into trouble when households have opportunities to accumulate savings because they then have a degree of insurance free of obligation to neighbors and kin. Similarly, when incomes of participants grow at different rates, richer households tend to opt out rather than face the possibility of systematically redistributing to others. Richer participants find themselves giving relatively more than they get back on average, and at a point they will leave, either to form a new group with other richer households or to fend for themselves individually. Evidence of these sorts of bifurcations is given by Platteau and Abraham’s (1987)
study of reciprocity in fishing villages in Kerala, India, and by Lund and Fafchamps's (1997) study of risk-coping mechanisms in the rural Philippines. In a reverse example, Platteau (forthcoming) argues that one reason for low saving rates in Sub-Saharan Africa is that rural communities and families discourage saving in order to avoid eventual cleavages.

These tensions explain how common mechanisms can solidify economic and social barriers along ethnic, gender, generational, and class lines (Fafchamps 1992; La Ferrara 1997) and contribute to “poverty traps” (Hoff 1997; Platteau forthcoming). Fafchamps (1992) draws on African experiences to suggest that instead of leading to cleavages, reciprocal exchange may instead lead to voluntary patron-client relationships. Rather than being asked to give more than poorer households, relatively rich households may find themselves in a position to extract surpluses from poorer households. Rich households, with their stocks of wealth, can offer a great deal to poor, vulnerable households. But the poor may have to offer labor at concessional rates to obtain protection from their patrons in hard times. The terms of reciprocal exchange may thus greatly favor the rich, although the terms are to everyone’s absolute advantage. This seemingly feudal scenario may play out in subtle forms throughout poor economies.

Fafchamps’s model shows how informal insurance may adapt to particular economic conditions, but observers suggest that despite the ability to adapt, these reciprocal mechanisms have started falling apart in recent years in Africa. The blame is put on economic and political upheavals, reinforced by increasing mobility and urbanization. In principle, urbanization and the increasing ease of mobility can both help and hinder the functioning of informal insurance. The negative is straightforward: by moving away, households are able to “default” on their obligations to relatives and neighbors. This may explain why until recently migrants from Kenya and elsewhere often moved as a family. Now, prohibitive costs and risks make that less prevalent, and workers often move on their own, adding to the likelihood of defaulting on obligations to their ex-neighbors.

The positive aspects rely on continued links. Migration allows geographic diversification of incomes, increasing the value of reciprocal relationships. Paulson (1995) shows evidence from Thailand that some family members migrate partly to diversify the family’s “portfolio” of earnings sources; Lucas and Stark (1985) and Rosenzweig and Stark (1989) make similar claims based on data from Botswana and rural South India, respectively. Because links must remain unsevered for informal insurance to work, only insurance among family-based groups can typically survive mobility.

In a recent theoretical contribution, Banerjee and Newman (1998) embed these ideas in a more general model of structural change. They suggest that the lack of insurance mechanisms in urban areas can inhibit mobility from villages. In the village, a worker can count on some security through group-based insurance mechanisms but will have relatively low earning opportunities. The city offers greater earning opportunities but weaker insurance mechanisms. The result is that only the
relatively rich (who can cope better without group-based insurance) and the relatively poor (who never had much group-based insurance to start with) will migrate. People in the large middle segment of the population will stay put, even though it may be economically beneficial for them to break their ties with the village and join the modern sector. The presence of informal insurance in villages can then be a drag on economic development. Drawing on data from Indian villages, Das Gupta (1987) provides evidence along these lines. In parts of Sub-Saharan Africa the greatest problems tend to be caused by excessive rural-urban migration rather than by insufficient mobility, but even here the Banerjee-Newman model can still provide useful insights. The basic ideas can be applied to explain inefficient mobility between economic sectors, for example, rather than just inefficient geographic mobility.

The final set of tensions centers on the role of the family. The family has been hovering in the background in this discussion because people generally turn to their relatives first—and often again as a last resort—in times of need. On the one hand, the institution of the family, stretching over generations and bearing well-understood protocols, greatly facilitates informal insurance. Most important, information and enforcement problems are mitigated. On the other hand, the family tends to have a much more limited pool of resources on which to draw relative to the broader community.

The most important tensions arise when the demographic structure of households is shaped to meet the purposes of informal insurance. For example, the old age security theory suggests that children are produced partly to provide informal social security. In situations with overcrowding and in cases in which parents do not take into account the negative externalities imposed by their children (through congestion and environmental degradation, for instance), social welfare may be enhanced by shifting to alternative social security mechanisms (Dasgupta 1993; Anand and Morduch 1999). For example, establishing secure, convenient savings programs may allow households to reduce the number of children they have without undermining their ability to cope with less income in old age and can provide a second round of benefits to the community through reductions in negative population-related externalities.

A number of other insurance-demographic links lead to similar tensions, including social pressure to migrate and to select marriage partners in order to provide the family with insurance (Rosenzweig and Stark 1989; Paulson 1995); family “churning” (that is, the turnover of responsibility) as a response to the death of a head of household; and the practice of taking in a foster child. Child fostering is common in Côte d’Ivoire, Ghana, and Sierra Leone, although explanations differ. Bledsoe and Isiugo-Abanihe (1989) discuss insurance-related motives for fostering; Ainsworth (1996), however, in a survey from Côte d’Ivoire, finds that the need for labor is a more important factor.
Tensions in Other Forms of Informal Insurance

Other insurance mechanisms also tend to be least effective just when they are most needed. In principle, buying and selling assets provides a good hedge against idiosyncratic risks (at a minimum). Rosenzweig and Wolpin (1993), for example, find that buying and selling bullocks is an important consumption-smoothing device in semiarid India. But even there, the correlation between poor harvests and the price of bullocks (the covariation of risk) can raise problems. In fact, Lim and Townsend (1998) suggest that covariant shocks and the nature of bullock transactions instead add volatility to cash holdings rather than protecting them, and after carefully sifting through the same data, they uncover little evidence that is consistent with the Rosenzweig-Wolpin findings. The Lim-Townsend finding is consistent with the tensions that are introduced when risks covary. Thus, asset prices may fluctuate widely when every household wants to buy goods—or dump goods—at the same time. As a result, it may not be surprising that Czukas, Fafchamps, and Udry (1998) find that in Burkina Faso, selling livestock protected households against only 20–30 percent of the income shortfalls suffered as a result of a drought. (It is possible, however, that effective insurance would have been stronger had the drought been less widespread.)

In addition, informal mechanisms are typically weak against repeated shocks. Simulations by Deaton (1992) show that the efficacy of using buffer stocks or savings accounts to smooth consumption is conditioned largely on the degree to which bad shocks follow one another over time. When harsh conditions are likely to persist for several years in a row, households would have to have very large stores of assets to achieve adequate protection. This is one reason that the consequences of droughts and floods may be especially bad: because they frequently entail adverse environmental changes (runoff; desertification; poor soil conditioning), they play out even after the climate has returned to normal.

Even where mechanisms work well in a narrow sense, they may do so only at large long-run social costs. First, many mechanisms are inherently costly. In risk-prone areas of India, for example, households may sacrifice as much as 25 percent of average income to reduce exposure to shocks (Walker and Ryan 1990). In principle, then, improving safety nets can increase average incomes by reducing reliance on these costly measures (Platteau 1991; Morduch 1993, 1994). Perhaps more important, the desire to stay with tried and true technologies limits experimentation and innovation, creating ongoing problems for households.

Improving insurance may also mitigate social inequalities. Many informal insurance mechanisms have a gender dimension as well. Women often bear the brunt of arranged marriages, migration, and child fostering. Women may also lose out more than men during downturns. In India, for example, Rose (1999) finds that child mortality rates increase during periods of very low rainfall and are significantly higher
for girls than for boys. Reducing households' vulnerability and instituting more flexible insurance instruments may thus have broader social implications.

Policy Implications

A first set of policy priorities includes actions to reduce risk itself. For example, improving governance can sharply reduce the vulnerability of households to downturns resulting from economic mismanagement. Increasing macroeconomic stability, reining in inflation, securing property rights, improving transport and communications, and creating a stable political environment can go a long way toward reducing the frequency and size of downturns and creating a supportive environment to facilitate private risk-reducing activities. Similarly, risk can be reduced through public health campaigns for immunization and sanitation, civil works projects (dams, retaining walls, irrigation), and, in some cases, price stabilization. Higher incomes and stable employment opportunities further enhance the ability to cope with risk. But these are all policy areas that are on the table for other reasons and are best judged by other criteria.

In richer countries, households typically prepare for income declines by acquiring savings accounts, lines of credit, pensions, insurance, and annuities. Where these actions run into limits, governments typically provide means-tested poverty alleviation programs, unemployment benefits, health insurance, and social security (Subbarao and others 1997, ch. 3). But neither the administrative capacity nor the funding exists in most low-income countries to build similar public safety nets. Public action can, however, help to address smaller, local hardships by providing regulatory and institutional frameworks that expand households’ access to insurance, credit, employment opportunities, and convenient ways to save. Limiting the government’s role conserves scarce administrative resources and avoids potential conflicts of interest between short-term political exigencies and requirements for longer-term institutional sustainability. These policies provide ways to strengthen informal coping mechanisms and broaden their accessibility rather than displace private actions.

Promoting Savings

It had long been thought that most poor households had little desire to save in banks, but the experience of Indonesia’s Bank Rakyat Indonesia (BRI) and similar microfinance programs are turning that view around. After BRI established a safe, convenient savings vehicle, consumers responded enthusiastically. BRI now has more than 16 million low-income depositors (compared with 2 million borrowers), greatly aiding the bank’s profitability. Although there is no systematic evidence on the in-
come levels of depositors, bank staff note that they tend to be poorer on average than borrowers and from diverse socioeconomic backgrounds. Partly as a result, savings mobilization efforts are now being renewed in microfinance programs in Africa, Asia, and Latin America. Public policy can aid by ensuring an appropriate regulatory environment and helping to keep inflation in check.

One promising program has shown the surprising demand for savings deposits among poor households in the slums of Dhaka, Bangladesh. SafeSave, a nongovernmental organization, patterned its savings program on that of the local rotating saving and credit associations, which operate by taking in small sums from participants (Rutherford 1999). The response to SafeSave has been much greater than expected, and depositors have been able to slowly build up usefully large sums of money. As a consequence, the program is being replicated by other nongovernmental organizations in South Asia.

Without easy saving opportunities, households are tempted to squander surpluses or are susceptible to calls for short-term help from family members or neighbors—often at the expense of long-term progress (Platteau forthcoming). In this way, savings instruments may well be much more important than the provision of credit in raising incomes and reducing risk—and easier to accomplish.

Such financial deposits can be particularly effective in helping households weather the difficult scenarios that undermine gift exchanges. Consider the various forms of shocks that households encounter. Events that occur infrequently—old age, death in the family, and chronic disability—can hit households hard and may require a continuing flow of transfers. Given that such transfers are not guaranteed, savings deposits can be critical in ensuring that people have enough income to satisfy basic needs. (Note, however, that if interest rates fall below inflation rates, the purchasing power of these deposits can erode quickly.) Savings also allow households to avoid borrowing from moneylenders at interest rates as high as 5–10 percent a month when emergency funds are needed and can be especially valuable in a crisis (Von Pischke 1991). A regional drought, for example, will lead to a decline in the price of assets as affected individuals simultaneously try to sell their holdings. Financial savings, however, will generally hold greater value (and could increase in value as prices fall).

Public policy that leads to better integrated savings programs can help to contain risk, allowing the financial system to handle shocks more easily. The theoretical relationship of deposit mobilization, efficiency enhancement, and the generation of economic growth is described by Bencivenga and Smith (1991). The keys to a successful savings program are providing long-term security and convenience, finding a way to hedge against inflation, minimizing costs, and exploiting opportunities to relend deposits safely but profitably. Existing banks and nongovernmental organizations may not be up to all of these tasks, and designing effective (but not overly intrusive) prudential regulations is a critical first step.
Microcredit

Microcredit programs have succeeded by creating hybrid institutions that channel formal-sector funds to poor households. The programs are not perfect. In Africa, for example, the challenge is to create mechanisms that work well in semiarid and arid rural regions where households tend to have less diversified income bases and where low population densities mean higher transactions costs for financial institutions. More generally, most poverty-focused programs face high costs that undermine attempts at profitability. A recent survey shows that such programs are able to cover an average of only 70 percent of their total costs (MicroBanking Bulletin 1998).

The benefits, though, may be considerable. Using a recent survey of 1,800 households in Bangladesh, I find that access to microcredit programs yields no appreciable increase in average consumption levels in the short term (Morduch 1998). For those with access, however, the volatility of consumption over the three main cropping seasons is roughly half that of control groups (after controlling for unobservable variables at the village level). This reduction in consumption variability turns out to be mainly a product of reduced income variability across seasons, which is made possible by the employment diversification that credit affords. Helping rural households to reduce risk further by diversifying into nonfarm labor is an overlooked, but important, return to microcredit, and more research needs to be done along these lines to inform discussions of the costs and benefits of supporting credit-based approaches (Khandker 1998; Morduch forthcoming).

Insurance

Crop insurance programs have been a disaster nearly everywhere—not unlike targeted credit programs in the 1970s (Yaron, Benjamin, and Piprek 1997). Imperfect information and high transactions costs have proven to be destabilizing, and there are no easy solutions. Although reform currently looks unpromising, in principle the problems of insurance markets are not much more intractable than those of credit markets—and microfinance programs have shown effective ways around some of the largest hurdles there. Some microfinance programs have introduced insurance successfully on a limited scale, offering term life insurance at very low rates (with benefits large enough to clear debts and provide for a burial but not much more). In addition, the Grameen Bank of Bangladesh, for example, appears to have had success with its “emergency fund” for borrowers. The fund aids with loan repayment and provides general help in the event of illness and other emergencies. Information and transactions costs are reduced by coupling these mechanisms with credit provision. Experimentation will be necessary to determine whether these types of insurance mechanisms can be provided separately from other microfinance services.
As Morduch and Siculcar (1999) suggest, there may be ways to insure poor households by drawing broader lessons from microfinance. In a study of northern China, we describe an insurance company that has found some success in selling crop insurance to groups (a whole village, say), rather than to individuals. At the moment, this group insurance is used just to lower transactions costs for insurers and is a poor analogue to the group-lending practices used by microfinance institutions. But if future premiums were tied to the history of losses, a group-based contract could provide incentives for peer monitoring along the lines that microcredit programs have found successful in addressing moral hazard. This is an area open to speculation, and many roadblocks remain—for example, how should a program discourage collusion by the entire group?

One important lesson from microfinance is that programs operated directly by governments tend to have inherent difficulties in generating compliance by participants; borrowers are far more likely to default on loans from government sources, and governments are more likely to tolerate defaults in the name of political expediency. This has proved disastrous for the long-term sustainability of public credit programs. There is a parallel in the case of insurance; insurees appear less likely to take due precautions when governments are the insurers. Facilitating insurance provision by nonprofits, nongovernmental organizations, and for-profit companies may thus be an important step forward.

**Employment Guarantee Schemes**

Direct public interventions can also help to reduce vulnerability, especially for the poorest households. Among the most promising are rural public works programs such as India's Employment Guarantee Scheme in Maharashtra State, mentioned earlier, and the Food for Work program in Bangladesh, both of which are described by Ravallion (1991). These programs provide wage employment in return for work on constructing and maintaining public infrastructure. Ravallion (1991) reports that India's scheme provided about 100 million person-years of employment between 1975 and 1989. On average, 500,000 people participated per month (of a total of 20 million rural workers). Walker, Singh, and Asokan (1986) find that the coefficient of variation of income among landless laborers in two villages in Maharashtra with access to the employment scheme was half that of a similar village without access.

The work requirement and low wage rates provide a way to target the aid to truly needy households, allowing the programs to avoid instituting costly means tests. Participants take advantage of the program only when needed, often in lean seasons before harvests. During peak seasons, alternative employment opportunities are generally more attractive. Thus the programs avoid long-term dependency and ever-growing lists of participants. It also means that the programs are set up to
help households cope with temporary hardships, not mainly as an answer to chronic poverty.

The success of such programs will depend on government budgets, eligibility criteria, and wage rates. According to Ravallion (1991:171–72), there should be “as few restrictions on eligibility as feasible, and wage schedules and the rights of participants should be well defined, well known, and nondiscriminatory. Ideally, all who want work at the going wage rate should be able to get it.” The principle of inclusiveness is a key to reducing vulnerability because households (or at least those with available workers) are reassured by knowing that they will have a place to turn when they fall on hard times.

Conclusion

Poor households throughout the world face twin disadvantages. The first is difficulty in generating income. The second is vulnerability to economic, political, and physical downturns. Inflation, recession, drought, flood, illness, and civil war hit hardest those households that are least well equipped to handle the shocks. Harder still, the two disadvantages reinforce each other. Poverty is a source of vulnerability, and repeated exposure to downturns reinforces poverty.

The circular nature of poverty and vulnerability does not, however, preclude effective responses. The evidence to date suggests several broad directions to pursue. In addition to helping households cope with large natural disasters, governments need to encourage flexible private interventions. Concern has arisen about whether public action will crowd out private informal insurance mechanisms. To the contrary, well-designed public action can strengthen and broaden the capacity of households to act independently through informal mechanisms.

Making saving safer and more convenient, helping to expand credit access, and fostering basic insurance programs are all promising ways to help households help themselves in the face of adversity. The possibility of crowding out existing informal arrangements should not be ignored, but in most low-income countries, it is unlikely to substantially undermine steps to help poor households. First, informal insurance is often very limited, and second, the crowding out of some private actions can have valuable social benefits.

Notes

Jonathan Morduch is a MacArthur Foundation Research Fellow at Princeton University. The author is grateful for comments from Harold Alderman, Angus Deaton, and the journal’s referees. He has also benefited from discussions with Marcel Fafchamps, Trina Haque, John Hoddinott, Emmanuel
Jimenez, Helena Ribe, and participants in a seminar at the World Bank in July 1997. This work was supported by the Africa Department of the World Bank. The first draft was completed while a National Fellow at the Hoover Institution, Stanford University, and the revision was completed with support from the John D. and Catherine T. MacArthur Foundation.

1. Economists' calculations of the welfare costs of vulnerability typically lead to relatively small numbers for the benefits of risk reduction (see, for example, Newbery and Stiglitz 1981). But consumer riots in the face of price swings, evidence on the lengths to which households and governments go to avoid volatility, and participatory assessments of poverty (such as World Bank 1994) suggest that approaches grounded in simple microeconomic theory have been too narrow to capture the full extent of concern.

2. The specific tests of consumption insurance are rooted in the theory of optimal social allocations, and they begin with the assumption that each household has a social planner’s problem. If household A's utility from consumption is \( u(C_A) \) and household B's utility is \( u(C_B) \), then marginal utilities are simply \( u'(C_A) \) and \( u'(C_B) \). In the benchmark theory without enforcement or information problems, transfers should be made in every period so that \( \theta_A u'(C_A) = \theta_B u'(C_B) \). That is, redistribution from household A to B or from B to A should occur until there is no possible way to increase the weighted sum of their utilities. The burden of idiosyncratic shocks should thus be shared by the participants. And, in principle, the relationship should also hold for households A and C, A and D, A and E, B and C, and so forth. With data in two periods, the relationships continue to hold, so that \( \theta_A u'(C_{A1}) = \theta_B u'(C_{B1}) \) and \( \theta_A u'(C_{A2}) = \theta_B u'(C_{B2}) \). Putting the two relationships together shows that marginal utilities must grow over time at the same rate for all households: \( u'(C_{A1})/u'(C_{A2}) = u'(C_{B1})/u'(C_{B2}) \). Under assumptions commonly made about the shape of utility functions—for example, \( u(c) = c^{1/(1-p)} \), where \( p > 0 \) is the coefficient of relative risk aversion—the relationship also holds for the growth of consumption itself (or its logarithm). This relationship has provided the basis for a test of the basic theory on risk sharing. If the proposition and the assumptions about the form of preferences are correct, once the consumption growth of any single household (or of a region in aggregate) is known, the consumption growth of everyone should be known. Moreover, no other variables (such as income or income growth) should have influence. Given the assumptions, the test boils down to whether or not coefficients on income and income growth are statistically significant in explaining patterns of household-level consumption once regional consumption aggregates are controlled for. If the proposition holds exactly, the marginal propensity to consume out of idiosyncratic income changes should be zero. Townsend (1994) and Cochrane (1991) give a more explicit presentation. The presentation here implicitly assumes that households have identical preferences for consumption over time, that consumption and leisure are separable in utility, that utility is additively separable over time, and that utility is a function only of consumption levels. If instead utility also depends on household characteristics (a reasonable view), the testable implication is that consumption growth depends only on preference parameters, not on budget parameters—but this is a harder proposition to test.

3. How and where deposits are invested appears to be far less important than that deposits are mobilized from poor households. In principle, there is no reason not to invest the money abroad, for example, if domestic options prove difficult and returns are unattractive. Where it is costly to set up savings bank branches, simple mechanisms such as post office savings plans (or innovations based on the African susu collector) may offer appealing options.

4. Reardon, Matlon, and Delgado (1988), for example, show that nonfarm income accounted for 30–40 percent of total income in drought-affected Burkina Faso and that it was only imperfectly correlated with crop income, providing protection against the drought. Some of the nonfarm income, though, may come as an ex post response to downturns in farm income, leading to negative correlations. Czukas, Fafchamps, and Udry (1998), however, find that nonfarm income was positively correlated with farm income during a similar circumstance in Burkina Faso, so that nonfarm income was not an important offset to crop income.
References

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


Using Household Surveys to Build Analytic Capacity

Lorraine Blank • Margaret Grosh

This article reviews the results of efforts in five countries to build a national capacity to analyze social policy; these efforts were undertaken in conjunction with Living Standards Measurement Survey (LSMS) projects. Case studies for Bolivia, Jamaica, the Kyrgyz Republic, South Africa, and Vietnam show that building such capacity requires explicit planning, significant time and money, open access to data, and strong support from policymakers. Lessons are drawn about four aspects of building capacity—training, technical assistance, research, and recruitment. The lessons learned from these LSMS projects apply in a more general sense to other kinds of capacity-building projects.

Strengthening the effectiveness, efficiency, and responsiveness of government is a priority for developing countries as they prepare for the 21st century. The ability to monitor the impact of economic and social reforms, evaluate the outcomes of public programs and projects, and undertake policy-relevant research is an important component of good governance, and together these institutional capabilities constitute what is referred to as the capacity for policy analysis.

Policy analysis uses multiple methods of inquiry to examine the quantifiable outputs of public policies to determine if the goals and objectives of a particular policy or program are being achieved (Waterman and Wood 1993; Dunn 1994). Such analyses are an important input into the process of formulating and evaluating government policies. The field is grounded in economics but draws upon other disciplines as well. The techniques used include quantitative modeling, statistical analysis, econometrics, qualitative research, and political and institutional analysis. This article summarizes the results of efforts in five countries to build a national capacity to analyze social policies covering welfare, education, health, fertility, nutrition, labor, and employment.

Because national capacity to undertake such operations is weak in most developing countries, the current practice is to rely largely on foreign researchers (contracted
by the government through externally funded projects or directly by international development agencies. We believe that a strong case can be made for generating good policy analysis locally, whether commissioned by the government or conducted by in-country analysts. Commissioned analysis is done at the time and on the topic that the government requires, and local researchers who know the environment are more likely to take into account domestic institutional issues and may be more knowledgeable about the full range of data that could be brought to bear on an issue. True integration of the results into the policymaking process requires a sense of ownership of the data and analysis that comes only with being part of the process from beginning to end. Local researchers may be better placed to disseminate the results and provide follow-up advice on their studies. Involving national policymakers and analysts in the process of setting the agenda for policy analysis and organizing the data collection to support it ensures that local needs are met. This means that building local analytic capacity should produce more effective social policies.

Several factors have contributed to the limited analytic capacity in many developing countries. Problems with the quality of the data are extensive. Where adequate data have been collected, opportunities for their use are often missed because researchers are given limited—or no—access. And countries often have very few researchers, especially those connected with established universities and trained in advanced statistical techniques and research methodologies. Although access to computers and software may be adequate, there is likely to be little in-country experience with statistical packages. Graduate training in policy analysis is either nonexistent or lacking in rigor. Finally, but most significantly, there is no culture of using social data to address policy questions. Policymakers have little experience posing policy questions in ways that are amenable to analysis; they also are frequently unwilling to provide open access to information. As a result government officials seldom demand policy analysis (Peterson 1991; Simonpietri and Ngong 1995; Synergie 1996; Thorbecke 1996). These factors are related in a vicious circle. If data are not collected or not released to analysts, aspiring students will not hone their skills in quantitative analysis but will move into other areas of their disciplines. If policymakers do not have access to good policy studies, they do not learn to integrate them into the policy process, and they are not interested in funding data collection efforts.

To address these limitations, international development agencies, including the United Nations, the U.S. Agency for International Development (USAID), and the World Bank, have provided financial support for household and regional surveys and other efforts to collect quality data by providing technical assistance, training, and computer hardware and software to policy analysis units (both independent and in government), research centers, and training programs. Other efforts to promote capacity have involved using local researchers in externally funded research projects and training programs (Paul, Steedman, and Sutton 1989; Grindle and Hilderbrand 1995; Myers 1997), establishing regional training programs and peer review net-
works for researchers (World Bank 1990; Fine 1995), and developing evaluation systems and a demand for evaluation (World Bank 1994; Mackay 1998).

Using Living Standards Measurement Surveys to Build Analytic Capacity

The World Bank’s efforts to strengthen analytic capacity in conjunction with LSMS survey projects date from 1985 (see box 1). The skills required to implement these and other large-scale surveys—and the statistical, econometric, and modeling techniques involved—are much the same as those used to analyze social sector and economic data, so the lessons learned apply in a general sense to other kinds of capacity-building efforts as well.

Initially, LSMS survey projects focused almost exclusively on data collection; little attention was paid to building analytic capacity. In the last few years, however, task managers have begun to incorporate activities to build analytic capacity into the design of LSMS projects. Some projects do not include capacity building as central elements, but we now know more about what small efforts can do and how to craft them and are beginning to have a clearer understanding of the factors and sequencing required for the larger efforts. In this article, we focus on analyzing social sector policies, and our primary concerns are how to remedy the scarcity of policy analysts and how to encourage the use of analysis in policymaking.

The Experience in Five Countries

The case study countries—the Kyrgyz Republic, South Africa, Vietnam, Bolivia, and Jamaica—are geographically dispersed, economically and politically varied, and of different sizes. Most important, they represent a wide range of levels of capacity and different degrees of involvement in capacity building. The first Kyrgyz project

---

Box 1. *Living Standards Measurement Surveys*

The World Bank has been sponsoring LSMS surveys in developing countries since 1985, often with partial or full financing from other development agencies. These multitopic household surveys, which sample 2,000 to 5,000 households, are designed to monitor welfare and analyze behavior. The questionnaires always include comprehensive measures of consumption, usually income from employment, self-employment in agriculture or household enterprises, and transfers; access to and use of social services; and a range of outcomes and behaviors associated with health, fertility, nutrition, education, migration, savings, and housing. Community questionnaires gather information on the local economic environment, the availability of services, and prices of basic goods. See Grosh and Glewwe (1998) or the LSMS Home Page, [http://worldbank.org/lsmshome.html](http://worldbank.org/lsmshome.html), for a more detailed description.

Lorraine Blank and Margaret Grosh
was not designed to build analytic capacity at all but is typical of projects where a single analytic imperative drives a survey project. The first Vietnam project was similarly focused on data collection but thoughtfully used small amounts of money for capacity building. In South Africa another single-survey effort did involve a large number of potential future analysts in the survey design and dissemination plan. In contrast, Bolivia and Jamaica were multiyear, multimillion-dollar efforts designed to build analytic capacity.

**Kyrgyz Republic**

The first Kyrgyz Multi-Purpose Survey (KMPs) was implemented in 1993 to help the government and other local institutions assess the poverty and employment situation in order to design a program of targeted social assistance.

The goal was to collect national data as quickly as possible. No budget was specifically allotted for building analytic capacity.

At the time, the country's existing capacity was limited to implementing nonrandom household surveys that were designed and subsequently analyzed in Moscow. Local researchers were not trained in quantitative analysis and had no experience in analyzing large data sets, either in government or in academic institutions. Computers were used, and programmers were available but were unfamiliar with statistical programs. Policymakers did not know how to use the survey to inform policy decisions. Driven by the need to design a social safety net project as quickly as possible, the survey project did not include building local capacity for data collection or analysis. The construction of a sampling frame would, however, provide the sampling basis for future collection of survey data.

The results of the analysis were presented at a seminar in September 1994, attended by 50 senior officials from government ministries, trade unions, and universities. Participants were invited to request specific tabulations related to their individual spheres of interest, and these were prepared and distributed the following day. The presentation of these customized tabulations generated considerable excitement among the attendees, who recognized the potential contribution of such analysis to policy decisions. This experience helped pave the way for a second, multiyear project with explicit goals for building analytic capacity. This project funded four rounds of the survey—two in 1996, one in 1997, and one in 1998. A formal user committee representing the social sector, economic ministries, and Goskomstat (the Kyrgyz statistical agency) reviewed the content of each questionnaire. The Research Triangle Institute, an independent nonprofit research and development organization headquartered in North Carolina, provided technical assistance, including training in data collection and analysis. Experts worked with staff from Goskomstat and government ministries to produce poverty profiles and statistical abstracts. Selected staff from these agencies participated in short in-country
and overseas training programs in survey research, data processing and analysis, construction of consumption aggregates, poverty measures, anthropometric measurement, and sampling theory. The project cost $3.7 million, of which $545,400 was for training.

Was the program successful? Although local staff participated in the design of each survey, their contribution was limited by their lack of experience with the uses of survey data. As anticipated, they were more active in the production of the 1996, 1997, and 1998 poverty assessments and helped to develop pension models based on the 1996 survey.

Two factors have limited the development of Kyrgyz analytic capacity: first, the statistical methods that had been in place for 70 years were considered perfectly adequate by policymakers, and second, Kyrgyz analysts were suspicious of the household survey because it had been designed, processed, and analyzed in a foreign country. Despite the declaration of open access to the data, researchers were not familiar enough with the new instrument to forsake traditional data sources. Thus it was hard to get people to demand training to use that data. Moreover, policymakers have not requested much information collected from the surveys. Although efforts to stimulate demand were made immediately after the first survey, they have had little lasting impact. Observation missions to other countries are scheduled, but these efforts are coming late in the project. Finally, it is generally agreed that in the rush to implement and analyze the surveys, technical assistance has been directed more at completing analytic tasks than at transferring analytic skills—a common phenomenon in survey projects.

South Africa

This 1992 survey was undertaken to create a credible national database on poverty for use in monitoring poverty, analyzing policy, and designing programs. A related objective was to strengthen the capacity for these tasks among individuals who were likely to join a new majority-rule government. The African National Congress and the Congress of South African Trade Unions requested the project, which was conducted outside of government. The Labor and Development Research Unit at the University of Cape Town coordinated the project. Originally envisaged as a two-phase project, the first step focused on upgrading and filling the gaps in national poverty information and the second phase on strengthening the capacity for continuous poverty monitoring. In the end, only the first part was implemented. The project cost about $1.5 million, of which approximately $500,000 went to activities most closely linked with building analytic capacity, including workshops, regional studies, and publication of the survey abstract.

The survey questionnaire was designed in a series of workshops attended by more than 30 social scientists over a one-year period. Most of the fieldwork was carried out
in the latter part of 1993. The data were available for the new majority-rule government in 1994.

Although there are 22 universities in South Africa, and many researchers have been educated abroad, few individuals have been trained in quantitative analysis. As a result, there were few analysts, either in government or in academic institutions, who were able to analyze household data sets. This was especially true among black South Africans. Three workshops were held to train researchers in processing and analysis, and training in reading the data was provided to staff members at several government ministries in 1995 and on a demand (and fee-paying) basis after that.

Access to data was completely open, and a very active data dissemination program was implemented. Within a year after the survey was completed, data sets had been distributed to 13 government agencies, 10 universities (7 of the 10 were traditionally white), 9 independent research institutes, international aid agencies, and various local and international individuals. By 1996 more than 36 research reports had been produced from the survey. The findings were used to inform government decisions regarding unemployment programs for women and pension programs for the elderly; a white paper on water and sanitation; and a report on key indicators of poverty. And the data were being used in university courses to strengthen interest and skills in policy analysis.

As a result, the country's capacity to implement a national household survey improved in a "learning by doing" framework with technical assistance from World Bank staff in questionnaire design, sampling techniques, fieldwork, data processing, and preparation of the statistical abstract. The capacity to analyze the data for basic information such as that contained in the statistical abstract was enhanced through a process that began with debating many versions of the questionnaire in many regions of the country. This experience provided an orientation for social scientists who had not previously worked with data from an integrated household survey, and it supported the Central Statistics Office's efforts to undertake household surveys that more or less duplicate the objectives of the LSMS survey.

Policy and research analysts in South Africa began talking about household data sets and the analytic capacity necessary to use them. As a result capacity-building efforts have been included in a three-year Mellon Foundation project that came on stream in 1998. The goal of the project is to increase the number of social scientists, policy analysts, and others with the skills needed to design, implement, and analyze data. The project will fund graduate work in demography at the University of Cape Town, including long- and short-term training in quantitative social science methods as well as the West Cape Area Study, which will be implemented with technical assistance from the University of Michigan. The area study will be used as a training tool for graduate students as well as a mechanism to generate reliable data. Total project costs will amount to $330,000.
The challenge is to respond to the demand for increased quantitative analysis by helping the government establish a more permanent system for monitoring poverty and to increase the capacity for policy research and policy analysis in relevant government agencies and in the less established and primarily black universities.

**Vietnam**

The first Vietnam Living Standards Survey was conducted in 1992–93 with funding from the United Nations Development Programme (UNDP). The goal was to generate a comprehensive data set to help guide the policy decisions being taken as part of the transition to a market-based economy. The project was jointly managed by the State Planning Committee and the General Statistics Office. The costs of implementing the survey were $516,000, including $80,000 for training.

At the time, the country's analytic capacity was limited to basic statistics, with little expertise in advanced analytic techniques. Where expertise existed, reliable raw data and computer hardware and software were absent. Capacity-building strategies in the first project focused on preparing and disseminating the published abstract; a study mission to Thailand to observe survey procedures and to look at how the data were utilized; and short-term training (two to three weeks) in the use of statistical software, the structure and contents of the data set, and policy analysis. The project supported learning by doing through preparation of the survey abstract and working papers based on the data. The Ford Foundation funded a six-week workshop in policy analysis, statistics, computer training, preparation of research papers, and an orientation to the data, followed by a two-week writing workshop on preparing research papers. The Ford Foundation provided participants in this workshop with a computer powerful enough to use the survey data, a copy of a statistical software package, and a copy of the survey data. Project-funded training targeted policy analysts and technical staff from the central government, including the General Statistics Office, the State Planning Committee, and social sector ministries; Ford Foundation training targeted researchers from universities and research centers outside central government. The Ford Foundation provided $70,000 for training and equipment and $50,000 to sponsor local analysis.

The first survey resulted in some positive changes in the policy analysis environment. To understand the extent of these changes, it is important to remember that the survey broke new ground in Vietnam in several ways—it was the first household survey with all the unit record data on computer, it was the first survey that included training programs for ministry staff, and it was the first time that the statistical package had been used in Vietnam. That the statistical abstract was produced, largely by Vietnamese analysts, just six months after the data were collected was a significant achievement.

The initial survey and surrounding activities created demand for further training in data analysis. A second survey was fielded in 1997–98 as part of a project to...
develop capacity to collect, analyze, and disseminate social statistics data. The project is funded by the UNDP and the Swedish International Development Agency (SIDA) with total project costs of $1,782,400, of which $700,000 is earmarked for training. The project is funding study tours, short-term in-country training in statistics and data analysis for participants from 22 agencies, and short-term overseas training in basic and advanced statistics, survey design and planning, sampling theory, and poverty measurement.

A wide range of international researchers and donor agencies have requested the data from the first survey, but the Vietnamese government itself has used the data only minimally. That is partly because analytic capacity is still limited but primarily because there is as yet little demand for data-driven policy analysis. Participants in training programs are rarely called upon to use their newly acquired skills once they return to their jobs.

Training has been extensive in terms of the number of institutions reached, but only a limited number of individuals within each organization is involved, and the training is fairly superficial. In the absence of long-range planning geared to developing the skills required to implement and analyze household surveys, strategies to build these skills progressively have been lacking. Consequently, training has been offered to as many participants as possible instead of being provided in-depth to a more limited number of participants. The workshops give trainees the basic technical skills necessary for descriptive analysis, but the skills required to formulate research questions and prepare research reports have yet to be acquired. The challenges for the future are to provide long-term training to participants who have been selected on the basis of ability, rather than seniority, and to develop a strategy for ensuring an ongoing supply of policy analysts.

Bolivia

The first LSMS-type survey was conducted in Bolivia in 1989. Five annual surveys followed, at a total cost of $2.6 million, including $1.5 million for training. Training was directed primarily at implementing and processing the surveys, and many of the changes that were adopted were applied in the government's subsequent household surveys. Despite the existence of the LSMS survey data and data from other sources, there was little interest in social policy analysis. Bolivian universities did not have a tradition of quantitative policy research. Similarly, there was very little analytic capacity in the social sector ministries; for example, researchers who intended to analyze Statistical Institute data would ask for tables rather than the raw data. External consultants conducted most of the studies that were done on household survey data.

By the late 1980s, however, as social policy issues rose to prominence on the national political agenda and the work programs of the international development
agencies, the gaps in in-country analytic capacity became obvious. As a result USAID undertook a $2.4 million project to build up the country’s domestic capacity through the establishment of a social policy analysis unit known as UDAPSO. The government contributed to operating costs with funding from other USAID-supported programs. SIDA paid for five staff positions through the Bolivian civil service program. The group included a multidisciplinary team of up to 12 professionals and concentrated on education, health, poverty, income distribution, and social sector expenditures.

The Harvard Institute for International Development (HIID) provided 32 months of long-term technical assistance and almost 40 months of service of short-term consultants. These consultants advised on best practices, interpretation of results, and future work and provided feedback on the quality and relevance of the work. In-house training (in the form of two-week workshops on the use of large data sets and statistical software, basic and intermediate statistics, econometrics, and other analytic tools) was also provided. Staff participated in workshops at Harvard in more general areas such as poverty measurement, education planning, and health planning. On-the-job training occurred through participation in research projects and preparation of research reports. The World Bank also provided a month of training in econometrics.

Between 1992 and 1995, these newly trained analysts produced 5 books, 2 research monographs, and 36 working documents. LSMS data were essential to a great deal of this research. The work was well received by other government agencies and international donors. Staff members participated in policy dialogues with senior government officials, and the dissemination of research allowed analysts to develop collaborative alliances with multilateral organizations.

The policy analysis unit’s early success resulted from a combination of intangible factors, including political support from the minister of planning, leadership, and esprit de corp. When the minister and director of the unit changed, most of the professional staff left, and recruiting difficulties plagued the organization. It was disbanded in 1997. Despite this setback, the government did not abandon its capacity-building commitment but instead created a division for social policy within the Ministry of Planning based on many of the former unit’s functions and staff. This move may actually increase the impact of previous capacity-building efforts because the new division is closely linked to the Ministry of Finance, providing it with the technical analysis that underlies budget decisions. As a result, it has achieved legitimacy among the sector ministries.

To ensure a continuing supply of trained policy analysts, USAID funded a project in 1994 to establish a new master’s degree program in public policy and management at the Catholic University of Bolivia. HIID has provided technical assistance for work with faculty and staff in managing the program, teaching courses for graduate students and for public and private sector managers, and acquiring library and teaching materials.
Jamaica

The Survey of Living Conditions (SLC) was designed to collect data to monitor the social impacts of public policies and the delivery of social services. The survey was first implemented in 1988 and has been carried out annually since then. By 1992 the SLC Jamaica’s capacity for data collection was carried out in a smooth and orderly fashion and was providing an exceptionally rich database. In-country analysis, however, was limited to production of the annual abstracts and some studies (of varying quality) produced by local academics; few in-country researchers were capable of undertaking sophisticated analysis.

With the assistance of the World Bank and funding from the Netherlands, a new project was developed to increase the analytic capacity of the Planning Institute, the sector ministries, and the University of the West Indies. The $3.4 million project came on stream in July 1993. About $2.8 million was devoted to building analytic capacity and the rest to conducting further rounds of the SLC. The project established a social policy analysis unit in the Planning Institute; outlined a program of research; and provided technical assistance for graduate courses in statistical analysis, research methodology, and policy analysis for senior researchers in the sector ministries and for faculty at the university, as well as graduate school curriculum development. The project set up a social indicators data bank; graduate fellowships; workshops in statistics, research methods, techniques of policy analysis, and computer processing for junior research staff in the sector ministries; and seminars to introduce policymakers to the uses of data in policymaking. It also funded computer equipment and technical assistance to the Ministry of Labor and Welfare (the only ministry not receiving assistance from other donors) and support for five rounds of the SLC.

The social policy analysis unit, known as the Policy Development Unit, was established and operational for the whole life of the project and serves as an analytic and a project implementation unit. The unit has produced policy studies on poverty and has developed a methodology for tracking poverty over time. Unit staff revised and updated the poverty line studies that the government had contracted out in 1989, and they played an important role in preparing the government’s poverty strategy. The SLC abstract is now produced solely by in-country analysts. The timeliness and complexity of the abstract have increased since the project’s inception.

When the Policy Development Unit was first established, staff salaries were competitive with those in the public sector. It was difficult to find qualified individuals to fill the analyst posts because the pool of trained individuals was extremely small. Salaries have eroded since then, relative to government pay scales, and are no longer competitive. The turnover of professional staff has been high. Although a significant number of trained individuals is available (as a result of the project), recruiting has been difficult. This is a reflection not only of budget constraints but also of the diminished political support for the unit.
The University of Toronto provided technical assistance to upgrade graduate education in social policy analysis at the University of the West Indies. Ten-week courses provided training in statistics, multivariate analysis, research techniques, cost-benefit and effectiveness analysis, and other techniques to graduate students, faculty, and senior researchers from the sector ministries and the Planning Institute. The series ran twice. Eighteen fellowships were awarded, and a much larger cadre of graduate students participated in the courses. The courses have been made part of the permanent curriculum, but their continued delivery is not guaranteed because none of the current faculty has the required skills to teach the quantitative courses. The university is making a serious recruitment effort, but its success is not assured.

A social indicators data bank, which houses a computer laboratory and a library of relevant social sector data sets, including SLC data, was established at the university, which now covers the recurrent costs of the data bank. The data bank is encouraging the integration of hands-on data analysis in a wider range of classes than heretofore possible and is thereby encouraging quantitative work in university training broadly.

A series of in-service workshops (averaging 30 contact hours) was held for junior staff in the line ministries on the subjects of statistics, computer literacy, research methodology, and other techniques of policy analysis. The goal was to improve the ability of these staff members to report basic statistics on services delivered in their ministry. Enrollment was large, but few people took more than one course, and it is not clear that they use the training on their jobs.

The SLC has continued to be conducted every year, with decreasing reliance on technical assistance. Recruitment of qualified computer analysts has been a problem, partly because of low public sector wages and partly because very few people are qualified to work with statistical software. Seminars for policymakers were scheduled to integrate policy research into decisionmaking but were never implemented. This diminished policymakers' interest in obtaining data-driven answers to policy questions and in ensuring staff involvement in various other parts of the project.

The project was designed with funds to support an annual agenda of social research. The initial call for proposals was issued in the first year of the project. Two years after the start of the project, no research had been funded. Twelve proposals were approved in the last three years of the project, but several of these are of more modest sophistication than originally hoped for. The poor showing on this component may reflect some shortcomings in the details of how the call for proposals and the review process were done. More likely, it indicates the difficulty (exacerbated by the failure to implement the policymaker seminars) that policymakers and government analysts have in posing policy-relevant research questions and understanding how these could be answered with quantitative data. Another factor was probably the limited experience with proposal preparation among researchers (which might have been addressed earlier in the university training courses).
In the 11 years since the first SLC, the capacity for monitoring poverty and the amount and quality of quantitative policy analysis done in Jamaica have gradually increased. As a result the nature of the public policy debate about social issues has changed somewhat. Capacity has yet to reach the level of sophistication originally envisioned, however. The government has expressed its interest in continuing to strengthen policy analysis capacity, but external funding will be required.

The Jamaican experience demonstrates that local capacity to implement an LSMS survey can be strengthened and that local analytic capacity can be built, but that it is a slow and incremental process. The challenge is to ensure that these gains are sustainable, including the ability to produce a continuous supply of trained analysts and ongoing demand for policy analysis by decisionmakers.

Lessons Learned

Experience in the five countries selected indicates that integrating efforts to build analytic capacity with data collection efforts provides an immediate, relevant, and useful data source that can be used as a learning device for analysts and policymakers. Although some of these lessons may appear obvious, experience shows that they have often been ignored and are thus worth highlighting. We break the lessons into two groups: four requirements for building capacity—explicit planning, resources, an open data access policy, and demand for policy analysis—and four tools for building capacity—training, technical assistance, participation in research activities, and the establishment of policy analysis institutes.

Requirement One: Explicit Planning

Real and sustainable capacity building is unlikely to occur in the absence of clearly articulated goals and strategies. Clarity is required from the outset regarding the extent of existing in-country policy analysis capability, the types of capacity to be built, and the plan for doing so. Planning for capacity building requires that the project planners make decisions about:

- The agency or agencies to be responsible for data collection, data processing, data analysis (initially and over time), and data dissemination.
- The aspect(s) of capacity to be strengthened. For example, capacity building may focus on data collection for production of the statistical abstract; on the capacity within sector ministries to utilize survey data, operational statistics, and other data for planning purposes; or on the capacity of public sector and university researchers and private consultants to undertake more complex analyses using more advanced...
statistical techniques. Capacity building may focus on each aspect individually or on a combination of these.

- The technical assistance, training and equipment that is required.

The objectives will differ from country to country. In countries like the Kyrgyz Republic and Vietnam, where there is almost no experience with unit record data, the initial goal should be to develop a cadre of government staff able to undertake simple means and cross-tabulations, to make these results available to a larger group of government staff, and to train a limited number of researchers in more sophisticated analytic techniques. Longer-range goals would be to develop more sophisticated analytic capacity among a wider number of analysts. In countries like Jamaica and South Africa, where a tradition of descriptive analysis of policy-relevant data already exists, the goal is to build a cadre of professionals able to do complex research (hypothesis testing, causal analysis, simulations of policy alternatives) on the social sectors for use in policymaking.

Activities to build capacity for social policy analysis should not be focused solely on economists. The multisectoral nature of social policy and of LSMS data means that substantive expertise in, for example, education, health, or poverty studies is required to formulate questions and to interpret data. Further, the technical skills required to analyze LSMS and other social sector data are common to several disciplines, including economics, sociology, education, public health, and management. Faculty and graduate students from the departments of economics, government, sociology, education, and public health participate in upgrading programs in Jamaica. UDAPSO included a multidisciplinary staff, with positive results. In all of the countries reviewed, the capacity-building activities successfully included staff from several agencies with different disciplinary backgrounds.

**Requirement Two: Resources**

Building analytic capacity is a slow and incremental process that takes significant investments of time and money. At least a decade is required to achieve a level of in-country policy analysis capacity that includes testing of hypotheses, formulation of models, and simulations of policy alternatives. Thus the time frame for capacity building extends beyond the life of most externally funded projects. More than one project will be required. A minimum investment of $5 million (not including survey costs) over several years can be anticipated, when the starting point includes some existing capacity in government and in local universities for at least the production of means and frequencies.

A lesson from the countries studied is that projects that support only a single round of a survey are not sufficient to upgrade analytic capacity significantly. At best, they can teach a limited number of people enough to do competent cross-
tabulations but not enough to pose policy questions or bring to bear more sophisticated analytic techniques (as the Vietnam experience illustrated). Single-round surveys can also help to build awareness of the need for a later, more extensive project and may serve as a pilot to explore particularly sensitive issues, assess risks, and fine-tune objectives and plans (as demonstrated in all five countries). While single-round survey projects will be implemented from time to time, they can provide only very modest aid in building analytic capacity.

**Requirement Three: Open Data Access Policies**

Efforts to build analytic capacity around a survey can not even begin unless and until data are broadly accessible. Both the formal policies governing access to data and the services for disseminating data sets affect the extent to which researchers will use data. Open data access is required because the bulk of policy analysis, and therefore of activities to build it, must take place outside the central statistics agency. Good policy analysis requires detailed knowledge of current and proposed policies and programs and often of advanced statistical techniques—skills that are not the bread and butter of survey institutions. Moreover, statistical agencies do not make policy. Their ability to conduct policy analysis is further limited in that they usually have only a handful of staff engaged in analysis, and they are usually subject to pressures to plan new surveys. Actors outside the statistical agency must therefore be included in programs to build analytic capacity. This, in turn, requires an open data access policy. The capacity-building program should, however, include the staff of the statistical agencies to help them be more aware of their clients’ concerns and to remain credible partners in the efforts to provide good data for policy formulation and monitoring.

The importance of access to data was somewhat hidden in the presentation of the individual case studies because, in fact, all five countries instituted open data access policies and included participants from a range of government agencies. Several of the countries organized data dissemination services; South Africa and Vietnam provided copies of the data sets to the participants in the workshops and training seminars. Open access to LSMS data does not necessarily imply open access to other data sets. The government of Vietnam, for example, maintains open access to the LSMS data but continues to deny researchers access to other data sets.

**Requirement Four: Demand for Policy Analysis**

Policy analysis should be driven by the needs of policymakers. The problem is that policymakers in developing countries frequently do not know how to frame policy questions or use the results of policy analysis. A common theme in the case studies is the need to convince government officials that data can enhance public policy decisions. Making sure from the outset that policymakers are included in the network of
activities to build analytic capacity is vital. For example, a presentation of the results of even a single round of a survey can be a useful vehicle for generating excitement regarding the potential contribution of such analysis to policy decisions. At the dissemination workshop in the Kyrgyz Republic, policymakers were invited to request custom tabulations that were conducted the very next day. In South Africa the dissemination workshops were tailored to interest policymakers in the results of work that had been done on the South African data and additional examples of work in other countries that could be replicated. In Vietnam analysts participated in the training along with their own computer technicians so that the policy analysts could pose questions for which the programmers would often generate answers. One-day workshops in each ministry could also be useful in orienting individuals to data-driven policy analysis. Disseminating the survey results, including the statistical abstract and working papers, can also stimulate the interest of program managers in data-driven analysis. But it is important to remember that one-day workshops do not generate sustained demand; continued efforts are necessary to convince decisionmakers of the gains to be expected from investing in such projects.

Ideally, there is a virtuous circle among policymakers, analysts, and data collectors. Analysts will, unbidden, produce analysis useful to understanding good policy. Policymakers will be influenced by the results and will commission work on specific issues or alternative policies. Both efforts will drive good data collection. Initiating this cycle requires breaking into the loop on both the demand (policymaker) and supply (policy analyst and data collector) sides.

**Tool One: Training**

A critical mass of trained policy analysts is required before significant changes in the policy analysis environment can be effected. Because sending large numbers of individuals abroad for training is neither feasible nor desirable, creative approaches to in-country training should incorporate a mix of short- and long-term studies.

Policy analysis encompasses a range of techniques, from simple descriptive tabulations to complex statistical analysis. Most analysts, especially those in sector ministries, need to be able to analyze a mix of operational statistics as well as budget, survey, and other data for monitoring and for planning. A smaller number, including senior research staff in government and researchers in universities and other research centers, need advanced analytic skills to answer more complex research questions about the efficiency and effectiveness of programs. Training will often be needed at both levels.

Evidence from the countries studied indicates that short-term training—probably at least six weeks in duration—can help to improve a nation’s capacity for basic analysis, including the preparation of statistical abstracts and simple two-way tables, and can be relatively effective in enabling researchers to produce useful output from
raw data. The ability to formulate research questions, to undertake more complex analysis, and to prepare research reports requires more extensive formal training, however. The inclusion of such training as a component of several projects has provided a wider perspective for assessing its effectiveness. The first lesson is that the demand for training is likely to be high. Second, selection procedures for training programs should be standardized to ensure that all participants have the necessary prerequisites for participation, including some knowledge of computers and access to a computer outside of class. Third, participation should be limited to staff whose jobs require such training. Finally, participants should be required to enroll in a comprehensive upgrading program rather than in individual courses. These mechanisms will reduce class size, permit more interactive teaching approaches, and help weed out less committed participants. Judging from experience in Jamaica and Vietnam, refresher courses in mathematics and computers may be required to prepare participants for higher-level studies.

An important lesson highlighted by the Vietnam study is that trainees need to have opportunities to interact with experienced analysts. Very few participants in this study used their newly acquired skills when they returned to work; skills are likely to atrophy when this happens.

One-shot training programs, no matter what the length or the level, do not address future needs for trained analysts. Sustainable capacity building will require strengthening graduate education to ensure the availability of ongoing training in policy analysis, including training in a range of sophisticated research and statistical techniques. One relatively easy and fruitful approach is to encourage the use of the survey data in whatever undergraduate or graduate statistical or policy analysis courses exist in national universities (through an open data access policy, well-documented data sets, and perhaps the funding of special pilot projects or a data bank). Even this requires some initial analytic and processing capacity, however. Assembling a critical mass of faculty to teach sophisticated research and statistics courses may be difficult. Upgrading existing faculty is a long-term effort. At the same time, it may be difficult to recruit already trained faculty. Clearly, fortifying or establishing graduate programs will not be feasible in every country. Regional approaches to capacity building, such as the African Economic Research Consortium, which is attempting to strengthen graduate training and research in Sub-Saharan Africa, present a viable alternative (Fine 1995). An unintended benefit of the Jamaica project is that the University of the West Indies in Jamaica is a regional institution and non-Jamaican graduate students are enrolled in the upgrading program.

**Tool Two: Technical Assistance**

The case studies demonstrate the importance of technical assistance for many aspects of the LSMS process. They show that the programs are most effective when the exter-
nal expert works closely with a local counterpart and that care must be taken to ensure that resident technical assistants do not end up substituting for—rather than training—local staff. This last lesson implies that extreme care must be taken in writing the terms of reference for the external experts and in arranging counterparts. It also means that short-term consultants and twinning arrangements (where an external agency serves as a mentor for a local agency over several years) may be more effective than resident technical assistants in promoting improved capacity. When foreign technical assistance is used to substitute for local staff, opportunities to improve in-country capability are lost. In Bolivia the long-term technical assistance provided by HIID seems to have been effective in building capacity because the technical assistance was designed as a teaching device rather than as a mechanism to accomplish data collection and analytic tasks.

**Tool Three: Research Apprenticeships**

Involvement in all stages of policy research, including identifying the research questions, preparing the proposals, implementing the research, and writing up the final report, is essential for learning and applying the skills acquired in training. The research can be designed to help refine the skills of the staff members—something like an apprenticeship. Thus the technical assistance should serve a mentoring function rather than actually carrying out the research. In such project components, it would be premature to set the research agenda before some work has been done on the demand side with policymakers and some capacity building has occurred on the supply side with analysts.

**Tool Four: Policy Analysis Units**

Policy analysis units are difficult to staff and even more difficult to sustain. Ample lead time is required to establish new units or to upgrade existing ones. Governments are likely to have difficulty staffing policy analysis units because civil service salaries make it almost impossible to recruit qualified and experienced analysts and because the shortage of qualified personnel means that competition will be stiff for the few people with the knowledge and experience required. In most cases, it will be necessary to upgrade the current staff, even though high turnover can be expected as experienced individuals leave for better-paying jobs outside government.

If the project design includes establishment of a policy analysis unit and if this unit is to coordinate many activities with other agencies (such as universities, the statistical agency, and sector ministries), vesting the administrative and research leadership functions in two different positions may be wise, for two reasons. First, it is difficult to find the right mix of research and administrative skills and experience in one person, and second, the time spent on routine administrative duties, fire-
fighting, and interagency diplomacy may mean that the administrator is unable to devote significant time to the substance of research.

The sustainability of such efforts frequently depends on the level of political support they enjoy. A minimal level of support will be required even to formulate a capacity-building effort, but the project itself can generate further demand for data-driven analysis. A household survey is a useful mechanism for sensitizing policymakers to the types of questions that can be answered with empirical data and to stimulate interest in quantitative analysis of policy questions.

Notes

Lorraine Blank is a consultant at the World Bank, and Margaret Grosh is a senior economist in the Human Development Network at the World Bank. This paper relies greatly on five case studies and the many people who contributed to them:

• The Bolivia case study was written by Manuel Contreras and benefited from comments by Jere Behrman, Christian Darras, Vicente Freles-Libils, George Gray-Molina, Menno Pradhan, Juan Carlos Requena, Eugene Szepesy, Rosa Talavera, and Nico van Niekerk.
• The Jamaica case study was written by Lorraine Blank, Margaret Grosh, and Pauline Knight, with comments provided by Pat Anderson, Paul Glewwe, P. B. K. Murthy, Jacques van der Gaag, and Colin Williams.
• The Kyrgyz Republic study was written by Raylynn Oliver, with comments provided by Jane Falkingham, Valentina Lomakova, Michael Mills, and Alexey Proskuryakov.
• The South Africa case study was written by Francis Wilson and Dudley Horner, with comments by Harold Alderman, Ann Duncan, Stephen Klassen, Carlo del Ninno, and Neeta Sirur.
• The Vietnam case was written by Lisa Drummond, with comments by Sara Bales, David Dollar, Paul Glewwe, Dominique Haughton, Jonathon Haughton, Jennie Litvack, Adam McCarthy, Nick Prescott, Jens Wandel, and Diep Vuong.
• The synthesis paper benefited from comments by Ann Duncan, Paul Glewwe, Jonathon Haughton, Emmanuel Jimenez, Jennie Litvack, John Newman, Raylynn Oliver, Laura Rawlings, Neeta Sirur, and participants in an informal seminar in the World Bank. It was edited by Fiona Mackintosh. The Development Research Group of the World Bank provided funding for the consultancies involved in all aspects of the study.

1. The survey was carried out under the direction of researchers from the University of North Carolina, Paragon Research International, and the Institute of Sociology of the Russian Academy of Sciences.
2. Modeled on the Detroit Area Study.
3. In Vietnam universities are teaching institutions, mainly at the undergraduate level, and normally do not have research programs.
4. In previous surveys data were collected and aggregated at the district level. Aggregate numbers were forwarded to the provincial level, where they were again aggregated and sent to Hanoi, where national-level aggregations were compiled. There was no file that contained the original information for each household.

References

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


Managing Government Exposure to Private Infrastructure Risks

Timothy Irwin • Michael Klein • Guillermo E. Perry • Mateen Thobani

The privatization of infrastructure should lead to the development of new infrastructure, improvements in the operation of existing infrastructure, and a reduction in budgetary subsidies. Whether countries reap the full benefits of privatization, however, depends on how risks are allocated. If, as is often the case in developing countries, governments assume risks that should be borne by investors, they may reduce incentives for efficiency and incur significant liabilities. To solve these problems, governments need to improve their policies and restrict their risk bearing to certain political and regulatory risks over which they have direct control. When a government provides guarantees, it should attempt to measure their cost and improve the way they are handled in the accounts and budgets. Measurement and budgeting are critical to improving decisions about the provision of guarantees, to improving project selection and contract design, and to protecting governments from unknowingly entering into commitments that might jeopardize future budgets.

The resurgence of private infrastructure in the past decade should have large benefits for developing countries. Compared with government enterprises, private firms typically have stronger incentives to build and run infrastructure businesses effectively and at low cost. If prices reflect costs and the firm’s profits depend on consumer demand, private firms tend to choose good projects. Privatization also encourages and facilitates the imposition of cost-covering tariffs, thus addressing the problem of underpricing that has afflicted many publicly provided infrastructure services. Greater efficiency and cost-covering prices allow firms to make investments and provide services that might not otherwise have been possible. They simultaneously improve the government’s fiscal position by making available the same quantity and quality of service with smaller budgetary subsidies.

Infrastructure subjects private investors to major risks because the investments are often large and their costs can be recouped only over long periods of time. Two special features of infrastructure create additional risks. First, the investments are largely sunk; the assets cannot be used elsewhere except at great cost. Second, infra-
structure projects often provide services that are considered essential and are provided by monopolists. As a result, services are highly politicized. This combination of factors makes investors especially vulnerable to opportunistic government actions.

Before the investment is made, the government has every reason to promise to treat the investor fairly—to allow cost-covering tariffs and to avoid changing regulations in ways that would adversely affect the investor. Once the investment is made, however, the government has an incentive to renege on its promises. The government can satisfy political demands to reduce prices or otherwise appropriate the investor's profits without causing the investor to pack up and leave. Lax and Sebenius (1981) discuss the problem in the context of mining. These incentives make returns to private investors uncertain and more sensitive to the host government's behavior.

To protect themselves from such risks as nonpayment by purchasers, cost overruns, and low demand, private investors often ask the host government to provide extensive guarantees. In other words, they ask the government to enter into some form of arrangement that results in the net wealth of the government—not the private investors—varying with the risky outcome. Eager to encourage investment without making any cash outlays, governments often consent.

Poorly designed guarantees threaten to undermine the benefits of privatization. First, they can blunt the private investors' incentives to choose only good projects and to run them efficiently. If the government bears the risk of the project's failing, the private investor will invest in projects that are more likely to fail; having invested in a project, the private investor has little interest in maximizing its chance of success. Second, guarantees may impose excessive costs on the host country's taxpayers or consumers and expose them to too much risk. Because guarantees rarely show up in the government's accounts or budgets, governments may not know the extent of their exposure. Moreover, economic crises can trigger many guarantees simultaneously; many of the government's contingent liabilities can thus become actual and current all at once. At worst, the issuance of guarantees may contribute to crises by encouraging excessive entrepreneurial risk taking (Merton 1978). Guarantees also may lead to asset stripping, in which a firm's insiders extract value from the firm even as they drive it into bankruptcy or excessive foreign borrowing (Akerlof and Romer 1993; McKinnon and Pill 1995).

Policy Reforms to Reduce Risks

Governments issue guarantees to make projects attractive to investors, often using risk bearing as a way to compensate for shortcomings in current and expected future policies. One of the best things governments can do to make projects more attractive without issuing guarantees is to put in place good policies that generally reduce risks and raise expected returns. Stable macroeconomic policies, for example, reduce the
likelihood of large changes in exchange and interest rates and therefore lessen the pressures on governments to prevent convertibility or transferability. The regular disclosure of timely and reliable information on the state of the economy and the government's finances makes it easier for investors to forecast future revenues. Liberal capital markets permit investors to spread risks more broadly—both locally and internationally—and allocate them to those most willing to bear them. Good regulatory policies reduce the risk that investors will be exploited after they have invested. Creating nonpolitical regulatory agencies, for example, reduces investors' fears that politicians will keep service prices too low to allow an adequate rate of return. Strengthening the independence and quality of the judiciary reduces investors' fears of being mistreated by the executive. Permitting international arbitration reduces investors' fears that they will be mistreated by local courts that are not independent.

Not surprisingly, governments that have established good policies and persuaded investors that those policies will be maintained can attract private investment without extensive risk bearing. In the United Kingdom, for example, the government attracts large amounts of private investment despite its policy of not bearing even regulatory risks except where they relate specifically to a project (United Kingdom 1995). When developing countries have introduced good policies and maintained them for a few years, they have also been successful in attracting private infrastructure capital without guarantees. In Argentina, for example, the complete restructuring and privatization of the power industry has permitted the government to attract private investment without having to assume major risks or issue guarantees (Klein 1997). In Chile private firms recently have invested in telecommunications, power, and gas without government guarantees (Jadresic 1997).

In this article, we focus not on the surrounding policy environment but on the appropriate allocation of risks of private infrastructure projects among governments, private investors, and consumers. We also discuss how governments should budget and account for risks.

Types of Risk

Governments, consumers, and private investors bear various types of risk, including demand and payment risk, exchange and interest rate risk, and political and regulatory risk. They also bear implicit risk and other types of risk. See box 1 for more details on defining risk.

Demand and Payment Risks

In privatizing toll roads, the host government often commits itself to ensuring that the private owner receives at least a minimum level of revenue when demand is lower.
Box 1. Defining Risk

In theoretical treatments of finance and government risk bearing, as well as in the applied fields of securities analysis and portfolio management, "risk" is often used to refer to the volatility of returns around an average or expected return (see, for example, Markowitz 1991 and Arrow and Lind 1970). In this sense, risk is equivalent to the statistical concept of variance, and a project's risk can increase without any change in the expected (or mean) return on the project. Investors who were risk neutral (in the sense used in economics and finance) would be indifferent to risk in this sense, and risk of this sort can be effectively eliminated by diversification if it is not systematic.

By contrast, in project finance, "risk" frequently refers to the ways in which actual results may be worse than planned. Here the benchmark is not the expected return of the project but the (generally higher) return that investors would receive if everything went according to plan. For example, investors estimate the returns they will earn on the assumption that the government will not expropriate their investment, while noting the risk of expropriation. An increase in expropriation risk in this sense does not just increase the volatility of returns, it reduces the expected return. Even risk-neutral investors would prefer to avoid these "risks." Diversification cannot eliminate this risk; it can only spread the loss among many people.

We use "risk" in the sense of variance or volatility around a statistically expected outcome. Expropriation risk, for example, is thus the volatility in returns around an expected return attributable to uncertainty over whether the government will expropriate.

than expected. This commitment shifts some of the risk of variation in demand to the government. In the El Cortijo–El Vino toll road project in Colombia, for example, the government undertook to reimburse the concessionaire if traffic was less than 90 percent of the specified level. The government agreed to pay the concessionaire an amount equal to the toll times the difference between 90 percent of the estimated number of vehicles and the actual number of vehicles (Lewis and Mody 1997).

Governments bear similar risks in other sectors. The Colombian government, for instance, provided a minimum revenue guarantee when it awarded a build-operate-transfer concession for a new runway at Bogotá's El Dorado airport in 1995 (Juan 1996). And many governments, through their utilities, have agreed to pay independent power producers a fixed amount each year that is independent of the actual level of power subsequently demanded from them.

An agreement by a state-owned utility to pay an independent power producer irrespective of demand protects the investor from the risk of falling demand for power or of new and cheaper generators coming on stream in the future. But it does not protect the investor from the risk of the utility defaulting on its obligations. To protect themselves against this risk, investors usually ask the government, which is more creditworthy than the utility it owns, to guarantee the utility's payments. In some cases, as when Pakistan sought to expand its power generation capacity, investors ask their governments to ask multilateral agencies to guarantee payments.

232

Exchange and Interest Rate Risks

Governments have sometimes borne the risks associated with adverse fluctuations in exchange and interest rates. The Spanish government, for example, had many private toll roads built during the 1960s and early 1970s and bore the exchange rate risk on the foreign loans that financed the roads. Gómez-Ibáñez and Meyer (1993:126) describe the government guarantees and their rationale:

The Spanish government had required the early concessions to finance a large part of their costs from foreign debt in order to ease Spain's balance-of-payments problems and to avoid drawing away domestic savings from other projects. The 1972 law on toll road concessions set standards that at least 45 percent of construction costs be financed from foreign loans, at least 10 percent from equity, and no more than 45 percent from domestic loans. The early Spanish highway companies had trouble raising funds from foreign capital markets, however, and in return the government agreed to guarantee some of these loans and to protect the companies from exchange rate fluctuations. The 1972 law specified that the government would guarantee up to 75 percent of the foreign loans; moreover, all foreign loans would be denominated in pesetas with the government assuming the full exchange rate risk.

If the peseta depreciated relative to the foreign currencies in which the loans were made, the concessionaire's loan repayments would remain the same, but the Spanish government would make an additional payment to ensure that the foreign lenders received no less foreign currency. In the end, the Spanish taxpayers spent about $2.7 billion as a result of the guarantees.

Political and Regulatory Risks

Governments often bear certain political and regulatory risks, even when they bear none of the risks mentioned above. In the Melbourne City Link, a private toll road in Australia, private-sector parties bear most of the demand, payment, and exchange and interest rate risks. The state government bears several risks that are tied to actions taken or influenced by the government. For example, if the government subsequently bans toll roads or takes actions that deliberately reduce the profitability of the private investor, the government will compensate the investor. The government bears the risk associated with possible court findings that aboriginal land rights have been violated. And it bears the risk that workers on the construction site will strike to protest against the state government rather than as part of site-specific disputes. In another example, the government of Pakistan has a policy framework for private power generation. The government agreed to "cover certain political and govern-
mental force majeure risks, provide protection against changes in certain taxes/duties, and ensure foreign exchange convertibility for the projects” (International Finance Corporation 1996:49).

Implicit and Other Forms of Risk Bearing

Government risk bearing need not be made explicit in contracts or laws. Sometimes everyone concerned expects that the government will in fact bail out a company if it would otherwise fail. The case of private Mexican toll roads may provide such an example. They were partly financed by commercial banks, which were owned at the time by the government. Some observers have argued that the commercial banks exercised less care than they should have in assessing the credit risks they assumed. Although the Mexican government did not explicitly agree to bear the credit risks taken on by the banks, it did in the end bail them out when they got into trouble, partly as a result of the poor financial performance of the private toll roads. Some argue that the banks had expected the bailout and that this expectation had an effect similar to that of an explicit government guarantee.

Governments bear risk in other, less obvious ways as well. They may lend directly to projects and bear repayment and perhaps interest rate risks. They may become part owners of a project and thereby bear a proportion of the overall risk of the project. Moreover, governments own a share of many firms, in an economic if not a legal sense, through the corporate tax system: if profits are high, the government gets more corporate income tax; if they are low, it gets less.

Principles of Risk Allocation

Infrastructure project risk can be allocated, at a broad level, to governments, firms, or consumers. How should governments decide whether to bear risks in a private infrastructure project? If they do decide to bear risk, which risks should they take on?

Two critical factors determine whether an agent should bear risk: the degree to which the agent can influence or control the outcome that is risky and the agent’s ability to bear the risk. These two factors often push in different directions—the group or organization that has most control over the risky outcome may not be in the best position to bear the risk. Other things equal, risks should be allocated to the agents who can best control the risky outcome and to agents who can bear the risk at the lowest cost. Those agents are the least risk averse because they can most easily insure or hedge against the risk, or because they can spread the risk among many people. The two critical risk factors must be balanced by three other factors: the incentives to reduce risk, the transactions cost of allocating risks, and second-best considerations stemming from credibility issues and policy transitions.
Control over Risk and the Cost of Bearing Risk

The benefits of allocating risks to those who can best control them must sometimes be weighed against the benefits of allocating risks to those who can bear them at least cost. Spreading a risk among many shareholders or taxpayers may lower the costs of risk bearing, but allocating a risk to a small number of agents who have control over the risk may help ensure the success of the project. Rather than diversifying risks completely, then, it usually pays to give managers and strategic investors significant stakes in the project. A single risk-averse entrepreneur would face a higher cost of bearing demand risk than a government, but the entrepreneur would face stronger incentives to increase demand and reduce overall risk.

In a telecommunications concession, for example, the demand risk would be the risk that demand may be higher or lower than forecast. The principle of control over risky outcomes suggests that the firm should bear this risk because it can effectively increase demand by keeping service quality high. By contrast, convertibility risk—the risk that local currency may not be convertible into foreign currency—should be borne by the government because it alone controls this risk.

Using the telecommunications example, the government of a large country with many taxpayers can bear demand risk at lower cost than a small telecommunications firm. But a large telecommunications firm with lots of small and diversified shareholders can assume demand risk at a lower cost than the government of a small country. The correlation of risk with other assets of the agent bearing the risk also matters. It may not be desirable for the government to bear demand risk because the government is likely to suffer the consequences of low demand just when its tax revenues have fallen. Foreign investors may be better placed to assume this risk because they probably hold a portfolio of assets whose value is not correlated with local business conditions.

Fiscal Incentives, Transactions Costs, and Second-Best Considerations

Even when governments have implemented good macroeconomic, legal, and regulatory reforms, firms may be reluctant to invest without government guarantees because they question whether the reforms will be maintained or protected by courts. Thus, even if a country’s laws and regulations are as good on paper as those of, say, the United States, investors will be considerably more wary about investing there. In such cases, governments may have to provide explicit undertakings—to allow convertibility or to compensate in case of expropriation, for example. In countries such as the United States, investors may think it unnecessary to seek certain explicit project-specific guarantees, either because the risks are negligible or because the investors are confident that the legal system and the courts will protect them in case of problems. In countries that have reformed their policies only recently, investors may want the
government to assume these risks explicitly in a contract, often one that is guaranteed by a multilateral agency.

Allocating risks that are better controlled by the government to the government will improve outcomes only if the government responds to financial incentives. Governments are generally less responsive than firms to financial incentives because government decisionmakers often do not act in the interests of the citizens. When the financial consequences of the government’s risk bearing do not show up in its budgets or accounts, government may be less responsive still. If, for instance, the likelihood of the government’s permitting convertibility is unaffected by any obligation to pay compensation in case of malfeasance, there is no value in allocating this risk to government. Similarly, as decisionmakers face weak incentives to lower costs, governments may be less adept than private investors at taking advantage of opportunities to reduce risk, for example, through diversification or hedging.

Transactions costs should also be considered. An allocation assigning each of a project’s many risks according to each party’s control over the outcome and its costs of risk bearing may require detailed analyses, tough negotiations, complex legal contracts, expensive monitoring arrangements, and possibly the high costs of settling disputes in court. The optimal allocation of risk takes these costs into account.

Second-best considerations may also come into play. According to the criteria for risk allocation discussed here, many governments in the developing world appear to bear too much risk when they privatize infrastructure. This situation, however, may still be better than the alternative of public ownership. When the government bears all the commercial risk, privatization almost always transfers some risk to the private sector. When the government guarantees the sponsor of a private toll road 90 percent of expected toll revenue, for example, the government bears less risk than it would if it owned the road and bore 100 percent of the risk. Thus, privatizing the road and providing guarantees to the concessionaire may be better than having the government build and operate it without private participation or not build the road at all. Similarly, governments that are unable to carry out all the necessary reforms to attract private infrastructure investment without government guarantees may prefer the second-best option of privatization without the full transfer of commercial risks to private investors. In many such cases, however, a direct fiscal subsidy may be preferable to a government guarantee.

Allocation of Typical Infrastructure Risks

Typical risks that governments are asked to assume in infrastructure privatization projects include political, regulatory, quasi-commercial, demand and construction cost, and exchange and interest rate risks. This section provides some practical guidelines on how to allocate these risks.
Political risks, such as expropriation, currency inconvertibility, and currency nontransferability, are directly under the control of the government. There is good reason to encourage the government not to create losses associated with any of these three risks. Thus, it makes sense for the government to assume these risks. The main issue is how the government can credibly commit itself to bearing the risk—that is, to commit itself not to create the conditions that would lead to loss or to compensate investors fully in the event of loss. Obtaining this commitment may involve approaching an agency such as the Multilateral Investment Guarantee Agency (MIGA).

Regulatory risks pose trickier questions. Should the government commit itself not to change the laws and regulations affecting the investment project or to compensate in case it does? On the one hand, these risks clearly fall under the government’s control. On the other, it is sometimes desirable for the government to change laws in ways that adversely affect investment projects. It may be beneficial to increase taxes to fund needed public investment, for example, or to impose regulations to mitigate newly recognized environmental problems. In many cases, such as that of new environmental regulations, the government can bear the risk and still change policy—it just needs to compensate firms for the policy change. In other cases, however, compensation cannot be reconciled with flexibility. If governments had to compensate everyone for imposing higher taxes, for example, they could never increase their (net) revenue. Smith (1997) argues for a case-by-case approach, noting that countries with better reputations for treating investors reasonably can adopt more flexible rules.

Quasi-commercial risks arise when an investor contracts with public suppliers or purchasers that may renege on contractual commitments. In many power projects, for example, investors ask for a guarantee from the government in case the government-owned power utility fails to pay for bulk power generated by the firm. If the utility has full autonomy, there is little to be gained by a government guarantee. But if the utility is completely beholden to government decisions, government guarantees may be desirable. Increasing the agency’s autonomy by privatizing it is a preferable solution.

In road, bridge, or tunnel projects, governments are often asked to bear demand and construction cost risks. These risks occur when variability in demand or construction costs makes profits higher or lower than their expected value. These risks are often critical, but the rationale for them in terms of the framework set out above is weak. The concessionaire usually has considerably more control than the government over construction costs and a greater incentive to avoid white elephant projects. Government policies can influence demand, but assigning demand risk to the government reduces the incentives to screen projects carefully with a view to investing only in those in which expected demand is sufficient to justify the project. At the same time, however, the toll road operator may have little control over the demand risk. As long as roads or bridges meet certain minimum standards, traffic may vary.
Little with increases in quality. Therefore, governments need to balance their desire to provide incentives for proper project screening with the recognition that operators often have limited control over demand.

By changing the way they regulate infrastructure, governments can reduce the demand risk faced by concessionaires and thereby reduce the concessionaires' demand for guarantees. Instead of auctioning the right to operate the service for a fixed period of time, as is typical in most road and bridge concessions, the term of the operating concession could be made to vary with demand. If demand is higher than expected, the concession will be shorter; if demand is lower, the concession will be longer. This method, which has been used in the United Kingdom for bridges, reduces risk.

In an ingenious variant of this method, Engel, Fischer, and Galetovic (1997) propose to award the concession to the bidder seeking the lowest present value of revenue, calculated with a discount rate specified in advance by the government. The concession ends when the concessionaire's revenue reaches the present value it had sought. The concessionaire still bears some demand risk—if demand is too low, revenue may never reach the target value—but it bears much less. Moreover, the investor still has an incentive to select only those projects that are likely to be financially attractive without government subsidies.

Floating-rate loans fund many infrastructure investments, making the projects' profits highly sensitive to changes in interest rates. Projects often involve considerable foreign financing. If project revenues are in local currency and the investors want to earn foreign-currency profits, foreign investors will suffer if the local currency depreciates. Ensuring that the right parties bear exchange and interest rate risk is thus important to the success of the project. Should the government accede to investor demands to assume exchange and interest rate risks?

At first glance, our framework appears to suggest that the government should bear infrastructure risks because it can better control them. Private investors have almost no control over the exchange rate or prevailing interest rates. If governments bear such risks, they have a financial incentive to adopt macroeconomic policies that tend to prevent depreciation or increases in interest rates. Such policies may also have a useful signaling effect. Governments that plan to adopt imprudent macroeconomic policies will find risk bearing expensive. Most likely, governments that choose to issue such guarantees will act reasonably (at least if politicians are concerned about the government's fiscal position).

Mas (1997) argues cogently, however, that three other factors often outweigh the incentive and signaling benefits. First, it is difficult to separate the effects on project profitability of exchange or interest rates and business decisions. Losses from currency depreciation, for example, could be blamed on the government, which allowed the currency to fall, or the firm, which left itself exposed by borrowing in foreign currencies. Second, in flexible exchange rate regimes, exchange rate guarantees may have undesirable as well as desirable incentive effects on the government. The guar-
antees discourage governments from allowing their currencies to depreciate in the wake of a terms-of-trade shock, for instance. Third, many governments and the taxpayers that back them may already be exposed to the risks associated with interest rate and exchange rate shocks. An adverse terms-of-trade shock, for example, might lead to both a depreciation and a decline in local incomes, forcing the government to compensate investors just when its tax base has shrunk. Foreign investors would not face this problem and may be in the best position to bear the risk.

Measuring and Budgeting for Risk

Whichever risks a government does take on, it needs to consider how it can measure them and incorporate them in its accounts and budgets. Otherwise, the government will have difficulty making good decisions about whether to assume risks and may even court financial disaster.

As a simple step in improving the monitoring and management of risks, the government can compile and publish a consolidated list of its contingent liabilities and the maximum amounts it stands to lose. The New Zealand government, for example, presents this information in its statement of contingent liabilities (table 1).

Calculating Expected Losses

The identification and listing of guarantees has limited usefulness. In particular, it does not provide information on the likelihood of losses. It reveals maximum possible losses, without indicating which losses the government should expect. Governments would find it useful to quantify not only the maximum possible loss but also the likelihood of losses and, therefore, the expected loss.

Sometimes it is simple to estimate expected losses. For example, the government might guarantee a payment of $1 million by one of its state-owned enterprises. For a 10 percent chance that the enterprise will default (and a 90 percent chance that it

<table>
<thead>
<tr>
<th>Table 1. Statement of Contingent Liabilities: Summary Table, New Zealand, 1997 and 1998</th>
</tr>
</thead>
<tbody>
<tr>
<td>(millions of New Zealand dollars)</td>
</tr>
<tr>
<td><strong>Quantifiable contingent liabilities</strong></td>
</tr>
<tr>
<td>Guarantees and indemnities</td>
</tr>
<tr>
<td>Uncalled capital</td>
</tr>
<tr>
<td>Legal proceedings and disputes</td>
</tr>
<tr>
<td>Other contingent liabilities</td>
</tr>
<tr>
<td><strong>Total quantifiable contingent liabilities</strong></td>
</tr>
</tbody>
</table>

will make the full payment), the government’s expected cost of the guarantee is $100,000. For more realistic cases, the expected cost may be more difficult to calculate. There may be more than two relevant possibilities, and the estimation of the probabilities may be extremely difficult.

Nevertheless, the calculation of expected losses is sometimes feasible using relatively straightforward techniques. The most tractable cases are those in which the government has issued a large number of similar guarantees for many years and has recorded information on defaults. In such cases analysts can calculate the expected cost of the guarantees in the same way as, say, car insurance premiums are calculated. The reforms the U.S. government enacted with the Federal Credit Reform Act of 1990 are examples of this (Lewis and Mody 1997).

The cost of some unique guarantees can also be estimated simply. Full credit guarantees, in which the government guarantees the repayment of a loan to another party can, for example, be calculated by looking at the interest rate charged on guaranteed and nonguaranteed loans. If a firm pays 15 percent interest on its nonguaranteed debt and 10 percent on loans guaranteed by the government, the annual value of the guarantee is 5 percent of the amount borrowed (Mody and Patro 1995).

Analysts can value guarantees and contingent liabilities, including more complicated ones, using the techniques developed in the past 25 years to value financial derivatives (such as options, futures, and swaps). Extending a credit guarantee, for example, is equivalent to the government’s selling a put option to the lender, which gives the lender the right to put the loan to the government. The valuation of some guarantees requires the skills of financial specialists, however, and the feasibility of timely, reliable, and cost-effective valuation has not yet been widely tested. But the possibilities are not merely theoretical: guarantees in both Colombia and the United States have already been valued using option pricing techniques (Lewis and Mody 1997).

Valuing the government’s guarantees and other contingent liabilities—and not simply noting maximum exposure—has important advantages. By calculating the expected cost of the government’s guarantees, analysts can more easily compare guarantees with cash subsidies. When guarantees are not valued, a government may prefer to provide a guarantee instead of a subsidy, even if the guarantee costs more than the subsidy, because a future administration may bear the hidden costs of the guarantee. When guarantees are valued, policymakers are more likely to make decisions on the basis of real rather than apparent costs and benefits.

**Incorporating Expected Losses in Accounts and Budgets**

If analysts can reliably calculate expected losses, the government should incorporate the estimates in its accounts and budgets. Most governments have cash-based bud-
gets and accounts. The budget authorizes the government to incur certain cash expenditures; the accounts show how much cash the government has received and spent. Noncash items, such as the depreciation of assets during the year and revenues earned but not received in cash (including taxes owed but not yet paid), do not appear in the budget or the accounts. Such governments do not report their balance sheets or net worth.

Although governments should include guarantees and other noncash items in their cash-based budgets and accounts, fully incorporating those items would require a switch away from cash-based systems. With standard accrual accounts and budgets, many noncash expenditures show up in the government's budget and in its operating statement. The government has no fiscal incentives to prefer these noncash expenditures to cash expenditures. Standard accrual accounting discloses guarantees (as in table 1), but it records them as expenses only for probable and quantifiable losses (Afterman 1997). From an economic point of view, this distinction between probable and improbable losses is not always useful; a 10 percent chance of losing $1 million is worse than a 90 percent chance of losing $100,000. Economists find it more useful to estimate the present value of the expected loss arising from the contingent liability.

An ideal system of accounting and budgeting would record the expected present value of all active contracts. Under such a system, a government would have no fiscal incentives to issue guarantees instead of giving subsidies of equivalent value. Both types of contracts would show up as expenditures affecting the deficit, and both would require appropriation by the legislature. Although full present-value accounting and budgeting are not feasible, governments can implement accrual accounting—and systematically record significant and quantifiable present values—even when losses are not probable. With accrual accounting, governments can take a crucial step toward the better management of guarantees.

Measuring Risk as Well as Expected Losses

Estimating, reporting, and budgeting for expected losses is important, but expected costs do not tell the government and those that monitor it everything they need to know. Just as guarantees with the same maximum exposure differ significantly if the expected loss differs, guarantees may differ even though they have the same expected but different maximum losses. The whole range of possible outcomes—that is, risk in the sense of volatility—matters.

Governments should develop systems for summarizing and reporting the major risks as well as the expected costs they face. They could use value-at-risk reporting, in which organizations report the largest loss that can be incurred with a probability greater than, say, 1 percent or 5 percent. Value-at-risk reporting applies statistical theory to the description of assets and liabilities. For example, a bank may report...
that its daily value at risk at the 1 percent level is $10 million, meaning that there is only a 1 percent chance, under normal market conditions, that it will lose more than $10 million in the next day (Jorion 1997). Although banks were the first organizations to report value at risk, the principles behind such reporting apply to all organizations. Governments should think about how they can apply these principles to their own operations and what useful estimates they might be able to publish.

Taking a Government-wide Approach to Risk Measurement

Expected losses can be measured individually and then aggregated; the government’s total expected loss from issuing several guarantees is simply the sum of the expected losses associated with each guarantee. Normally, risks cannot be estimated individually and then summed because the total risk depends on the relationships between the individual risks.

A government worries not about the risk relating to any one guarantee, but the riskiness of its portfolio of assets and liabilities. The reporting of risk exposures is most useful when done for the government as a whole. Therefore, measuring the true risks associated with infrastructure privatization requires an assessment of the riskiness of all of the government’s operations. Exchange rate guarantees, for instance, are likely to be more risky if the government also has net debt denominated in foreign currencies. A depreciation of the local currency may simultaneously increase debt service payments and trigger payments under the guarantee. A portfolio-wide approach also allows the government to judge the importance of the risks associated with infrastructure privatization relative to the risks associated with government debt, pensions, the banking system, and debt owed by city and provincial governments.

The creation of consolidated accrual accounts—incorporating present values where practicable—is a step toward portfolio-wide risk monitoring. Combined with other information on likely future revenues and expenditures, the balance sheet can provide an indication of the government’s vulnerability to exchange rate, interest rate, and other shocks. It can indicate the relative importance of monitoring and managing infrastructure guarantees on the one hand and debt on the other. Ultimately, governments should aim to measure portfolio-wide value at risk. The achievement of that goal appears to be a long way off, however.

Risk Management

A government is in a good position to consider managing its portfolio only when it has good information on the risks to which its total portfolio is exposed. It can manage to reduce those risks, for example, by entering into derivative contracts to hedge against changes in exchange rates or commodity prices. Whether the govern-
ment should act to reduce risk (rather than just reduce expected costs) is perhaps an open question. It is analogous to the firm's decision about whether to attempt to reduce the variance of shareholders' returns or simply to maximize the expected value of the returns.

Citizens, like shareholders, are usually risk averse. The policy decision is whether the government should manage risk on their behalf or publicize its risk exposure and permit taxpayers to diversify and hedge their portfolios in ways that give them the risk exposure they want. In practice, citizens often lack sophistication in considering risk. Moreover, because of transactions costs or government restrictions, citizens may have limited opportunities to hedge against government risk. In that case, governmental risk management may be desirable. It would aim to achieve a level of risk the government finds acceptable, given its citizens' risk preferences, at the lowest possible cost.3

**Conclusion**

Whether infrastructure privatization will realize its potential depends on how government allocates the risks facing privatized businesses. Government can increase the benefits of privatization by assuming risks it can control itself (convertibility risk, for example), but it should normally avoid bearing other risks. That way, investors face strong incentives to select projects well and to run those that they do select efficiently.

In many infrastructure privatizations, governments have assumed risks that investors should bear, both because the investors have been understandably wary of taking on the considerable risks involved and because governments have been able to offer guarantees without incurring any immediate cash costs. A government can thus take two steps to improve the environment for risk allocation. It can reduce the extent of the risks investors face by pursuing stable macroeconomic policies, disclosing information, implementing good laws and regulations, and liberalizing financial markets. And it can improve the way it measures, budgets, and accounts for the guarantees it does give, so that the costs and risks are clear at the time the guarantees are issued—not only when the government must subsequently pay up.

**Notes**

Timothy Irwin is an economist in the Private Participation in Infrastructure section, World Bank; Michael Klein is chief economist of Shell International; and Guillermo E. Perry is chief economist and Mateen Thobani is principal economist in the Latin America and the Caribbean Region, World Bank. The paper summarizes papers and discussions of a conference held in Cartagena, Colombia, in May 1997 on the topic. The authors are grateful to conference participants for their contributions and owe special thanks to Katherine Brewer for her excellent organization.

*Timothy Irwin, Michael Klein, Guillermo E. Perry, and Mateen Thobani* 243

2. In the extreme case of a guarantee for a risk that is negatively correlated with the value of the government’s portfolio, assessment of risk in isolation would lead the government to think the guarantee created risk, when in fact it reduced it. In other words, because the value of this guarantee tends to fall when the value of the rest of the government’s portfolio rises, and vice versa, the total volatility of the government’s wealth would be greater without the guarantee.

3. A useful short summary of risk management for countries (as opposed to governments) can be found in Claessens (1992). Claessens and Qian (1991) apply the techniques to African countries.

References

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


Klaus Deininger • Hans Binswanger

This article examines the evolution of policy recommendations concerning rural land issues since the formulation of the World Bank’s “Land Reform Policy Paper” in 1975. That paper set out three guiding principles: the desirability of owner-operated family farms; the need for markets to permit land to be transferred to more productive users; and the importance of an egalitarian asset distribution. In the 25 years since that paper was published, these guiding principles have remained the same, but it is now recognized that communal tenure systems can be more cost-effective than formal title, that titling programs should be judged on their equity as well as their efficiency, that the potential of land rental markets has often been severely underestimated, that land-sale markets enhance efficiency only if they are integrated into a broader effort at developing rural factor markets, and that land reform is more likely to result in a reduction of poverty if it harnesses (rather than undermines) the operation of land markets and is implemented in a decentralized fashion. Achieving land policies that incorporate these elements requires a coherent legal and institutional framework together with greater reliance on pilot programs to examine the applicability of interventions under local conditions.

In the rural areas of most developing countries, land is not only the primary means for generating a livelihood but often the main vehicle for investing, accumulating wealth, and transferring it between generations. Thus the ways in which access to land is regulated, property rights are defined, and ownership conflicts are resolved has broad implications beyond the sphere of agricultural production. These regulations, rights, and procedures affect not only the ability of households to produce for their subsistence and for the market but also their social and economic status (and often their collective identity), their incentive to work, their willingness to use the land sustainably, and their ability to self-insure or to obtain access to financial markets.

© 1999 The International Bank for Reconstruction and Development / THE WORLD BANK
The importance of land issues in fostering economic growth and reducing poverty was the impetus for the World Bank’s 1975 “Land Reform Policy Paper.” At the time, this dialogue was complicated both by an economic environment in which government interventions often caused the prices of rural land to deviate significantly from the net present value of agricultural profits and by a political context in which land was at the heart of a broader ideological struggle. In many developing countries today, far-reaching macroeconomic reforms have removed distortionary policies, the ideological divide has narrowed or disappeared, and the need to tackle structural issues has greatly increased the demand for policy advice. These considerations provide an opportune moment to review earlier policy recommendations and to use experience to assess the role of such policies in the broader process of development. This article reviews the analytical underpinning for policy recommendations and examines the effectiveness of such advice in the areas of tenure security, land markets, and land reform.

The broad consensus underlying current thinking about land issues can be summarized in four key principles:

- The desirability of owner-operated family farms on both efficiency and equity grounds
- The importance of secure property rights to land in eliciting effort and investment and in providing the basis for land transactions
- The need for a policy and regulatory environment that promotes transfers to more efficient land uses
- The positive impact of an egalitarian asset distribution and the scope for redistributive land reform where nonmarket forces have led to a highly dualistic ownership and operational distribution of land, that is, a distribution characterized by very large and very small holdings.

Although these principles remain valid, experience with land reforms challenges earlier assumptions in four areas. First, the 1975 World Bank land reform policy recommended that communal tenure systems be abandoned in favor of freehold titles and the subdivision of the commons. Today it is recognized that some communal tenure arrangements can increase tenure security and provide a (limited) basis for land transactions in ways that are more cost-effective than freehold titles. Where that is the case, governments may find it useful to reduce the cost of cooperation, improve accountability, and facilitate a gradual evolution of communal systems to meet emerging needs, possibly for greater individualization of property rights over time.

Second, although individual titling has great potential to increase investment and productivity, several preconditions must be satisfied for this to be a desirable intervention. The circumstances under which title is conferred are important; for example, titling should be area-based (that is, it should cover an entire area at once)
and fit within a broader strategy of rural development. Otherwise, imperfections in other factor markets may undermine or even eliminate the advantages from possession of title, at least for the poor.

Third, the earlier skeptical view of land rental markets has given way to a recognition of their critical role as a means for providing the poor with access to land. The removal of remaining restrictions on land rental is therefore a top policy priority. In contrast, however, removing the restrictions on markets for land sales may not be the most urgent requirement for increasing efficiency—and may have a negative impact on equity. Measures thus should be sequenced properly, emphasizing rentals rather than sales, and should be integrated with the development of other rural factor markets.

And fourth, a growing literature has made the case for redistributive land reform on efficiency and equity grounds. Most of the land reforms undertaken during the last 20 to 30 years, however, were politically motivated and have not lived up to expectations. Recently, a new approach has emerged: encouraging community-managed agrarian reform based on voluntary negotiation. Provided that careful monitoring permits officials to make the necessary changes in program design and that political pressures to provide free handouts to influential lobbies can be resisted, this approach can help to overcome long-standing problems of asset distribution and social exclusion, which are key factors leading to rural violence.

In addition to such changes in specific recommendations, land policy is increasingly viewed as an integral element of a broader development process rather than as a string of narrowly oriented technical interventions. This view is based on experience showing that a lack of consensus on the broader subject of land policy has often compromised the effect on development of specific interventions, such as land titling. In countries where land issues have in the past resulted in civil strife, revolution, and war, reaching a consensus requires time and involves all the relevant sectors of civil society. Initiating such a process, ensuring its integration into a broader framework of rural development initiatives, and strengthening the analytical capacity of key players have become important components of the World Bank’s approach to land issues. At the same time, the political sensitivity of such issues and the need to adapt to site-specific conditions often dictates that specific approaches be explored on a small scale before they are implemented broadly.

The Conceptual Basis of the World Bank’s Land Policy

A stylized fact, confirmed by a large literature, is that owner-operated smallholder farms are desirable from both an equity and an efficiency perspective. Secure individual property rights to land would therefore not only increase the beneficiaries’ incentives and provide collateral for further investment but, if all markets were competitive, would automatically lead to socially and economically desirable land mar-
ket transactions. The arguments advanced in favor of secure property rights are based on three observations.

First, clear property rights can prevent wasteful "overinvestment" in protective measures by individuals eager to claim and defend their property rights. As Malik and Schwab (1991) point out, property rights are a public good, and in the absence of public enforcement of these rights, individuals will overinvest in protective measures to claim and secure their rights. This privately optimal spending will be inefficient from a social point of view, particularly if the claim is secured through negative environmental externalities, as is the case in many frontier situations (Alston, Libecap, and Mueller 1999; de Meza and Gould 1992; Feder and Feeny 1991). Government regulations have often directly encouraged such behavior (Binswanger 1989).

Providing farmers with residual rights to production, even if these are only temporary, will increase the incentive to clear and cultivate land, as illustrated by the tremendous increases in output and productivity associated with the transition from collective to individual (usufruct) rights in China (Lin 1992; McMillan, Whalley, and Zhu 1989). The link between secure ownership rights (although not necessarily a formal title) and investments in farm improvements also emerges in evidence from Burkina Faso (Brasselle, Gaspart, and Platteau 1997), China (Jacoby, Li, and Rozelle 1998; Yao 1996), Ghana (Besley 1995), and Niger (Gavian and Fafchamps 1996).

At the same time, the lack of enforceable—though not always formal or individual—property rights has been associated with the unsustainable use and degradation of natural resources. A breakdown in the ability of communities to enforce rules governing the use of communally held land and the inefficiency of collective forms of production (Deininger 1995) were at the root of environmental degradation in Mexico (Key and others 1998; McCarthy, de Janvry, and Sadoulet 1998). Even in situations where individual property rights are infeasible, helping communities develop structures that overcome these coordination problems and establish effective property rights can enhance the sustainability of resource use, prevent environmental degradation, and promote the overall efficiency of land use (Baland and Platteau 1996).

A second issue relates to credit access. In addition to increasing demand for investment, as discussed above, secure landownership and the associated ability to use land as collateral can increase the supply of credit from formal sources. This can also contribute to the evolution of financial markets in more general terms (Alston, Libecap, and Schneider 1996; Carter and Olinto 1996; Feder and others 1986; Lopez 1997).

A third benefit is that written records of landownership improve the transferability of property. By reducing asymmetric information about landownership and quality, land transactions are less costly to implement, thus increasing the liquidity of the land market and making it possible to transfer land from less productive to more productive individuals. The ability to transfer land may be of limited importance, however, in the early stages of development, when nonagricultural opportunities...
and regional migration are limited, and in marginal environments, where economic opportunities are more constrained. The importance and value of being able to transfer use or ownership rights to land increase with economic development, specialization, and better-functioning markets, and one would expect the transferability of land brought about by better-defined property rights to be of increasing relevance with higher levels of population density and nonagricultural development. Indeed, in coastal China, more secure transfer rights are associated with higher allocative efficiency in the economy (Yao 1996).

Given the distribution of agricultural production and the need (especially at low levels of technology) to adjust constantly to variations in the environment, owner-operated farms have an advantage over large operations, which are associated with the large agency costs entailed in managing wage labor (Jensen and Meckling 1976). Conceptually, it has long been recognized that supervisory capacity is an important determinant of the mode of operation of large tracts of land (Eswaran and Kotwal 1985a, b; Feder 1985). Empirical evidence indicates not only that hired labor is less productive than family labor but also that the intensity of supervision matters (Frisvold 1994). A large number of studies based on aggregate or cross-sectional data confirmed the existence of a negative relationship between farm size and productivity for all but the smallest farms (Benjamin 1995; Berry and Cline 1979; Carter 1984; Kutcher and Scandizzo 1981; Newell, Pandya, and Symons 1997). The relationship weakens if adjustments for soil quality are made (Benjamin 1995; Bhalla and Roy 1988). Still, several studies using panel data with household- or plot-specific effects show a negative relationship between farm size and productivity that is likely to originate in labor market imperfections of the kind mentioned (Burgess 1997; Olinto 1995; Udry 1996).

Economies of scale and imperfections in other markets can outweigh the cost advantages of owner-operated farms. Scale economies can arise from the use of machinery and the advantages of professional management and marketing, the use of which would lead to declining average costs with firm size. Empirically, the indivisibilities associated with machinery rarely increase optimum farm size beyond the level at which, with existing technology, the labor of a family (possibly complemented by hired labor for specific seasonal tasks) is fully utilized. Even where they would, rental markets can help to overcome this indivisibility, at least to some extent. Most empirical studies (for example, Burgess 1997 and Feder and others 1989 for China, Lanjouw 1995 for India) are therefore unable to reject the hypothesis of constant returns to scale in agricultural production.

The number of cases in which true technical economies of scale apply is therefore limited. One example is plantation crops, such as bananas, sugarcane, and tea, where production is often organized on a scale that corresponds to the optimum scale of the processing factory. Even in this case, however, the supervision advantages of owner-operators have frequently led to the adoption of contract grower arrange-
Economies of scale in processing or marketing are therefore important for the size of farming operations only so long as markets for outputs and inputs are either unavailable or malfunctioning.

Imperfections in other markets, such as lack of access to capital and insurance markets, put small farms at a disadvantage. Their limited ability to cope with risk could offset the cost advantage that small family farms enjoy. For example, in the Sudan, capital market imperfections led to a positive relationship between farm size and productivity (Kevane 1996), while Burkina Faso recorded an inverse farm size–productivity relationship (Udry 1996), and in other contexts an optimum farm size emerged (Carter and Mesbah 1993). To the degree that imperfections in those markets, rather than an inherent productivity advantage of large farms, are at the source of differences in the shadow price of land across categories of farm size, improvements in these markets through regulation, better information, or cooperatives to reap economies of scale or input supply could lead to productivity gains.

The Importance of Land Market Transactions

That well-functioning land markets can promote efficiency-enhancing land transfer is well recognized. The extent of such land transfers is affected by government policies, informational constraints, and incomplete credit markets and their impact on land prices, producers' ability to access financial markets, and transaction costs associated with land rentals and sales.

Land Price Formation

Mortgaged land cannot be used as collateral for working capital. Thus the borrower would not reap the production credit advantage and would be unable to repay a loan from the income generated from the land. In addition, the value of the ability to use unmortgaged land as collateral in formal credit markets would be capitalized into land prices, and the equilibrium price of land at given credit costs would exceed the present discounted value of the income stream that can be produced from the land (Binswanger and Elgin 1988; Just and Miranowski 1988). Credit subsidies, tax advantages, and the use of land as an inflation hedge would have the same effect. The only income stream available to the poor for consumption is the imputed value of family labor. The remainder of the profits would have to be used to pay for the loan, reducing the purchaser's utility below what could be achieved in the labor market and implying that land purchases would normally have to be financed out of household savings. Even if credit access were perfect, mortgage-based land acquisitions would be unlikely to lead to redistribution in favor of landless households (Binswanger and Elgin 1988). Any additional subsidies, tax advantages, and other factors that
further increase land prices above the net present value of agricultural profits would, of course, further reduce the scope for participation of the poor (Gunal, Williams, and Romain 1996).

**Credit Market Imperfections**

The same credit constraints that make it difficult for landless farmers to finance the purchase of land also affect their ability to pay for land in the rental market. Credit rationing reduces their ability to make productive use of land, possibly outweighing the advantage of owner-operated farms, reducing their reservation price for land, and possibly leading to (relatively inefficient) rentals to large producers.

One outcome of these difficulties is that smallholders are forced to adopt costly (and relatively inefficient) insurance substitutes to enable them to deal with unexpected productivity and idiosyncratic shocks, such as adjusting their crop and asset portfolios to a low-return, low-risk combination that reduces their vulnerability. Another possibility is to replace land with a more liquid asset such as grain, which even though less productive, would provide them with greater security in case of subsistence risk (Zimmerman and Carter 1999). Indeed, for India, Rosenzweig and Wolpin (1993) show that the lack of insurance substitutes affects farmers' investments. Despite this less-than-optimal allocation of assets, small farmers are still more productive than large farmers (Rosenzweig and Binswanger 1993). Dercon (1996) and Dercon and Krishnan (1996) find that in Africa the levels of liquid assets and capital available to the household determine its ability to enter into high-risk but high-return activities.

The market for land is also affected by credit market imperfections that deny smallholders insurance against shocks, such as bad harvests or accidents, and may force them to sell off land in periods of distress. Where the covariance of weather risks would imply wide fluctuations of land prices over time, it is hard or impossible for smallholders to recover from those asset losses because they would have to sell at low prices during disaster (when there is little effective demand) and buy during normal times, when prices are high again (Bidinger and others 1991). In research on Bangladesh and India, Cain (1981) illustrates this view that transactions in the land sales market are driven by credit and insurance limitations rather than by cultivators' productive efficiency. He finds that in villages that had access to a safety net program, farmers generally sold land in response to shocks and to obtain cash for major investments such as drilling wells, purchasing pumps, and educating and marrying their children. By contrast, where such consumption-smoothing devices were absent, the majority of sales were prompted by distress (to obtain food and medicine). Whether or not households were able to buffer consumption during crisis situations had a significant impact on whether markets helped to equalize or disequalize land endowments. Indeed, distress sales have not only played a major role historically but
also are linked in the literature to the elimination of traditional mechanisms for coping with risk (Kranton and Swamy 1997).

**Transaction Costs and Incentive Issues**

Because land is immobile, land markets are localized, with several important consequences. As Balcazar (1990) and Carter and Zegarra (1995) note, land sales markets are normally highly segmented, especially in countries with a dualistic distribution of ownership and where sales by large farmers to smallholders are virtually absent. One explanation is that it is costly to subdivide large farms so as to make them suitable for smallholder cultivation. Similarly, as noted earlier, transaction costs either discourage small land transactions or drive them into the informal market.

Further, even though land rental markets do not permit perfect adjustment to the desired size of operation (Skoufias 1995), the transaction costs incurred are smaller than in land sales markets. There has, however, been long-standing concern about the scope for incentive issues to lead to efficiency losses in rental markets. Under a land rental arrangement, tenants have few incentives to undertake long-term investments unless they receive the value of the investment back during the rental term or at its end (in the form of compensation). Moreover, with share tenancy, that is, in situations where the landlord receives a share of the harvest as rent, tenants receive only a fraction of their marginal product. It is therefore difficult to motivate tenants to work hard enough, a phenomenon that is known as "Marshallian inefficiency." Share tenancy arrangements are still more efficient than wage labor, however. They may be an "optimal choice," given the constraints faced where markets for credit and insurance are incomplete. For risk-averse tenants (where risk aversion can arise out of the need to satisfy a minimum subsistence constraint), a share contract can provide insurance against fluctuations of output and income (Cheung 1969). For landlords, share tenancy insures against rent default by tenants whose wealth is insufficient to pay the rent or who, because they are credit constrained, underuse inputs (Shetty 1988). Thus the poorest tenants often receive wage contracts, while richer individuals (who have a lower risk of default) cultivate under share contracts with progressively increasing tenant shares; fixed-rent contracts are limited to wealthy farmers (Laffont and Matoussi 1995; Lanjouw 1995; Shaban 1991).

The degree to which policymakers need to be concerned about preventing efficiency losses from share tenancy contracts depends on the magnitude of this inefficiency. Although different methodologies produce widely diverging results, Shaban (1991) indicates that the losses may be relatively modest—about 16 percent. Thus even if the government imposed regulations to replace the tenancy system with more efficient forms of production, the impact is likely to be modest. The case for government intervention is further reduced because inefficiency decreases with monitoring.
and social control, that is, by embedding share contracts in long-term social or kinship relationships. In fact, Sadoulet, de Janvry, and Fukui (1997) find that share tenancy contracts between kin (but not between others) were not associated with any disincentive effects. The study of the efficiency implications of contracts should therefore be complemented by a focus on the contracting parties’ opportunities outside their specific contract and on possible changes in the economic environment that might lead to the adoption of different types of contracts (Mookherjee 1997).

**Implications for Land Sales and Rental Markets**

When the 1975 land reform paper was written, policy advice focused on land sales markets to achieve efficiency-enhancing transfers of property and took a decidedly negative stance toward rental markets for land. If other markets are imperfect, however, as noted above, sales may be less efficient and less equitable than rentals. Even where attempts to liberalize markets for land sales are embedded in a well-sequenced program of integrated factor market development, land sales may not be the most important constraint on higher productivity, and their liberalization may not be an immediate priority.

Consider the effect of land transfers in rental markets. By renting out, landowners would not forgo possible benefits from credit access associated with land ownership (and could even advance the credit thus obtained to a credit-constrained tenant under an interlinked contract) and would at the same time benefit from any efficiency advantage of the tenants’ family labor. Credit market imperfections tend to affect rental markets less than they do sales markets; moreover, rental markets are associated with lower transaction costs and generate positive externalities by facilitating the acquisition of agricultural knowledge by the tenant and adaptation to changing labor availability. Thus, rental markets may contribute more to efficiency than sales markets (Carter and Olinto 1996). Governments should therefore aim to create conditions conducive to the development of rental markets, rather than implicitly or explicitly restricting the scope for tenancy, as they have often done in the past.

**Land Reform**

The World Bank’s 1975 policy paper strongly supported redistributive land reform on equity and efficiency grounds, pointing to the success of Asian land redistribution and the Kenyan “million-acre scheme,” which redistributed land from European settlers to African farmers. The practical difficulties associated with implementing land reform notwithstanding, the conceptual attractiveness of such a policy rests on three pillars.
First, in situations where credit and product markets are incomplete, access to land can make a significant contribution to food security, households' nutritional well-being, and their ability to withstand shocks (Bardhan, Bowles and Gintis forthcoming). Evidence from China, where land was distributed largely independently of economic status, suggests that even though access to land insures household income only moderately against shocks, it provides almost complete insurance against malnutrition (Burgess 1997). Second, landownership affects economic growth and poverty reduction through credit-financed investment. The underlying idea is that the lack of collateral precludes landless individuals from making investments (in education, livestock, wells, and so on) that would require credit, even though the investments would profit both the individual and society (Eckstein and Zilcha 1994; Galor and Zeira 1993). Poor people who do not have access to assets might remain impoverished not because they are unproductive or lack skills but because they never get the opportunity to utilize their innate ability (Fafchamps and Pender 1997; Jalan and Ravallion 1997). And finally, several studies have argued that a more egalitarian distribution of assets (not necessarily land) would improve political stability. Because this issue does not relate directly to land issues, we refer the reader to the literature on this topic.

The ease of actually implementing land reform has varied considerably between "landlord estates," which had been cultivated by tenants, and "haciendas," whose tenants received a small plot of their own in return for working on the landlord's farm. In landlord estates, all that is required is a reassignment of property rights; land reform is generally easy to implement, and stable systems of production emerge. Since the end of World War II, landlord estates in Bolivia, large areas of China, Ethiopia, eastern India, Iran, Japan, the Republic of Korea, and Taiwan, China, have been transferred to tenants in the course of successful land reforms. The productivity gains associated with these land transfers were modest in cases where security of tenure had already been high, where cash rent (rather than share rent) contracts had prevailed before the reform, and where landlords had provided tenants with market access (and no substitute was available). Both welfare and productivity increased where investment opportunities were available (Callison 1983; King 1977; Koo 1968), where land ownership enabled the new owners to access markets for credit and insurance that had previously been beyond their reach (Dorner and Thiesenhusen 1990), and where new technology could be readily adopted (Otsuka, Chuma, and Hayami 1992).

By contrast, land reform in hacienda systems has been very difficult, and the "game of Latin American land reform" has been declared lost (de Janvry and Sadoulet 1989). In the great majority of these systems, large landowners responded to the threat of land reform by evicting all hired workers or tenants who could have claimed ownership under a reform program. The landlords either switched to livestock production and ranching or—aided by significant credit subsidies—shifted to highly mechanized cultivation (Binswanger, Deininger, and Feder 1995). As a result, programs of
redistributive land reform reached far fewer people than intended and were often accompanied by a decline in tenant welfare that may have outweighed the benefits of the programs. Several factors account for this lack of success.

First, if land is transferred from large to small farmers through government programs, the ability of the latter to make economically productive use of this asset is contingent on a change in the pattern of production, subdivision of the farm, and construction of complementary infrastructure. Second, because the main productivity advantage of land reform is linked to the increased incentives of owner-operators, it is important not only to avoid collective forms of production but also to ensure that owners operate their own farms. Third, beneficiaries are unaccustomed to making independent entrepreneurial decisions, an ability that is particularly important to make individual family farming a success. In many cases in which the farms acquired under a land reform program were not farmed at full capacity, the lack of funds for pastures, fencing, and so on or for startup capital was often the reason for the lack of success. Similarly, programs that were limited to transferring land to existing workers without providing those workers with complementary investment, training, technical assistance, and resources were generally associated with very limited equity and efficiency benefits.

Without access to credit markets, land reform beneficiaries may well be worse off than they had been when the landlord provided them with inputs and possibly even credit for consumption smoothing (Guinnane and Miller 1997). Restricted access to credit together with insecure property rights led beneficiaries of land reform in Nicaragua (Jonakin 1996) and the Philippines to sell off their new holdings—often at prices well below the productive value of the land. The key to avoiding such an outcome is the ability to access output and financial markets (Brooks and Lerman 1994). Arrangements where financial intermediaries provide input credit and help with marketing of the farm produce have in some cases helped beneficiaries overcome the obstacles posed by market imperfections (Deininger 1999).

Implications for Policy: Communal Tenure Systems

Communal tenure systems are dominant in most countries of Africa, in China, in indigenous areas in Latin America, and in Mexico. When the community rather than the individual owns the land, whatever market exchanges (sale or rental) exist are normally limited to the community. Individuals have very secure and normally inheritable rights to land even after a period of absence, but they do not have permanent property rights to a specific plot, a limitation that may reduce investment incentives. In some cases, communal systems also permit periodic redistribution of land by the village chief to accommodate population growth.
In the past, communal tenure arrangements were often considered economically inferior and equivalent to collective production. The establishment of freehold title and the subdivision of the commons were proposed to prevent the efficiency losses that were assumed to be associated with communal ownership. More intensive study of communal tenure systems in a broader framework and the recognition that these systems perform multiple functions has led to a reassessment of these recommendations, however.

On the one hand, the efficiency losses associated with communal tenure systems may be more modest than generally assumed, for a number of reasons. First, arable land (in contrast to pasture, forest, or fishing grounds) is, in most communal systems, cultivated by individuals who enjoy inheritable rights, which means that the static (and maybe even dynamic) efficiency losses possibly associated with communal tenure may be quite limited. Second, communal resource ownership is often maintained because it either provides public goods or takes advantage of synergies that would be difficult to provide under individual cultivation, including risk reduction through diversification (McCloskey 1991; Nugent and Sanchez 1993), economies of scale to help with seasonal labor bottlenecks (Mearns 1996), and investment in community-level infrastructure (Boserup 1965; Dong 1996). Third, when population density is low and the payoffs from land-related investments are limited, the investment disincentives associated with communal tenure are likely to be of little consequence because people do not tend to invest under either system of tenure. With arable land becoming increasingly scarce, many communal tenure systems either recognize a user’s property rights if the land has been improved or compensate the user for improvements when the land is redistributed, thus attenuating tenure-related investment disincentives (Sjaastad and Bromley 1997). Finally, although communal systems prohibit land transactions with outsiders, rentals—and often even sales—within the community (and possibly beyond) are normally allowed, providing scope for efficiency-enhancing transfers.

On the other hand, in environments with low population density and limited access to infrastructure and markets, the costs of delimiting and enforcing boundaries for individual plots are high, so the economic benefits of formal titling may not offset the expenses involved. Indeed, in several African countries, titles that were generated at high cost have lost their value as landowners failed to keep them updated. Thus, in cases where there is no clear demand for demarcation of individual plots, communal titles that are administered internally in a transparent fashion could provide tenure security at a fraction of the cost of individual titles (Bruce and Migot-Adholla 1994; Heath 1992). Communal titles also might provide a more effective safety net to ensure against risks and substitute for more costly redistributive mechanisms (Burgess 1997). Thus, instead of reinforcing an often artificial dichotomy between private and communal rights or trying to privatize land rights to “modernize” land tenure in an environment where few of the conditions for such moderniza-
tion are present, policymakers should focus on ways to increase secure property rights within given constraints.

More secure land rights may be highly valued by cultivators even under conditions of relatively low population density. For example, in Zambia (with a population density of 12 people per square kilometer and where 75 percent of the land is suitable for farming), almost 50 percent of farmers feel their land tenure is insecure and would be willing to pay something (US$40, on average) for land titles (Deininger and others 1998). Disputes, efficiency losses arising from limiting transfers and barring certain groups from land rights, investment disincentives, and land grabbing in anticipation of future appreciation are all indicators that existing land rights are inadequate. Clarification and formalization of informal property rights in a process that increases the accountability of local leaders, establishes a transparent and implementable legal basis, and provides for adjudication of boundary disputes across communities must precede any effort to award formal titles. Adopting a flexible institutional structure that gives communities freedom of choice in accomplishing these goals is therefore of great importance. The draft land policy adopted by Zimbabwe provides a good example in this regard (Zimbabwe 1998).

In countries where land ownership has traditionally been vested in the state, policymakers are concerned that a shift to individual land ownership is likely to lead to an undesirable reconcentration of land ownership. Experience suggests that this concern can be accommodated without forgoing major productivity benefits by giving producers long-term tradable leases rather than full ownership rights. For example, the household responsibility system in China (which gave 15-year lease rights and at the same time made individuals residual claimants to output) has led to tremendous increases in output and productivity. To increase investment incentives, the government has decided to replace the 15-year leases with 30-year contracts (Prosterman, Schwarzwalder, and Ping 1998). Because the degree to which earlier leases were honored varied greatly from village to village, inferences can be made regarding the impact of tenure security; studies find that more secure tenure did increase the level of investment (Jacoby, Li, and Rozelle 1998).

Establishing Formal Tenure Security

Land registration and titling have long been viewed as the main instruments for increasing tenure security, empowering a flourishing land market, and facilitating the use of land as collateral in credit markets. Although numerous studies have confirmed the positive impact of titling where the conditions are right, experience with World Bank projects has also demonstrated that titling is not a panacea for achieving a wide variety of divergent goals at the same time. The objective—whether it is to improve credit access, increase tenure security, or activate land markets—must be
clear. In addition, the ways in which individuals gain access to land before titling, whether through collective, communal, or informal means, as well as the broader trajectory of economic development, will affect the costs and benefits of specific titling instruments, their incidence across population groups, and the scope for public intervention.

**Conceptual and Implementation Issues**

Improved credit access has repeatedly been shown to be one of the major benefits from formal title. Thus, titling will confer the highest benefits where informal land transactions are common, a formal credit market that permits use of title as collateral exists, and profitable investment opportunities are available. Title is unlikely to increase the banks' willingness to lend to the rural sector where, for cultural or economic reasons, land cannot be repossessed or where land sales and mortgages are restricted (Atwood 1990; Ensminger 1997).

Even if the above preconditions are satisfied, the effect of titling may vary across groups of producers, an issue that is of particular relevance if the initial distribution of land endowments is unequal. If the transaction costs associated with lending to specific groups of producers exceed the benefits they can derive from the use of credit, title would not be expected to increase credit access. In such cases, the title might make it easier for large producers to access credit but would not make small landowners creditworthy, a situation that would deepen preexisting inequalities. To prevent this and help titling contribute to broad-based growth, concurrent measures to improve access to credit markets, and possibly a differentiated scheme of recovering the costs of establishing title, will be necessary.

If a case can be made for formal titling, it must be systematic and areawide to take advantage of economies of scale in measurement, adjudication, and conflict resolution. Similarly, complementary infrastructure (such as programs to ensure access to credit markets) can be provided more easily and cost-effectively under an area-based program. To achieve equity, titling needs to be combined with a mechanism for resolving disputes on the spot and an information campaign explaining the legal background, the titling process used, the rights of different parties, the rules of evidence, and the benefits of the appeal process. If a decision is made to title on demand, the status of individual plots will still have to be investigated on a case-by-case basis, and any reduction in the transaction costs associated with titling will thus be minimal. Moreover, titling on demand has often had disastrous consequences for the poor because individuals with good political connections can often bypass the land rights of indigenous people, women, or other vulnerable groups (Bruce 1988; Platteau 1996).

The titling process requires a clear legal basis and a streamlined institutional infrastructure that is capable of administering the process efficiently. Numerous World
Bank projects have either underestimated the complexity of the technical issues involved in titling or assumed that titling could be initiated even if agreement over complex policy issues had not been reached. Many countries have a plethora of institutions, programs, and projects—often with overlapping competencies and responsibilities, contradictory approaches, and high resource requirements—that make it impossible to administer a titling program effectively or to instill confidence in the validity of the titles issued.

The absence of clear property rights increases the costs of land transactions and may drive them into the informal sector, but empirical evidence on the magnitude of this effect is limited, and government regulation of rental and sales markets appears to have been quantitatively more important. For example, in many Eastern European countries, land rental and even sales transactions emerged long before individuals were able to obtain formal land title. By contrast, the threat of expropriation of rented lands in Colombia and Mexico appears to have deterred land rental transactions even with a formal title. Evidence from Mexico suggests that formal individual title is not always necessary to facilitate operation of rental markets. The codification of property rights through proper procedures significantly reduced the transaction costs and increased the amount of land rentals in the market (Olinto, Davis, and Deininger 1999).

**Examples**

In the aftermath of the 1915 revolution in Mexico, about half of the national land area was granted to communities (ejidos) under communal title. Well-intended restrictions to prevent ejido land from falling into the hands of the wealthy proved to be highly inefficient. Although farmers invented ingenious ways to circumvent these restrictions (Heath 1992), commercial credit was difficult to obtain, and the transaction costs imposed by the various restrictions were high, involving, among other things, the threat of loss of land. In areas where nonagricultural opportunities had increased and farmers engaged in seasonal migration, communal tenure became increasingly dysfunctional.

This issue was addressed in 1992 by legislation that lifted the restrictions on transfers of land, subject only to an upper limit, and allowed ejidos to decide on the admission of members and the tenure regime under which they would operate. They can opt for communal tenure, contribute part of the common lands to a corporation or to a joint venture with outsiders, parcel all or part of the land out to members under freehold title, or even convert the whole ejido from communal to freehold tenure. Contrary to some fears, the law has not led to a widespread sell-off and pauperization of the majority of ejido members. Instead, allowing communities for which existing regulations had increasingly become a constraint the option to shift to a different tenure regime increased the owners' flexibility without giving up the core principles (and the advantages, such as the insurance function provided by joint land ownership) associated with communal types of tenure (de Janvry, Gordillo, and Sadoulet 1997).
The second example involves a reorientation of Bolivia's legal framework and overall institutional environment. Bolivia had a long history of arbitrary allocation of frontier lands to influential individuals. Corruption within the responsible agencies resulted in double titling and conflicting rights, which created considerable tenure insecurity. Titling was highly arbitrary; the process could take more than 12 years. The need to develop a broad consensus on the development of a more conducive land policy framework was time consuming, however, and required political support at the highest level, including presidential intervention (Munoz and Lavadenz 1997). Two areas of the resulting legislation are of particular interest.

Recognizing the possibility of multiple claims on land, the law established three classes of property rights: private homesteads, communal lands, and land under private freehold title. Homesteads are family residences up to half a hectare in size; so long as there are no conflicting claims, families can receive title to their homesteads merely by showing proof of possession and occupancy in a relatively simple and quick process. Title to communal lands (those that have been continuously used by a community or indigenous group) will be awarded if the community acquires legal status and specifies the by-laws under which it intends to govern internal matters. Private agricultural properties under freehold can be established on the remainder of the lands. In addition, public lands can be allocated free of cost to indigenous communities and landless peasants or, if there is no demand from these groups, sold at market prices through a competitive auction. The legislation declares all land titles that have been acquired illegally to be null and void.

Conflicting claims are adjudicated, land rights are regularized, and a legal cadastre is established by a newly founded institution. Areas to be reformed first are selected on the basis of existing land conflicts and demand for regularization. A desk investigation based on existing registry information and aerial photography is followed by a field investigation that involves all the claimants. The results of this investigation are then cross-checked with the land records and published in the communities to elicit public comments. Once complaints and objections have been attended to, the results are posted publicly for two months. If no further complaints are lodged, the rights are finally registered. Experience thus far suggests that all but a tiny minority of claims can actually be regularized using this process and that large landowners are happy to cede illegally acquired parts of their “property” in exchange for legally recognized title to the rest.

Improving the Functioning of Land Markets

Governments in many developing countries maintain regulations that restrict land use and transfers. In many cases these restrictions have been adopted to avoid an unequal concentration of landholdings under a distorted policy regime or to reduce
the political and fiscal cost of implementing land reform. A review of these policies finds that they have rarely achieved their goals. We examine three main issues: restrictions on land ownership and use; restrictions on land sales and rentals; and interventions to improve the functioning of land markets.

**Land Ownership and Use**

Ceilings on land ownership have been imposed primarily to facilitate the breakup of large farms and the associated sales of land to small producers. Even where such measures have had a strong economic and social justification and where conditions for implementing them should have been favorable, ownership ceilings have had only a marginal impact. In India, for example, 35 years of ceiling laws have, in all except three states, transferred less than 1 percent of the agricultural area to the target group (Appu 1996). Ceiling laws have been expensive to enforce, have imposed costs on landowners who took measures to avoid them, and have generated corruption, tenure insecurity, and red tape.

Such ceilings might be justified as a temporary measure in situations such as Eastern Europe if there are large imperfections in markets for credit, inputs, and outputs and if (new) landowners are ill informed about the productive value of their endowment. In such a situation, measures to reduce the scope for rapid land accumulation by individuals with better market access or information might be justifiable—although a temporary sales moratorium may be a better way to achieve this than ownership ceilings.

Governments may also adopt zoning laws that classify certain land as either agricultural or nonagricultural. In rural areas, zoning land for agricultural use provides tax credits, exemption from assessments for urban services, and protection from nuisance suits and forecloses the option of selling the land as residential property. Zoning is justified if negative externalities need to be reduced by more than the cost of zoning enforcement (Brandão and Feder 1995), but the cost of enforcing zoning regulations that run counter to economic incentives should not be underestimated. Especially if the institutional infrastructure for enforcement is weak, zoning may lead to rent-seeking and corruption that reduce the economic benefits to a point where they become negative.

**Restrictions on Land Rentals and Sales**

Sales restrictions have frequently been imposed on beneficiaries of land reform or on settlers on formerly state-owned land to prevent them from selling or mortgaging their land. These measures could be justified to prevent beneficiaries from taking undue advantage of a land reform program or as a temporary measure to prevent land sales based on a lack of information or in response to imperfections in product
and financial markets. Even temporary restrictions on land sales can be counterpro-
ductive, however, because they prevent landholders from accessing credit when it is
most needed. Hayami and Otsuka (1993) describe a situation in which farmers were
forced to resort to inefficient arrangements such as usufruct mortgaging and wage
labor to gain access to credit. Precluding beneficiaries of land reform from renting or
selling their land is likely to prevent adjustments that reflect the settlers' abilities and
could, if combined with restrictions on rentals, cause large tracts of land to be
underutilized. The goal of preventing small landowners from selling out in response
to temporary shocks would be better served by ensuring that they have access to
output and credit markets and technical assistance and by providing them with safety
nets during disasters to avoid distress sales. A moratorium on land sales might be
justifiable to discourage speculative purchases, but alternatives such as limiting the
amount of land that can be allocated to one individual or requiring that the land be
cultivated before title is granted should be considered instead.

Rental restrictions aimed at eliminating the efficiency losses that are assumed to
be associated with share tenancy are not justified. They should be eliminated because
rental markets are likely to acquire increased importance with economic develop-
ment (in most industrial countries, between 40 and 70 percent of all cultivated agri-
cultural areas is rented rather than owned). As noted above, the efficiency gains from
rental restrictions are likely to be modest even in the most desirable case, and the
danger of less favorable outcomes is high. The historical root of most rental restric-
tions in developing countries is in tenancy reforms that sought to improve the status
and welfare of the tenant farmer by imposing rent ceilings, awarding permanent
rights to tenanted land (subject to landowners' right to retention), and transferring
land ownership to lands not claimed by landowners. The inability to implement
these reforms swiftly has negatively affected the functioning of rental markets. In
most Latin American countries that tried to give tenants secure tenure, landlords
thwarted the reforms by undertaking large-scale evictions or shifting to ranching,
highly mechanized cultivation, or the use of wage labor. In India, tenancy reforms
meant to benefit the poor seem, in the aggregate, to have damaged them. Although
the impact varies by state, tenant evictions associated with tenancy reforms have
cased the rural poor to lose about 30 percent of the total cultivated area, and, by
threatening landowners who lease with the loss of their land, the reforms have com-
pletely undermined land access through rental markets (Appu 1996). Even in coun-
tries such as Egypt and Uganda where tenancy reforms could be implemented, fail-
ure to separate clearly the rights of landowners from those of tenants has led to
overlapping claims to the same piece of land, causing uncertainty and inhibiting
investment. Landowners (who normally are precluded from raising rents) have no
incentive to invest, while tenants' rights cannot normally be used as collateral for
formal credit.
Replacing rental restrictions with a clear regulatory framework for land rental markets could do much to improve agricultural productivity and the welfare of potential tenants. Furthermore, in situations characterized by overlapping property rights resulting from incomplete implementation of tenancy reforms, mechanisms allowing the parties to come to a mutually agreeable solution—whereby one party buys out the other or each party receives full property rights to part of the land—could boost investment and productivity.

**Land Market Interventions**

Many governments now realize that the social benefits from better-functioning land rental markets far outweigh the advantages of most of the restrictions that have historically been imposed on the operation of such markets. The key question, therefore, is to identify the most important impediments to better development of land markets and to sequence their removal in a way that does not jeopardize the poor. Earlier discussions indicated that the key issues are to enhance investment by clarifying property rights and establishing an institutional framework that guarantees the security of these rights; to increase efficiency by facilitating increased transferability of land (use) rights in rental and possibly sales markets; and to improve the integration of land and other (financial) markets. In addition, governments can consider imposing a land tax and establishing land information systems.

A land tax that is enforced at the municipal level not only could provide an incentive to large landowners to utilize their land more productively but could also make an important contribution to decentralization. On the one hand, a land tax is one of the few cases of a lump sum tax where—using asset, rather than production, values—the effective tax rate decreases as the income generated from the land increases, thus encouraging more productive use of the resource. Several countries are currently experimenting with a land tax, either using a flat tax rate as in Nicaragua, or basing land taxes on self-assessed land values as in Chile (Bird 1974). Land taxes have proven very useful in a wide range of urban contexts in developing countries and—if accompanied by appropriate institutions to help with accounting and implementation—should be feasible in rural ones as well. Because the value of land tax revenues in any given municipality is linked to land values, however, potential land tax revenues will obviously be meager in poor and remote rural municipalities. Land taxes therefore cannot redistribute wealth from rich to poor neighborhoods, which means that local governments will need additional sources of revenue if the interests of horizontal equity are to be served.

Several countries are also attempting to establish market information systems that would reduce transaction costs and improve the availability of information about land prices and markets. These systems would help expand participation in
sales and rental markets and thereby improve the acceptance of land as collateral by financial institutions. Such information systems would also help in developing, fine-tuning, and evaluating the broader framework for land policy, particularly in determining the degree to which distortions continue to apply, who exactly participates in these markets, and whether the interaction between land and credit markets is efficient.

Redistributive Land Reform

Many of the impediments to a smooth functioning of land, labor, and product markets date from the colonial era; because such longstanding barriers maintain a highly unequal distribution of land, large tracts of productive land lie idle, while peasants have to eke out a living on marginal and often environmentally fragile lands. In addition to reducing productivity, unequal land ownership is also linked to social unrest and violence. But the practical difficulties of implementing a land reform program and the ease with which the economic imperatives might be subordinated to political pressure are daunting obstacles.

In the past, instead of aiming to increase productivity and reduce poverty, land reform often aimed at calming social unrest and allaying political pressures by peasant organizations (Horowitz 1993). Even where there was a genuine commitment to breaking the power of landed elites, agrarian reforms were generally designed by urban intellectuals with little idea of the realities of agricultural production and a suspicion that small-scale cultivators could not farm on their own—let alone increase productivity (Barraclough 1970). Moreover, the individuals who were targeted to benefit from these programs were often politically powerful and well-connected rather than those who could make productive use of the land or who were the most deserving on poverty grounds.6

Furthermore, the continued existence of implicit and explicit distortions (for example, the use of land as a tax shelter) raised the costs of land reform by driving land prices above the capitalized value of the agricultural profits the land would produce. Such distortions also reduced the sustainability of land reform and, by encouraging beneficiaries to sell out to large farmers, contributed to a reconcentration of holdings. As noted earlier, attempts to impose legal restrictions often made matters worse. A recent census of Brazilian land reform settlements reported that only about 60 percent of land reform beneficiaries were actually tilling their land.

Finally, rather than improving the way land markets function and using these markets to complement government efforts to redistribute agricultural land, previous programs often aimed to provide substitutes for these markets, resulting in complex regulations that stretched available administrative capacity (Lipton 1974). Centralized government bureaucracies—charged with providing technical assistance and
other support services to beneficiaries—proved to be corrupt, expensive, and ineffective in responding to beneficiary demands.

**Land Banks**

In view of these difficulties, land banks and frontier settlements were seen as alternative mechanisms to land reform. Land banks provide loan financing at commercial rates for small farmers to acquire land, while frontier settlement, or colonization, aims to transfer individuals from congested areas to remote areas where lack of infrastructure means that land is cheap. With hindsight, it can be said that these alternative mechanisms were ineffective. Expecting beneficiaries to repay the full price of land has resulted in widespread default and nonrecoverable loans. Frontier settlement is no longer seen as a way to equalize land distribution. In addition to high administrative costs and associated environmental hazards, it has reinforced, rather than eliminated, unequal land ownership patterns in many countries (Thiesenhusen 1991). Thus most land reforms have relied on expropriation and have been more successful in creating bureaucratic behemoths and in colonizing frontiers than in redistributing land from large to small farmers, although redistributive land reform was shown to have positive social returns.

**New Opportunities for Viable Reform**

The fall in land prices associated with macroeconomic reforms, along with the loss of the privileges that had been conferred on large farms by discriminatory laws, trade protection, and credit subsidies, provides an opportunity to address land reform that is less detrimental to the functioning of markets. Several countries (Brazil, Colombia, Guatemala, the Philippines, and South Africa) are experimenting with a new “community-based” model of land reform. In this instance, the government’s role is limited to providing groups of poor people with technical and financial assistance to buy land in a way that is similar to demand-driven social investment funds. This approach has a number of advantages. First, because there is an upper limit on the amount of the grant, beneficiaries have an incentive to seek run-down, unproductive farms. This approach also aims to replace the confrontational atmosphere that has characterized land reforms with a more collaborative attitude. In fact, because anything that improves the buyer’s productivity is likely to increase the land price, the seller will, in a competitive market, have a strong incentive to help buyers improve the quality of their product—for example, through technical advice and marketing assistance. Second, in a clear departure from the traditional approach, the new model would stimulate, rather than undermine, land markets. Finally, by drawing on the private sector, nongovernmental organizations, and the community to develop, finance, and administer projects, the approach promises to overcome some of the
informational imperfections that have plagued the implementation of land reform by government bureaucracies. This also would help to develop a menu of project options more attuned to the specific needs of different groups within the target population (such as female-headed households).

These programs are too new for their impact on productivity and on the poor to be assessed. But initial evaluations underscore both the potential and the importance of the incentive framework and close monitoring. In South Africa the lack of local government structures, the continued existence of the land subdivision act (which was repealed only recently), and a very centralized and bureaucratic process initially posed limits to private sector participation, the outreach of the program, and the economic viability of the projects. Based on this experience, efforts are under way to reduce the administrative requirements for “livelihood projects” that involve very limited amounts of subsidy, to strengthen incentives for beneficiaries' own contributions, and to decentralize implementation of the entire land reform program. The success of several “share-equity schemes,” where beneficiaries form joint ventures with private investors (including former farm owners), together with evidence from land transactions in the market outside of the program (Graham and Lyne 1999), point toward considerable commercial potential for land reform.

In Colombia evaluations show that the results of a community-based pilot program are clearly superior to those of previous programs and that formerly landless cultivators are able to establish highly productive agricultural operations (Forero 1999). The large size of the grant (70 percent of the land value), however, together with the legal requirement that it be used only to purchase land and not for complementary investments, reduces the economic and fiscal viability of the program. In addition to establishing small areas of perennials and vegetables, where productivity actually exceeds expectations, beneficiaries have purchased large tracts of relatively unproductive (pasture) land that often generates less revenue than is necessary to service interest on the debt (30 percent of the land value) incurred to purchase the land. Changing the program structure to avoid this problem and allowing each beneficiary family to purchase and invest in an area sufficiently large to fully occupy the family’s labor (about 2 hectares) could greatly increase the economic return as well as reduce the fiscal cost.

In Brazil, where individual states sought to increase the pace of land reform, a pilot program to allow market-based acquisition of land by beneficiaries has had impressive results, accomplishing the land reform faster than expected. The new approach is now being implemented nationwide. Because of its decentralized nature, there is ample scope for innovative ways to ensure that the program is targeted to the poor, that it is economically viable, and that it provides incentives for repayment of the land credit, all issues that are of critical importance if the program is to be replicated on a broad scale (Buainain, Da Silveira, and Teófilo 1998; Navarro 1998).
Conclusion

Within the last two decades, considerable advances have been made in understanding the principles underlying land relations and in the way in which they might be affected by specific policy interventions aimed at growth and poverty reduction. At the same time, the number of countries where policymakers believe that the issues surrounding land relations must be addressed has expanded.

It is now recognized that formal title, under conditions of low population density, is not necessarily the most cost-effective and desirable way to ensure secure tenure and facilitate land transfers. One alternative is to award property rights to communities, which then decide on the most suitable tenure arrangements. This system not only should reduce transaction costs but also should allow a more flexible evolution of the structure of property rights while at the same time restoring some of the traditional social functions of land through secondary common property uses. Evaluations of such approaches, which are in increasing demand all over Africa, would be highly desirable. Another option is to award long-term and transferable leases, which could increase investment and expand the scope for using the rental market to transfer land to more productive uses.

Experience shows that the undesirable outcomes that have been attributed to the free operation of land markets were caused more by policy distortions and imperfections in other markets than by the operation of land markets per se. The fact that land sales are more affected than rentals by such factors suggests that the liberalization of rental markets should be a high priority. Indeed, the plethora of land market interventions has greatly reduced opportunities for the poor to rent land. A number of countries inherited a dualistic landownership distribution that is not conducive either to efficiency and investment or to equity and that has often been at the root of violence and protracted social struggle. After macroeconomic liberalization, some of these countries have started to implement a new model of community-based, market-driven land reform. Additional research is needed to determine whether such programs have affected land access, investment, productivity, and social indicators such as violence. The results of that research not only will allow policymakers to make changes as individual programs evolve but also will provide lessons for countries that are struggling to make land policies more effective.

Notes

Klaus Deininger is an economist in the World Bank’s Development Research Group; Hans Binswanger is sector director in the Rural Development and Environment Department in the Africa Region. The authors would like to thank Michael Carter, Alain de Janvry, Dina Umali-Deininger, Ruben Echeverria, Gershon Feder, Gustavo Gordillo de Anda, Isabel Lavadenz, Shem Migot-Adholla, Pedro Olinto,  

Klaus Deininger and Hans Binswanger
Elisabeth Sadoulet, and participants at a World Institute for Development Economics Research (WIDER) workshop in Santiago and various World Bank seminars for detailed comments.

1. The breakdown in collective forms of production occurs because members do not receive the full benefits of increased effort (the free-rider problem), because members' ability to benefit from the collective's assets ends with termination of membership, thus diminishing members' investment incentives, and because there is an incentive to reduce the number of members, often coupled with government subsidies to embark on a capital-intensive development path—implying that collectives generate much less employment than do small (or even large) farms.

2. Farm management and supervisory skills are of importance not only because farmers with better management skills would operate larger farm units but also because they will generally want remuneration for their management comparable to what they could obtain in other sectors of the economy. This leads farm operators to substitute capital for labor as nonagricultural wage rates increase (Kislev and Peterson 1982). Such an increase in farm size over time does not necessarily indicate the presence of increasing returns to scale.

3. A positive relationship between asset distribution and growth is ascertained, for example, by Birdsall and Londono (1997); Deininger and Olinro (1999); Deininger and Squire (1998); Fajnzylber, Lederman, and Loayza (1998); and Rodrik (1998). Besley and Burgess (1998) extend this to land reform legislation.

4. Indeed, some studies have found that in cases where no formal credit markets existed, title had little impact on farm income or investment (Carter and Wiebe 1990; Migot-Adholla and others 1991).

5. As in the case of Mexico, communities can decide to subdivide the communal lands and distribute parcels to individual members under freehold title if they so wish, subject to an upper size limit on the holding of any individual in the group and adherence to proper processes in doing so.

6. This is in line with evidence from a number of Eastern European countries, where political constraints generally led to a relatively "inefficient" way of implementing land reform—through physical restitution of plots rather than compensation of former owners through fungible cash payments.

References

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


Buainain, Antônio Márcio, José Maria da Silveira, and Edson Teófilo. 1998. "Reforma agrária, desenvolvimento e participação: Uma discussão das transformações necessárias e possíveis." Paper
presented at the workshop on Reforma Agrícola e Desenvolvimento Sustentável, Fortaleza, Brazil, November (also at http://www.dataterra.org.br/Documentos/documentos.htm).


McCarthy, Nancy, Alain de Janvry, and Elisabeth Sadoulet. 1998. “Land Allocation under Dual 


Mears, Robin. 1996. “Community, Collective Action, and Common Grazing: The Case of Post-

in Sub-Saharan Africa: A Constraint on Productivity?” The World Bank Economic Review 5(Jan-
uary):155–75.


Processed.

e político-institucionais de seu desenvolvimento recente.” Brasilia. Processed (also at http://


Olinto, Pedro V. 1995. “Land Quality and the Inverse Relationship between Farm Size and Pro-
ductivity: A Panel Data Analysis of Paraguayan Farm Households.” University of Wisconsin, 
Department of Agricultural Economics, Madison. Processed.

Olinto, Pedro V., Benjamin Davis, and Klaus Deininger. 1999. “Did the Poor Benefit from Land 


Prosterman, Roy, Brian Schwarzwalder, and Li Ping. 1998. “China’s Breakthrough 1998 Land Man-
gement Law: Recommendations for Implementing Regulations.” Rural Development Institute, 
Seattle, Wash. Processed.


and the Accumulation of Durable Production Assets in Low-Income Countries: Investments in 


Climate Change, Agriculture, and Developing Countries: Does Adaptation Matter?

Robert Mendelsohn • Ariel Dinar

Because most developing countries depend heavily on agriculture, the effects of global warming on productive croplands are likely to threaten both the welfare of the population and the economic development of the countries. Tropical regions in the developing world are particularly vulnerable to potential damage from environmental changes because the poor soils that cover large areas of these regions already have made much of the land unusable for agriculture.

Although agronomic simulation models predict that higher temperatures will reduce grain yields as the cool wheat-growing areas get warmer, they have not examined the possibility that farmers will adapt by making production decisions that are in their own best interests. A recent set of models examines cross-sectional evidence from India and Brazil and finds that even though the agricultural sector is sensitive to climate, individual farmers do take local climates into account, and their ability to do so will help mitigate the impacts of global warming.

As scientists have become more confident that greenhouse gases will lead to a rise in global temperatures (Houghton and others 1996), developing countries have grown increasingly concerned about the economic impact of climate change on agriculture (Watson and others 1996). Most of the empirical work to date has focused on the industrial countries (Bruce, Lee, and Haites 1996; Reilly and others 1996), and although experts have extrapolated the results of their findings worldwide (see, for example, Fankhauser 1995, Tol 1995, or Pearce and others 1996), little research has focused specifically on the developing nations.

To assess the likely effects of climate change, researchers have pursued three approaches: agronomic models, agroeconomic models, and Ricardian models (which draw on the work by Ricardo showing that land values reflect a site’s productivity).
The agronomic research that applies to developing countries focuses on the vulnerability of farmland in less productive tropical climates (Rosenzweig and Parry 1994; Reilly and others 1996) and on the likelihood that warming will push even more farmland into this zone. Experts are further concerned that small-scale farmers who have very little capital will not be able to pursue the new strategies that will be required to adapt to the change in climate. Unfortunately, agronomic studies have not examined the actual behavior of individual farmers. Farmer responses are strictly hypothetical in these models, although the results demonstrate that adaptation will significantly affect production outcomes (Rosenzweig and Parry 1994). Economists argue that only efficient adaptations should be included in forecasts of climate impacts, claiming that farmers will adopt new methods only if the benefits exceed the costs.

Agronomic models of climate sensitivity indicate that higher temperatures are likely to be harmful in many developing countries where the climate is marginal, water is inadequate, and temperatures are high (Rosenzweig and Parry 1994; Reilly and others 1996). A further increase in temperatures will make many agricultural areas less productive—and some completely unsuitable. In these models, no effort is made to examine the impacts of warming on all crops. Studies of vast territories, such as Rosenzweig and Parry’s, have examined only grains and thus do not take account of crops that prefer tropical climates. Nor, as just noted, have these studies examined efficient adaptation. Thus such research may overestimate the damages from global warming. Their perspective has been limited to arbitrary adaptation measures; even these, however, indicate that adaptation will have a major impact on outcomes.

Research suggests that climate change is not likely to have an adverse effect on agriculture in the United States. The results come from two sources—agroeconomic and Ricardian models. Both approaches find that adaptation by farmers would reduce some of the damages from climate change. Using agroeconomic models to examine farmers' alternatives and determine the most efficient choices under each climate scenario, Kaiser and others (1993a) and Easterling and others (1993) predict that farmers in midwestern regions of the United States will make decisions regarding crops, varieties, and farming practices to mitigate potential reductions in yield. The Ricardian models, which examined a cross-section of farming systems across counties in the United States, found that the effects of higher temperatures ranged from mildly harmful to unequivocally beneficial (Mendelsohn, Nordhaus, and Shaw 1994, 1996). Moreover, substantial evidence from laboratory and field experiments shows that elevated levels of carbon dioxide serve as a fertilizer and can stimulate growth and make plants more drought resistant. A doubling of the amount of carbon dioxide in the atmosphere is predicted to raise crop yields by an average of 30 percent (Reilly and others 1996). Although the exact magnitude of this carbon fertilization is uncertain, the positive outcomes were almost universal. In a study that
incorporates the assumptions of carbon fertilization and farmer adaptation, Adams
and others (1999) suggest that climate warming is likely to increase crop yields in the
United States. These benefits, however, do not necessarily extend to the rest of the
world because the climate in the United States is temperate and would remain tem-
perate even with a little warming and because Americans rely on capital-intensive
agriculture that can adapt to a range of climates. It is not clear whether labor-
intensive agriculture has the same flexibility.

In an approach based on the Ricardian model, Dinar and others (1998) and Sanghi
and Mendelsohn (1999) use a cross-sectional approach to examine the sensitivity of
agriculture in developing countries to changes in climate. By examining the actual
performance of farms across India and Brazil, these studies explore this question
using the Ricardian model first developed for the United States (Mendelsohn,
Nordhaus, and Shaw 1994). By regressing farm performance (land value or net in-
come) on a set of environmental factors, traditional inputs (land and labor), and
support systems (infrastructure), it is possible to measure the contribution of each
factor to the outcome and to detect the effects of long-term climate change on farm
values. Because farmers adjust to their local climates, the cross-sectional method
automatically incorporates farmer adaptation. The results suggest that warming will
do less damage than predicted by the agronomic models (which do not include ad-
aptation). When Ricardian results are compared with agronomic models that do not
include efficient adaptation, however, the results are consistent. It appears that the
Ricardian model does a good job of including efficient adaptation in its predictions.

Methodology

Three techniques have been used to measure the impacts of global warming on agri-
culture: agronomic-economic simulation, agroecological zone analysis, and Ricardian
cross-sectional analysis.

Agronomic-economic models. Agronomic-economic simulation uses a crop model
that has been calibrated from carefully controlled experiments in which the crops are
grown in field or laboratory settings that simulate different climates and levels of
carbon dioxide (Adams, Glyer, and McCarl 1989; Adams and others 1990, 1993,
1999; Easterling and others 1993; Kaiser and others 1993a, b; Rosenzweig and Parry
1994; Kumar and Parikh 1998b). Farming methods are not allowed to vary across
experimental conditions so that all differences in outcomes can be assigned to the
variables that are being tested (temperature, precipitation, or carbon dioxide). The
estimates do not include adaptation. The yields are then entered into economic models
that predict aggregate crop outputs, prices, and net revenue. Because each crop re-
quires extensive experimentation, only the most important crops have been studied.
Thus, almost all of these studies have focused on grains. A notable exception is the
work by Adams and others (1999), which includes citrus fruits and tomatoes along with grains to account for more heat-tolerant crops.

**Agroecological zone analysis.** In this approach, crops are assigned to each agroecological zone and the yields are predicted (FAO 1996). As climate changes, the agroecological zones—and crops—change. By examining these changes, it is possible to predict the effect of alternative climate scenarios on crop yields. The yield changes can then be entered into an economic model that will predict overall supply and market effects (Darwin and others 1995; Darwin forthcoming). The climate scenarios can be relatively simple stories of uniform changes across a country, or they can involve complex geographic distributions of changes. These geographic distributions vary substantially across global climate models. Consequently, most impact studies examine multiple climate scenarios.

**Ricardian models.** The Ricardian cross-sectional approach, which has been used to value the contribution that environmental measures make to farm income, includes work by Mendelsohn, Nordhaus, and Shaw (1994, 1996, 1999); Kumar and Parikh (1998a); Sanghi (1998); Sanghi, Mendelsohn, and Dinar (1998); and Sanghi and Mendelsohn (1999). In all these studies, the countries (Brazil, India, and the United States) are large enough to contain a sample with a wide range of climates. Table 1 presents the range of seasonal temperatures and precipitation in Brazil and India. The range of climates in both countries is large relative to the predicted 1–3.5°C change in temperature within the 21st century (Houghton and others 1996). Many uncertainties cloud the forecast of expected climate change, but the IPCC (Intergovernmental Panel on Climate Change) report (Houghton and others 1996), which includes the most recent research on the topic, is used in this article. By estimating the economic performance of farms across this range of climates, one can

<table>
<thead>
<tr>
<th>Table 1. <strong>Mean and Range of Temperature in India and Brazil</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Season</strong></td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td><strong>India</strong></td>
</tr>
<tr>
<td>January (winter)</td>
</tr>
<tr>
<td>April (spring)</td>
</tr>
<tr>
<td>July (summer)</td>
</tr>
<tr>
<td>October (fall)</td>
</tr>
<tr>
<td><strong>Brazil</strong></td>
</tr>
<tr>
<td>June (winter)</td>
</tr>
<tr>
<td>September (spring)</td>
</tr>
<tr>
<td>December (summer)</td>
</tr>
<tr>
<td>March (fall)</td>
</tr>
</tbody>
</table>

*Note:* All temperatures are Celsius.

measure climate sensitivity in each country. Economic performance is measured using farmland value in the United States and Brazil and annual net income in India.

Advantages and Disadvantages of the Methodologies

Researchers using all three methods generally agree that the extent to which farmers adapt to the new conditions can be very important. Agroeconomic and agroecological models must explicitly model adaptation, however, in order to include it. The analyst must be able to determine which adaptations are economically desirable. In practice, such determinations are difficult to make, and so they have been done largely on an ad hoc basis. The adaptation involves a change in agricultural practices in response to a change in climatic conditions. It includes changes in management practices, such as timing of sowing and harvesting, the intensification of inputs, and changes in the crop mix. Of course, adaptation assumes that farmers have access to alternative practices and technologies that are already practiced elsewhere.

Adaptation and Agronomic Studies

The agronomic literature (which includes the agricultural components of the agronomic-economic approaches and the agroecological zone analyses) addresses adaptation by simulating changes in the growth parameters of various crops according to the latest scientific advances. This approach does not take into account economic considerations and human capital limitations, both of which affect actual farm decisions. Therefore, it is hard to interpret the adaptation scenarios frequently explored by agronomists. El-Shaer and others (1997) identify possible climate-related adaptation strategies for Egyptian agriculture, but they do not provide quantitative estimates of the changes in crop performance associated with these strategies. Kapetanaki and Rosenzweig (1997) identify several adaptation strategies for maize in Greece, including adjusted planting dates and the introduction of new maize varieties. Simulations for three sites suggest that earlier planting dates (10–30 days earlier than the norm) increase yields by nearly 10 percent at all sites. The introduction of new varieties fully mitigated the negative impacts of climate change on yield at one site but only partially at the two southern sites. A combination of earlier planting dates and new varieties completely offset the negative impact of climate change at all sites. Iglesias and Minguez (1997) evaluate several adaptation strategies for wheat and maize in various climatological regions in Spain and find no reductions in yields. The adaptation strategies they tested include combinations of new hybrids, changes in sowing dates, and double cropping, using short-cycle maize varieties as a second crop along with lentils and a vetch–forage barley mixture. This strategy not only reduced the impact of increased temperatures on yields but also
permitted more intensive use of water and land. In Spain water efficiency improved by 1–10 percent in southern regions and 40–80 percent in northern regions. Jin and others (1994), who examine adaptation strategies in southern China, find that a new rice cultivar increased yields at five out of seven sites. Changing the planting dates of the currently used cultivars increased rice yields at the northern sites, but not at the southern sites. Combining both changes—the cultivars and the planting dates—significantly increased yields at six of the seven sites.

Schimmelpfennig and others (1996) and Lewandrowski and Schimmelpfennig (1999) review the literature that examines how farmers adapt to climate changes in the United States. For example, Kaiser and others (1993a, b) include adaptation practices such as crop mix, crop varieties, sowing and harvesting dates, and water-saving technologies (tilage). Based on a comparison of nearby geographic sites, crop models (for example, Rosenzweig and others 1994) and farm-level models find that adaptation reduced the negative impact of warming on crop yields by up to 50 percent (Kaiser and others 1993a, b; Mount and Li 1994; Reilly 1994, 1995; Reilly and others 1996).

**Pros and Cons of Agroecological Models**

The biggest advantage associated with agroecological zones is that they have been carefully studied and the geographic distribution of the zones in developing countries has been published (FAO 1992). The current models using the zonal approach, however, have many problems. The climate zones represent large temperature categories, so that subtle shifts within a zone have no effect but a small shift from one zone to another has a dramatic consequence. The key measures of productivity have not yet blended soils and climate together; the effect of each is computed independently. Nor is it clear how tightly climate zones can predict which crops should be grown or what their yields will be. The approach is subject to the same limitations as the agroeconomic models in that researchers must explicitly account for adaptation. Finally, the existing application of the method predicts large price changes along with small changes in aggregate supply, suggesting that there may be problems with the calibration of the underlying economic model (Darwin and others 1995; Darwin forthcoming). Although the technique has potential, the available models are currently crude.

**The Ricardian Approach**

The most important advantage of the cross-sectional Ricardian approach is its ability to incorporate the changes that farmers would make to tailor their operations to a warmer climate. Because these adaptations benefit the farmer, there is every reason to expect that they will occur. One of the most crucial adaptation strategies is the
choice of crop. Depending on the effects of a warmer climate, a particular crop will be the optimal choice. Of the three grains shown in figure 1, for example, each is best suited for a specific temperature (and level of precipitation). If the temperature warms, however, wheat yields will fall and net revenues will fall as well. If the farmer switches to corn, net revenues will rise. Optimal crop switching is an important component of measuring the agricultural impact of climate change.

One of the drawbacks of the Ricardian method is that the experiment is not carefully controlled across farms. Farms may vary for many reasons in addition to those incorporated in the variables of interest. To control for this problem, the studies include other important variables such as soil quality, market access, and solar radiation. But it is rarely possible to get perfect measures of all these variables, and one cannot guarantee that all of these processes and interactions have been taken into account; some may not be measured at all. Paradoxically, this weakness is a strength of the agronomic model, which relies on carefully controlled experiments and thus does not fall prey to this problem of extraneous variables.

**Figure 1. Crop Choice: Adapting to a Warmer Climate**

Crop revenue

Source: Authors' calculations.
Another valid criticism of the Ricardian approach is that it does not consider price variations; all farms face the same prices. The models have consequently been forced to assume that prices are constant, leading to a bias in the welfare calculations (Cline 1996). The cross-sectional Ricardian studies measure only the loss to producers. By ignoring the price change that would occur if supply changed, the approach overlooks any loss in consumer surplus and consequently underestimates the damages and overestimates the benefits.

Although it is easy to criticize these studies for assuming that prices are constant, including price effects is difficult in any method. In most cases prices are determined in a global market. A global model would be needed to predict how climate changes would affect crop yields. Unfortunately, we do not have accurate global crop models, so predicting what will happen to the global supply of any crop as a result of a change in climate is difficult. Moreover, because the few global analyses completed to date predict that the range of warming in the 21st century should have only a small effect on aggregate supply, the bias from assuming that prices are constant is likely to be small (Reilly, Hohmann, and Kane 1994; Reilly and others 1996). For example, even if aggregate supply changed by 25 percent, the bias from assuming constant prices would be less than 7 percent (Mendelsohn and Nordhaus 1996).

The Ricardian Approach Modified

The application of a cross-sectional approach to agriculture in developing countries raises some additional difficulties that researchers addressed in the Brazilian and Indian studies. Although many prices are constant throughout the sample, some are not. Not only are these prices endogenous, but because they may not be accurate, it is difficult to control for their influence. For example, household members form a large fraction of the agricultural labor pool in developing countries. No wages are paid to household members, nor are there data on the number of hours that family members work (Bennholdt-Thomsen 1982; Grepperud 1997). In Brazil and India, therefore, researchers were forced to control for this factor using a dummy variable that identified those farms that relied heavily on household labor. This dummy variable is difficult to interpret because it signifies unpaid labor, which implies a positive sign on net revenue, but the dummy also signifies a smaller and more marginal farm, which implies a negative sign. Another input that is difficult to price is animal work. In India, for example, there is an official price for a bullock—the purchase price, in other words. But bullocks also have to be fed and managed. Farms that already grow feed may find it cheaper to maintain a bullock than do farms that have to purchase animal feed. To try to control for the price of animal power, the Indian study (Dinar and others 1998) includes the number of bullocks per hectare as a control variable. Although this is an imperfect solution because the number of animals is endogenous,
the researchers believe that it reduces the potential bias animal power may have on the climate coefficients.

Another issue on the input side is technology. Both India and Brazil have mounted large and successful drives to enhance farming technology. These drives tended to be concentrated on the more temperate farmlands in both countries. Farm technology centers were originally concentrated around São Paulo in Brazil and around the Ganges River delta in India. Consequently, it was possible that the technology was facilitating improvement in temperate climates, but not in tropical climates, making warm areas relatively less productive. McKinsey and Evenson (1998) examined this hypothesis for India and found that technology had increased the output of farms over the last two decades, but because technological development had not specifically addressed the problems of heat tolerance, the interaction between technology adoption and climate appeared to be minimal.

Although technological adoption appears to have occurred in all climate zones in India, new technologies can still affect climate sensitivity by giving farmers increased flexibility (Antle 1995). The adoption of modern farming methods can free farmers from previous environmental constraints with new crop varieties, irrigation technologies, and chemical controls (Dinar and Zilberman 1991; Dinar, Campbell, and Zilberman 1992). For example, figure 2 compares the climate sensitivity function for the United States and India as it applies to Indian conditions. Both climate response functions suggest that India is likely to suffer damages because initial temperatures are so high. The results using the Indian response function are more damaging, however, than the results from the American response function. It would appear that capital-intensive agricultural systems are less sensitive to climate, perhaps because they can control so many more inputs. Alternatively, the modern technologies may simply be more able to substitute purchased inputs for climate. Of course, the results also suggest that as developing countries get richer, their agricultural sectors will become less sensitive to changes in climate.

Results of Studies

The agronomic results around the world vary, reflecting alternative methods, starting conditions, and climate scenarios (Reilly and others 1996). Some of the studies included some adaptation and carbon fertilization, whereas other studies ignored both factors. Some of the studies used old climate scenarios that tended to predict large temperature changes, whereas other studies examined the more modest climate scenarios now considered likely. Each of the studies was based on agronomic experiments in selected locations, and so they started from different initial conditions. Some of the studies were based on narrowly defined locations, whereas others ranged over large territories. Although table 2 shows great variation across all regions, some
general observations can be made. In the developing countries studied (China and countries in Africa, South Asia, and Latin America), there are 25 negative and only 6 positive outcomes (of 43 studies). The results for industrial regions (Europe, United States, Japan, and Oceania) are more positive, however, with 9 outright positive results and only 3 negative outcomes (of a total of 27 studies). The agronomic studies suggest that the countries of the temperate and polar zones could gain productivity, whereas developing countries in the subtropical and tropical zones are likely to lose productivity.

If carbon fertilization and adaptation are ignored, agronomic studies of India suggest that extensive warming could cause significant reductions in yields. Grain yields would fall 25–40 percent if temperatures rose by 4°C (Rosenzweig and Parry 1994). These findings are confirmed by Kumar and Parikh (1998a, b), who predict that rice yields will fall by 15–25 percent and wheat yields by 30–35 percent. Not all grains
Table 2. Agroeconomic Results: Change in Yields

<table>
<thead>
<tr>
<th>Region</th>
<th>Crop</th>
<th>Negative</th>
<th>Mixed</th>
<th>Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>Wheat</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Maize</td>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>South Asia</td>
<td>Wheat</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Rice</td>
<td>5</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Maize</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>China</td>
<td>Wheat</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Rice</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Latin America</td>
<td>Wheat</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Maize</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Europe</td>
<td>Wheat</td>
<td>0</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Maize</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>United States</td>
<td>Wheat</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Maize</td>
<td>2</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Japan</td>
<td>Wheat</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Rice</td>
<td>0</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Oceania</td>
<td>Wheat</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Source: Reilly and others (1996).

...are necessarily temperature sensitive; Rao, Rao, and Acharya (1989) find that sorghum and millet are more stable across climates than other grains.

The cross-sectional studies suggest only modest agricultural damage in India (table 3). Using pooled analysis, Sanghi, Mendelsohn, and Dinar (1998) find that a 2°C warming would reduce average net income by only about 4 percent. Using repeat annual analyses, Kumar and Parikh (1998a) determine that a 2°C warming would decrease net income by about 8 percent. Even with a 3.5°C warming, Sanghi,

Table 3. Ricardian Results: Change in Net Income Resulting from a Temperature Increase

<table>
<thead>
<tr>
<th>Country</th>
<th>Temperature increase (Celsius)</th>
<th>Change in income (percent)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td>2.0</td>
<td>-3 to +2</td>
<td>Mendelsohn, Nordhaus, and Shaw (1994)</td>
</tr>
<tr>
<td>United States</td>
<td>2.0</td>
<td>-3 to +2</td>
<td>Mendelsohn, Nordhaus, and Shaw (1996)</td>
</tr>
<tr>
<td>India</td>
<td>2.0</td>
<td>-3 to -6</td>
<td>Sanghi, Mendelsohn, and Dinar (1998)</td>
</tr>
<tr>
<td>India</td>
<td>3.5</td>
<td>-3 to -8</td>
<td>Sanghi, Mendelsohn, and Dinar (1998)</td>
</tr>
<tr>
<td>India</td>
<td>2.0</td>
<td>-7 to -9</td>
<td>Kumar and Parikh (1998a)</td>
</tr>
<tr>
<td>India</td>
<td>3.5</td>
<td>-20 to -26</td>
<td>Kumar and Parikh (1998a)</td>
</tr>
<tr>
<td>Brazil</td>
<td>2.0</td>
<td>-5 to -11</td>
<td>Sanghi (1998)</td>
</tr>
<tr>
<td>Brazil</td>
<td>3.5</td>
<td>-7 to -14</td>
<td>Sanghi (1998)</td>
</tr>
</tbody>
</table>

Note: These estimates do not include carbon fertilization, which is expected to add 30 percent to crop productivity. Climate scenario assumes a 7 percent increase in precipitation.
Mendelsohn, and Dinar find damages of only about 15 percent, while Kumar and Parikh predict damages of about 23 percent. The Ricardian study of Brazil (Sanghi 1998) suggests that land values would fall by about 8 percent with a 2°C warming and by about 11 percent with a 3.5°C warming. These estimates are considerably smaller than the agronomic predictions.

Comparing the damages predicted by the agronomic simulations with the results of the cross-sectional studies provides an estimate of the importance of adaptation. In India, for example, the agronomic approach predicts damages of about 28 percent for severe warming, whereas the cross-sectional results predict damages of between 15 and 23 percent. If this difference is due to adaptation, private adaptation could reduce potential climate damages by between one-fourth and one-half. Note that private adaptation does not involve technical change; farmers simply adjust their techniques using existing technology.

The cross-sectional studies also reveal that climate has important seasonal patterns (table 4). Although using a uniform value to represent the change in climate conditions across various regions may not represent the situation in each region, it provides useful comparable information. For example, it allows differentiation among growing seasons that may have different levels of impact across regions. In the cases

<table>
<thead>
<tr>
<th>Variable</th>
<th>India</th>
<th>India</th>
<th>Brazil</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Temperature</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Winter</td>
<td>-133.0</td>
<td>-95.0</td>
<td>-12.0</td>
</tr>
<tr>
<td></td>
<td>(3.38)</td>
<td>(6.81)</td>
<td>(13.12)</td>
</tr>
<tr>
<td>Spring</td>
<td>-372.0</td>
<td>-174.0</td>
<td>16.3</td>
</tr>
<tr>
<td></td>
<td>(16.71)</td>
<td>(11.96)</td>
<td>(14.62)</td>
</tr>
<tr>
<td>Summer</td>
<td>-103.0</td>
<td>-141.0</td>
<td>-19.4</td>
</tr>
<tr>
<td></td>
<td>(2.84)</td>
<td>(5.20)</td>
<td>(6.62)</td>
</tr>
<tr>
<td>Fall</td>
<td>486.0</td>
<td>458.0</td>
<td>10.1</td>
</tr>
<tr>
<td></td>
<td>(7.35)</td>
<td>(13.06)</td>
<td>(5.95)</td>
</tr>
<tr>
<td><strong>Precipitation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Winter</td>
<td>18.5</td>
<td>7.5</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>(6.11)</td>
<td>(4.39)</td>
<td>(7.71)</td>
</tr>
<tr>
<td>Spring</td>
<td>-14.4</td>
<td>-4.5</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>(8.00)</td>
<td>(8.91)</td>
<td>(0.94)</td>
</tr>
<tr>
<td>Summer</td>
<td>-0.4</td>
<td>-0.4</td>
<td>-0.31</td>
</tr>
<tr>
<td></td>
<td>(2.11)</td>
<td>(3.86)</td>
<td>(15.53)</td>
</tr>
<tr>
<td>Fall</td>
<td>2.3</td>
<td>6.4</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>(2.23)</td>
<td>(10.78)</td>
<td>(11.10)</td>
</tr>
</tbody>
</table>

*Note*: Marginal effects measured at the mean for each season. *t*-statistics are in parentheses.


c. From Sanghi (1998). Dependent variable is the log of farm value.
presented in table 4, for example, it allows a comparison of the impact of a change in climate on agriculture during the four seasons in Brazil and India. Net incomes in India decline precipitously with warmer winter, spring, and summer temperatures, whereas warmer fall temperatures increase net revenues. Land values in Brazil also decline with warmer summer and winter temperatures and rise with warmer falls. These results are similar to patterns found in the U.S. studies. The only seasonal exception is in Brazil, where warmer springs are beneficial. The harmful effects of warmer spring and summer temperatures in India are expected because the temperatures are quite hot already in India during this period. In Brazil, however, a warmer spring may simply extend the growing season. The effect of a warmer fall in all locations is expected to be beneficial, as the higher temperatures help ripen and dry the harvest. The winter temperature effect is more controversial. Some agronomic models ignore winter temperatures because targeted crops are not growing at that time. Farm income may be very sensitive to winter temperatures, however, because cold temperatures help control pests. This can be important even if winter temperatures remain above freezing, as they do in most of India and Brazil. Net revenues are also sensitive to seasonal precipitation, but the effects are smaller and offsetting. Wetter winters are beneficial, but wetter summer and springs are not. In India additional summer rains are not helpful because most of the country already enjoys a monsoon during this period.

The cross-sectional studies reveal that the effect of climate change is not uniform across India. Even if the warming were the same throughout the country, some areas would lose heavily, most would be moderately damaged, and some areas would benefit slightly. Warming would damage the western coastal districts most heavily; districts in several eastern states along the coast would benefit. Interestingly, the desert and marginally dry areas are not very sensitive to warming; productivity in these areas is already so low that additional warming cannot harm them much further.

Policy Implications

Many agronomic studies predict large agricultural losses in developing countries. These estimates appear to be too pessimistic for three reasons. First, the pessimistic results generally do not account for the powerful fertilizing effect of carbon dioxide. Second, these studies tend to underestimate the importance of efficient adaptation as a mechanism to reduce damages. Third, almost all of the agronomic studies focus on grains, which tend to prefer temperate climates, and do not include tropical and subtropical crops. Taking these mitigating factors into account, climate change is not likely to reduce dramatically aggregate productivity in developing countries.

In contrast, global warming is likely to increase productivity in industrial countries in the temperate and polar regions. As these cooler regions become more pro-
The increased supply is likely to depress world prices, making farmers in developing countries worse off. Although these price effects are likely to be small, developing country agriculture will be relatively worse off.

The adaptation measured in the cross-sectional studies entails individual actions, but enlightened public policy could facilitate further changes. First, public policy could help farmers adjust their cropping patterns and methods by monitoring the weather and providing better climate forecasts. Second, the government could advise farmers on how to adjust to alternative climates. Third, the government could invest in new technology by, say, funding research on heat-tolerant crops as an incentive to introduce such crops into warmer climate zones.

Although scientific models provide important insights into the sensitivity of developing country agriculture to changes in temperature, little is known about many regions, especially Africa and Oceania, and there is little information about how warming will affect subsistence farmers. Agricultural GNP in developing countries may not be severely damaged by warming because many developing countries—even those near the equator—have significant pockets of highly productive temperate farmlands. But poor farmers in marginal territories may be very vulnerable. Studies are also vital to the future of other sectors of developing economies, such as timber, energy, water, and coastal properties, all of which may be damaged by warming. Nonmarket factors such as changes in ecosystems, health, and aesthetics need to be considered. As we enter a world where climate is likely to change, it is important that we learn as much as possible about the consequences of warming for the Earth’s systems in order to prepare ourselves and to avoid the most serious impacts.

Notes

Robert Mendelsohn is the Edwin Weyerhaeuser Davis Professor at the Yale School of Forestry and Environmental Studies, and Ariel Dinar is principal economist in the Rural Development Department of the World Bank. This research was funded in part by the Electric Power Research Institute.

1. Although we refer only to “private adaptation,” it is assumed implicitly that the availability of some management and other technologies by which farmers can adapt to climate change may also be the result of public research and development.

2. Variables that were used in the Ricardian models applied to India and Brazil include the following (not all nonclimate variables were used in both countries): temperature and precipitation (annual, monthly, seasonal); various measures of soil property; latitude; labor; machinery; animals; human capital variables; and technology level (measured in shares of high-yielding varieties). Variables measuring infrastructure were hard to obtain at the level at which the analysis was conducted, so proxies were used in the form of population density and distance from market.

References

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


Robert Mendelsohn and Ariel Dinar


———. 1998b. “Climate Change Impacts on Indian Agriculture: Results from a Crop Modeling Approach.” In Dinar and others, eds.


*The World Bank Research Observer,* vol. 14, no. 2 (August 1999)
What Does Climate Change Mean for Agriculture in Developing Countries? A Comment on Mendelsohn and Dinar

John Reilly

Mendelsohn and Dinar review much of the important work on the implications of climate change for agriculture, focusing particularly on developing countries. Their message is that efficient economic adaptation significantly reduces the estimated effects of climate change. Few dispute that some amount of adaptation is likely and that its potential contribution to reducing the negative impacts of global warming is large. One such study (Darwin and others 1995), which analyzed the global impacts using an ecozone (land class) methodology, found that without adaptation, average cereal production yields fell roughly 20 to 30 percent in four different climate scenarios. Through various channels of adaptation (modifying crops and techniques on existing farmland, shifting crops to new land, and responding to changing market prices), these losses were reversed, resulting in small increases in production worldwide (0 to 1 percent) even before considering the positive effects of carbon dioxide (CO₂) fertilization (table 1). Striking, however, are both the initial shock in cereal production in the study reported in table 1 and the range of impacts on yields (without adaptation) estimated by a variety of studies for different sites around the world (shown in table 2).

The Ricardian method reported by Mendelsohn and Dinar and the ecozone (land class) method of Darwin and others (1995) are similar in that they use cross-sectional evidence to estimate the adaptation response to climate change that occurs over time. Darwin and others (1995) use this evidence to estimate productivity shocks that are introduced into a general equilibrium model. As Mendelsohn and Dinar note, the Ricardian method is limited because it does not account for market effects, that is, the fluctuation of prices reflecting market conditions. The result is thus strictly applicable only to a closed economy. Mendelsohn and Dinar note that this bias will be small if the global price effect is small, and they cite a study by Reilly, Hohmann, and Kane (1994) that shows small price effects in some scenarios. This single study is
Table 1. Percentage Changes in the Supply and Production of Cereals for the World

<table>
<thead>
<tr>
<th>Study</th>
<th>No adaptation, land use fixed</th>
<th>On-farm adaptation, land use fixed</th>
<th>On-farm adaptation, market response, land use fixed</th>
<th>On-farm adaptation, market response, land use response</th>
</tr>
</thead>
<tbody>
<tr>
<td>GISS</td>
<td>-22.9</td>
<td>-2.4</td>
<td>0.2</td>
<td>0.9</td>
</tr>
<tr>
<td>GFDL</td>
<td>-23.2</td>
<td>-4.4</td>
<td>-0.6</td>
<td>0.3</td>
</tr>
<tr>
<td>UKMO</td>
<td>-29.6</td>
<td>-6.4</td>
<td>-0.2</td>
<td>1.2</td>
</tr>
<tr>
<td>OSU</td>
<td>-18.8</td>
<td>-3.9</td>
<td>-0.5</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Note: Climate change scenarios from the Goddard Institute for Space Studies (GISS), Geophysical Fluid Dynamics Laboratory (GFDL), United Kingdom Meteorological Office (UKMO), and Oregon State University (OSU) general circulation models (GCMs) that have been logged at the National Center for Atmospheric Research for use by other researchers. These scenarios represent simulated changes in climate when CO₂ levels are doubled in the atmosphere.


Table 2. Impact on Crops of Climate Resulting from a Doubled CO₂ Environment

<table>
<thead>
<tr>
<th>Region</th>
<th>Crop</th>
<th>Yield impact (percent)</th>
<th>Discussion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latin America</td>
<td>Maize</td>
<td>-61 to +62 increase</td>
<td>Argentina, Brazil, Chile, Mexico. Range is across GCM scenarios, with and without the CO₂ effect.</td>
</tr>
<tr>
<td></td>
<td>Wheat</td>
<td>-50 to -5 increase</td>
<td>Argentina, Brazil, Uruguay. Range is across GCM scenarios, with and without the CO₂ effect.</td>
</tr>
<tr>
<td></td>
<td>Soybean</td>
<td>-10 to +40 increase</td>
<td>Brazil. Range is across GCM scenarios, with CO₂ effect.</td>
</tr>
<tr>
<td>Former Soviet Union</td>
<td>Wheat</td>
<td>-19 to +41 increase</td>
<td>Range is across GCM scenarios and region, with CO₂ effect.</td>
</tr>
<tr>
<td></td>
<td>Grain</td>
<td>-14 to +13 increase</td>
<td></td>
</tr>
<tr>
<td>Europe</td>
<td>Maize</td>
<td>-30 to +50 increase</td>
<td>France, Spain, Northern Europe. With adaptation, CO₂ effect. Longer growing season; irrigation efficiency loss; northward shift.</td>
</tr>
<tr>
<td></td>
<td>Wheat</td>
<td>Increase or decrease</td>
<td>France, United Kingdom, Northern Europe. With adaptation, CO₂ effect. Longer growing season; northward shift; greater pest damage; lower risk of crop failure.</td>
</tr>
<tr>
<td></td>
<td>Vegetables</td>
<td>Increase</td>
<td></td>
</tr>
<tr>
<td>North America</td>
<td>Maize</td>
<td>-55 to +62 increase</td>
<td>Canada, United States. Range across GCM scenarios and sites; with/without CO₂ effect.</td>
</tr>
<tr>
<td></td>
<td>Wheat</td>
<td>-100 to +234 increase</td>
<td>United States. Less severe effect or increase in yield when CO₂ effect and adaptation considered.</td>
</tr>
<tr>
<td></td>
<td>Soybean</td>
<td>-96 to +58 increase</td>
<td></td>
</tr>
</tbody>
</table>
Table 2 (continued)

<table>
<thead>
<tr>
<th>Regiona</th>
<th>Crop</th>
<th>Yield impact (percent)</th>
<th>Discussionb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>Maize</td>
<td>-65 to +6</td>
<td>Egypt, Kenya, South Africa, Zimbabwe. With CO₂ effect; range across sites and climate scenarios.</td>
</tr>
<tr>
<td></td>
<td>Millet</td>
<td>-79 to -63</td>
<td>Senegal. Carrying capacity fell 11–38 percent.</td>
</tr>
<tr>
<td></td>
<td>Biomass</td>
<td>Decrease</td>
<td>South Africa; agrozone shifts.</td>
</tr>
<tr>
<td>South Asia</td>
<td>Rice</td>
<td>-22 to +28</td>
<td>Bangladesh, India, Indonesia, Malaysia.</td>
</tr>
<tr>
<td></td>
<td>Maize</td>
<td>-65 to -10</td>
<td>Myanmar, Philippines, Thailand. Range over GCM scenarios and sites, with CO₂ effect; some studies also consider adaptation.</td>
</tr>
<tr>
<td></td>
<td>Wheat</td>
<td>-61 to +67</td>
<td></td>
</tr>
<tr>
<td>Mainland China and Taiwan, China</td>
<td>Rice</td>
<td>-78 to +28</td>
<td>Includes rainfed and irrigated rice. Positive effects in NE and NW China, negative in most of the country. Genetic variation provides scope for adaptation.</td>
</tr>
<tr>
<td>Other Asia and Pacific Rim</td>
<td>Rice</td>
<td>-45 to +30</td>
<td>Japan and Republic of Korea. Range is across GCM scenarios. Generally positive in northern Japan; negative in south.</td>
</tr>
<tr>
<td></td>
<td>Pasture</td>
<td>-1 to +35</td>
<td>Australia and New Zealand. Regional variation.</td>
</tr>
<tr>
<td></td>
<td>Wheat</td>
<td>-41 to +65</td>
<td>Australia and Japan. Wide variation, depending on cultivar.</td>
</tr>
</tbody>
</table>

Note: Except as noted, model results use "equilibrium" scenarios from doubled CO₂ GCM experiments. In these experiments, atmospheric carbon is doubled and the climate model is run for 10 to 15 model years (until the climate stabilizes under the new CO₂ level). The newest generation of climate model experiments, so-called "transient" climate scenarios, attempt to more realistically describe a time path of climate change reflecting gradual increase in CO₂ levels, but these scenarios have only recently become available for crop model analysis.

a. In all regions except Latin America and Other Asia and Pacific Rim, the comments apply to all the crops studied.

b. Indicated here is the basis for the range of crop yield estimates given, including the countries in which site studies were conducted; whether the range is across different sites, different GCM scenarios, or both; whether adaptation was considered; and whether the studies included the direct effect on crops of higher ambient levels of atmospheric CO₂—the so-called "fertilization effect." Apart from its effect on climate, CO₂ has direct physiological effects on plants. Generally, experimental evidence shows that higher levels of ambient CO₂ increase crop yields. The magnitude of this "fertilizer" effect on crops is much debated. As a result, many studies estimate the impact of climate change both with and without this effect.

Source: Summarized from Reilly and others (1996).

not conclusive. Mendelsohn and Dinar, in fact, argue that global crop models are poorly calibrated but still use the price results from those models to support the validity of the Ricardian method.

In fact, unless one estimates the effect of climate change worldwide, there is no obvious bound on how much the world market price for agricultural products can
change (in either direction) and hence no way to determine the direction or mag-
nitude of the bias. If adaptation proves to be as effective as Mendelsohn and Dinar or
Darwin and others (1995) estimate, and if the CO₂ fertilization effect does increase
yields by 10 to 15 percent, the prices of agricultural commodities may, in fact, de-
cline sharply. Although a price decline would certainly be an economic benefit for
consumers, agricultural exporting nations could sustain significant welfare losses.
Reilly, Hohmann, and Kane (1994) make the point that exporting countries bear
the largest per capita losses (among cases with CO₂ fertilization and adaptation that
they examine) under scenarios in which the world prices of agricultural commodities
fall. The point is that nobody has good estimates of the global impact. If the goal is
to provide guidance for individual nations or regions, then results based on the hy-
pothesis that the net global impact will be zero (or at least small) must be treated
with extreme caution.

The list of concerns about using evidence from cross-sectional data to estimate the
impact of time-series phenomena is long. One problem is that of controlling for all
the other phenomena (either included in the estimated relationship but poorly mea-
sured or not included for lack of data) that might be affecting the estimated relation-
ship between climate and agricultural production. Nordhaus (1996), who investi-
gated the relationship between wages and climate to get at the direct value of climate
in people's everyday lives, used sophisticated econometric techniques to obtain bet-
ter estimates of the parameters. The study showed that the impact of global warming
on climate amenities could not be reliably determined. (As used here, climate ame-
nity refers to the value people place on living in a warm and sunny climate rather
than a cold and snowy or hot and humid climate.) The relation between agricultural
productivity and climate in cross-sectional evidence would seem to be much stron-
ger, on the face of it, than the relation between wages and climate. Nevertheless,
more robust measures of the reliability of the statistical estimates would be useful.

A second major concern with cross-sectional evidence is that it represents at best a
long-run equilibrium response. The Ricardian method and similar reduced-form
approaches do not provide much information on how one gets from point A (current
climate and current production practices) to point B (new climate and new
production practices). Darwin and others (1995) provide a bit more insight into the
channels of adaptation by dividing the response into three categories: changes that
occur on the farm, in the market, and in land use. Although these distinctions are
somewhat artificial, they show that farmers are able to adjust even without much
market response and without moving agricultural production to entirely new areas.
Table 2 illustrates, however, that without adaptation, the impacts at individual sites
can be dramatic in both directions. Although tables 1 and 2 are difficult to compare
directly, if one assumes that the overall picture presented in table 2 is roughly consis-
tent with the “no adaptation” column in table 1, it appears that at finer geographic
detail the response can be much more varied. In fact, many of the crop yield esti-
mates in table 2 were part of a study by Rosenzweig and Parry (1994), which, when aggregated to a global estimate, generated reductions in yields that were almost identical to those reached by Darwin and others (1995) in the case of "no adaptation."

In comparison with the reduced-form statistical approaches, agronomic models provide evidence on which technological solutions would increase yields (for instance, more fertilizer, changes in the planting date, new varieties of crops), but they do not offer any insight into whether farmers will actually choose these techniques or even whether these strategies would be economically beneficial responses. Reilly and Schimmelpfennig (forthcoming) point out that the techniques used by most studies maintain hypotheses about whether adaptation will occur autonomously or not. Hence, Mendelsohn and Dinar are concerned that crop response models introduce adaptation in an ad hoc manner, whereas cross-sectional evidence assumes agents will detect the changed climate even in a highly variable environment and will know which adaptations will work. Time-series data can be misleading as well because they capture the response to unexpected weather events, whereas in the process of climate change, agents may learn that some of these events are becoming more or less frequent and thus decide to adapt. If one assumes that dynamics do not matter, as implied by the use of cross-sectional evidence, then adaptation can and should be left to the market. If detection is expected to be difficult and agents need to learn the correct response, then the cross-sectional evidence shows the ultimate potential of adaptation. But public policy actions may be needed to realize this potential fully. If irreversibilities that slow the adaptive response are present, the costs may be greater than those estimated by cross-sectional methods unless or until the climate stops changing. Thus I believe it can be said only that the potential of adaptation is large.

The growing literature reviewed by Mendelsohn and Dinar and presented briefly here raises at least three broad questions. First, how are these estimates to be used—of what policy relevance are they? Second, how certain are researchers of these estimates? Third, given these estimates, what should be done now?

What Is the Policy Relevance of These Estimates?

The research agenda behind much of this climate change work is to develop estimates that clarify the damages associated with increased greenhouse gas emissions and the benefits of reducing emissions, as proposed, for example, under the Kyoto Protocol of the Framework Convention on Climate Change (FCCC 1998). Integrated assessment efforts sometimes represent the problem as a generalized and dynamic cost-benefit analysis, where the benefits of the mitigation policy are the avoided damages to agriculture, coastlines, health, and other sectors (Nordhaus 1998). Most of the estimates focus on climate change associated with the equivalent of doubling
the pre-industrial levels of CO₂ in the atmosphere, with global average temperature changes of 2.5°C to 5.2°C. The low end of this temperature range is not predicted to occur until 2070; the high end is not predicted until well after 2100.

A push to foster adaptation through research on the likely effects of both climate change and adaptations to that change has been growing, for three reasons. First, it may be economically sensible to spend something on adaptation and a bit less on reducing greenhouse gas emissions. Second, if one despairs about reducing emissions, given the costs and difficulties of reaching and enforcing a global agreement, adaptation may be the only defense. Finally, because inertia in earth and energy systems means that several decades of climate change are virtually inevitable, those who are ill prepared to adapt (either to avoid losses or to take advantage of new opportunities) may lose comparative advantage to those who are better prepared. In fact, work by Rosenzweig and Parry (1994), as reported in Reilly and others (1996), shows the paradoxical result that cereal production in developing countries was lower with adaptation than without. This decline occurred because the adaptation response was stronger in the industrial countries. As a result, world prices were lower, agricultural comparative advantage shifted to the industrial countries, and developing countries had less incentive to grow cereal crops. This finding does not mean that adaptation is a bad idea—if developing countries had not adapted at all, the shift would have been greater. It does indicate, however, the danger of basing results on partial equilibrium models or even on market or general equilibrium models of a single nation or region.

The general conclusion that adaptation (to the extent it is economically justified) makes sense is tautological. But the value of the empirical work for identifying particular adaptation options is negligible or nonexistent. First, most of the work assumes that adaptation occurs without intervention from anyone. Researchers figuratively position themselves in low Earth orbit and observe that food continues to be produced and people continue to inhabit the land. The contrast between the results in table 2 and those reviewed by Mendelsohn and Dinar (or between the first and last columns in table 1) suggests that something quite powerful must happen to get from estimated yield losses of 20 percent (or more) to the conclusion that effects are minor or positive for the globe and for most countries. To the extent that one is interested only in adding up the damages, perhaps one can assume that everything that needs to happen will happen. But part of what may need to happen is for other researchers to muck around on farms, in agribusiness, and in government agricultural institutions to help point the direction.

Second, the time frame of 2070–2100 and beyond is irrelevant for decisions today about possible adaptation measures. Most of the capital in agriculture will be replaced several times over in the next 70 years. It would be nonsense to optimize a system today for conditions far in the future and ignore the next three decades. It would be nonsense to optimize for conditions in 2070–2100 when most decisions...
can wait until 2069 or at least 2050, when much better forecasts will be available (if for no other reason than that the conditions in 2050 will already be known). Even where the lifetime of a project is long (for example, a large dam), almost any positive discount rate will make irrelevant to today's decision the question of whether there is water in the river in 2100 or farmers who need it.²

Third, the level of uncertainty in these forecasts is unknown. For any particular country, evidence and other simulations of doubled CO₂ effects suggest that predicted crop yields will vary, in either direction, by up to 100 percent of the nation's average predicted yield under the same scenario. I discuss some of the reasons for this large, subjective assessment of uncertainty later. If the assessment is reasonable, this level of uncertainty poses significant challenges for the development of adaptation strategies. It is extremely dangerous to develop a strategy based on two or three scenarios when so little is known about where these sit within the distribution of possible outcomes.

Fourth, these studies are insufficiently detailed or too incomplete—or both—to be of much guidance. In work using these crop response model results and a fairly detailed food trade model, Reilly, Hohmann, and Kane (1994) showed that in most countries the economic impact had as much or more to do with the effect of climate change on world prices as with the impact of climate on agricultural yields within the country. In fact, net exporters of agricultural commodities generally benefited economically from climate change if world prices rose (climate change was, on net, bad for world production) even if they suffered yield losses. They suffered economic losses if world prices fell regardless of whether the climatic effects on agriculture in the country were positive or negative. The situation for food-importing countries was reversed. The difficulty with the argument in Darwin and others (1995) suggesting that global changes in prices may be small is that their study aggregates agricultural commodities to only three categories—grains, other crops, and livestock—and so cannot begin to investigate realistic changes in comparative advantage in the key export crops that are important for specific countries.

Some of these limitations affect the usefulness of these forecasts even for the global cost-benefit calculus. It would be useful to have uncertainty bounds and to know more than just a few point estimates of impacts 70 to 100 years in the future. The limitations are fatal for adaptation actions other than the most general. It would be more useful to recommend climate monitoring, more research, or better forecasts—but even for these, it is unclear how much more money and effort should be spent. When researchers are forced to come up with robust strategies, the adaptation story is similar to the literature on reductions in emissions. In other words, researchers should look for adaptations that will improve resiliency to existing weather variability—so-called no-regrets adaptations. Even such seemingly innocuous recommendations might go wrong. One might well regret investing heavily in irrigation to reduce vulnerability to drought if climate change means that the river itself will dry up.
Are the Estimates of the Impact of Climate Change Valid?

The body of work referenced here presents a somewhat negative result. Researchers went looking for the impact of climate change, and even under the fairly extreme scenarios of warming that might not occur until after 2100, they found little or no effect. Logically then, if the problems, even in these extreme cases, are so slight, less warming between now and 2070 should have even smaller effects. One would then conclude that it is unnecessary to reduce emissions or do anything else to adapt to climate change.

Are there errors in this logic? At issue is whether these are in fact “extreme” scenarios. There are both socioeconomic and biophysical reasons why these scenarios are extreme only in terms of average surface temperature change. Yet the evidence is that mean changes in temperature have little impact on agriculture production; extremes of temperature, rainfall, and storm events are what cause negative agricultural outcomes.

On this subject the literature offers few strong conclusions. Will tropical storms (hurricanes and cyclones) increase in number or intensity? The hydrological cycle will speed up; will that mean more intense rainfall and more frequent droughts? Will seasonal changes from cold to warm or from wet to dry become more variable? Will the El Nino-Southern Oscillation phases become more intense, or will they remain in one phase for longer periods of time? Will monsoons and other rainfall patterns change their seasonal or geographic pattern? The lack of convincing evidence forecasting these changes does not rule out any chance of their happening. A shift in rainfall patterns of 100 or so miles or by a month or two could lead to far larger changes in precipitation in a particular region than is suggested by the estimated global average changes of 7 to 15 percent. Agricultural studies have largely imposed mean warming and precipitation changes from climate predictions on the climate of today without exploring the implications of the many dimensions of climate that could change.

Are there catastrophes (low probability–high consequence events) that could upset even the global mean estimates? The executive summary of the report of the Intergovernmental Panel on Climate Change (Bruce, Lee, and Haites 1996:5) argues that the “consideration of risk aversion and application of the precautionary principle provide rationales for action beyond no regrets.” The possibility of catastrophic consequences occasionally enters discussions of climate change. The melting of the West Antarctic ice sheet, a runaway greenhouse effect from the release of methane hydrates in permafrost or shallow coastal regions, and changes in the ocean’s conveyor belt are events that have been suggested, and in some cases examined and dismissed, as being highly improbable. If these are real possibilities, there may be adaptation actions that governments could take to minimize the consequences. The strategies that would be helpful will not be understood if researchers persist in considering only the center of the distribution.
The adjustment process and the potential that adjustment could increase costs has not been factored into many of the recent analyses. Thus, while the “dumb farmer” studies referred to by Mendelsohn and Dinar are perhaps overly pessimistic, their recent studies may be overly optimistic. If one could trust that the rate of change in the global climate is an indicator of the rate of change in local climate, then one could comfortably dismiss adjustment costs for market sectors. But a smooth response to climate change can hardly be assumed, particularly in the case of precipitation that can change dramatically for local areas if the storm track changes by 50 or 100 miles. Because no realistic transient climate scenarios have been developed, it is impossible to rule out a pattern of climate change in which local areas are stable for some period of time and then change rapidly over a few years. Such a pattern could impose serious adjustment costs even with accurate forecasting and forward-looking behavior.

With regard to the socioeconomic response, a real question remains about the ability of agents to detect and adapt successfully to climate change, given the huge variability in weather from year to year. Moreover, misguided responses to changes are possible, if not likely. Countries that lose comparative advantage in agricultural exports may erect trade barriers to protect their market share; existing conflicts over water rights within and among countries may prevent the efficient allocation of water if it becomes more scarce; investment in irrigation may expand in areas that should be abandoned; cropping in flood-prone regions may continue if insurance and disaster assistance encourage such behavior or if farmers cannot detect whether the flood is part of the normal weather pattern or a signal of a major shift in the hydrological regime.

The long-run response estimated using cross-sectional evidence essentially assumes that farmers rely on decades of weather records and experience in farming to guide their selection of farming strategies. With climate change, this historical experience is no longer automatically relevant. If the signal is simple and clear—gradual warming—farmers can look to nearby warmer regions for guidance. But because patterns of rainfall, temperature, storms, and extreme events over the season are more important than mean changes, the weather record and the farming experience of nearby warmer regions will not be relevant unless climate change involves the wholesale shifting of climate with all moments of the distribution and patterns of extreme events intact.

What Should Be Done Now?

Researchers like myself who have been looking at this subject for nearly 20 years sometimes forget about the air of unreality that taints the discussion of climate change. It is easy to seize on one or two scenarios for a period 100 years hence and overinterpret the predictive content of the estimates. The best remedy for this lack of reality is to
think seriously about what should be recommended today. It is important to take this question seriously, applying the model used in predicting hurricanes. If the hurricane hits with no evacuation warning, the costs are high. But evacuating millions of people if the hurricane turns away or if the forecasted point of landfall is imprecise is also disruptive and costly. Too many false alarms, and no one will believe the forecast when it is right. Given the large degree of uncertainty in the estimates of the effects of climate change on agriculture, researchers can only wish that their forecasts were more precise. Additional work is needed to clarify whether adaptation will indeed resolve any problems and to explore the full range of ways in which climate could change. A better assessment of the uncertainty involved in the forecast (although not as helpful as a firm prediction) is more helpful than a firm prediction that is wrong.

Notes

John Reilly is an associate director of the Joint Program on the Science and Policy of Global Change at the Massachusetts Institute of Technology.

1. Equivalent doubled CO\textsubscript{2} refers to an additional radiative forcing in the atmosphere as if atmospheric concentrations of CO\textsubscript{2} had doubled. Some of this forcing may be due to other gases such as CH\textsubscript{4} (methane) and N\textsubscript{2}O (nitrous oxide).

2. There are some well-known issues with discounting that have been discussed in the context of climate change (Lind and Schuler 1996). From a normative perspective, discount rate decisions imply judgments about intergenerational equity. If there is growth in per capita income, then any equity criterion that favors the poor should be biased toward higher discount rates to allocate more consumption to the present, poorer generation. From a positive perspective, there are also well-known problems with evaluating a few decisions in the economy (for example, adaptation to climate change) at a different rate than other investments in the economy. Agricultural adaptation to climate change largely involves normal investments farmers make in equipment and machinery rather than the large international commons problem of controlling long-term climate change. Weitzman (1998) makes the compelling case that with uncertainty in the discount rate, “the far distant should be discounted at the lowest possible rate,” although he is hazy about what is the “far distant future” or what would be the “lowest possible discount rate.” If economic stagnation and falling incomes are imaginable, then negative discount rates are possible. He also notes that one must evaluate the problem in the context of the life of the investment. A dam is among the longest-lived projects related to agriculture, with a lifetime of perhaps 50 to 100 years (Reilly 1995). This lifetime is not so different from those of transportation infrastructure, power plants, and major building projects. The Weitzman result of declining discount rates for the far distant future thus may be of some importance for dams, but one would want to apply such a rate consistently across other similarly long-lived investments in the economy. And, under any circumstances, it is useful to know the flow of benefits over the entire life of the investment, rather than just a single year near the end of its life.

References

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


Would you like to have more information about

THE WORLD BANK

Research Observer?

You can view the tables of contents and the abstracts of articles from current and past issues at:


We’re interested in your comments. Please write us at:

Editor, The World Bank Research Observer
The World Bank
1818 H Street, N.W. (MC-4-392)
Washington, DC 20433