The World Bank
Research Observer

INSIDE

Ghazala Mansuri and Vijayendra Rao on Community-Driven Development

- Jock R. Anderson and Gershon Feder on Agricultural Extension's Intentions and Realities

- David Coady, Margaret Grosh, and John Hoddinott on Targeting Outcomes Redux

- Sunita Kikeri and John Vellis on Assessing Privatization

- Julia Devlin and Sheridan Titman on Managing Oil Price Risk

ISSN 0257-3032
www.wbro.oupjournals.org
The World Bank Research Observer is intended for anyone who has a professional interest in development. Observer articles are written to be accessible to nonspecialist readers; contributors examine key issues in development economics, survey the literature and the latest World Bank research, and debate issues of development policy. Articles are reviewed by an editorial board drawn from across the Bank and the international community of economists. Inconsistency with Bank policy is not grounds for rejection.

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

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

Community-Based and -Driven Development: A Critical Review

Ghazala Mansuri • Vijayendra Rao

Community-based and -driven development projects have become an important form of development assistance, with the World Bank's portfolio alone approximating $7 billion. A review of their conceptual foundations and evidence on their efficacy shows that projects that rely on community participation have not been particularly effective at targeting the poor. There is some evidence that such projects create effective community infrastructure, but not a single study establishes a causal relationship between any outcome and participatory elements of a community-based development project. Most such projects are dominated by elites, and both targeting and project quality tend to be markedly worse in more unequal communities. A distinction between potentially “benevolent” forms of elite domination and more pernicious types of capture is likely to be important for understanding project dynamics and outcomes. Several qualitative studies indicate that the sustainability of community-based initiatives depends crucially on an enabling institutional environment, which requires government commitment, and on accountability of leaders to their community to avoid “supply-driven demand-driven” development. External agents strongly influence project success, but facilitators are often poorly trained, particularly in rapidly scaled-up programs. The naive application of complex contextual concepts like participation, social capital, and empowerment is endemic among project implementers and contributes to poor design and implementation. The evidence suggests that community-based and -driven development projects are best undertaken in a context-specific manner, with a long time horizon and with careful and well-designed monitoring and evaluation systems.

Community-based development and its more recent variant, community-driven development, are among the fastest-growing mechanisms for channeling development assistance. Community-based development is an umbrella term for projects that actively include beneficiaries in their design and management, and community-driven development refers to community-based development projects in which...
communities have direct control over key project decisions, including management of investment funds.

By conservative calculations, the World Bank’s lending for such projects has risen from US$325 million in 1996 to $2 billion in 2003—or from $3 billion in 1996 to $7 billion in 2003 when lending for an enabling environment for such projects is included.¹ The World Bank’s Poverty Reduction Strategy Paper Sourcebook (Dongier and others 2001) views community-driven development as a mechanism for enhancing sustainability, improving efficiency and effectiveness, allowing poverty reduction efforts to be taken to scale, making development more inclusive, empowering poor people, building social capital, strengthening governance, and complementing market and public sector activities. Community-driven development is said to achieve all this by reducing information problems (by eliciting development priorities directly from target communities and allowing communities to identify projects and eligible recipients of private benefits), expanding the resources available to the poor (through credit, social funds, capacity building, and occupational training), and strengthening the civic capacities of communities by nurturing organizations that represent them.

The potential gains from community-driven development are large. It has the explicit objective of reversing power relations in a manner that creates agency and voice for poor people, allowing them to have more control over development assistance. This is expected to make the allocation of development funds more responsive to their needs, improve the targeting of poverty programs, make government more responsive, improve the delivery of public goods and services, and strengthen the capabilities of the citizenry to undertake self-initiated development activities.

This vision has become one of the cornerstones of the World Bank’s Comprehensive Development Framework, with its increasing emphasis on empowerment (Dongier and others 2001; Narayan 2002). This vision is not universally shared, however. Skeptics have misgivings about the basic precepts of the approach and more practical concerns with the challenges of implementing such projects. Summers (2001), for example, is concerned that local institutions promoted under the aegis of such projects could undermine democratically elected governments. Harriss (2001), Mosse (2001), Cooke and Kothari (2001), and others have focused on what happens when complex and contextual concepts like community, empowerment and capacity for collective action are applied to the needs of large development projects on tight timelines. Project implementers, whose incentives are often poorly aligned with the needs of the project, may choose to gloss over differences within target groups that underscore local power structures and to short-change the more difficult task of institution building in favor of more easily deliverable and measurable outcomes.

Many critics note that evidence on community-driven development initiatives lags well behind the rate at which projects are being implemented and scaled up.
However, the diversity of views and the intensity of their expression make a review of the available evidence both necessary and timely. Because of the considerable overlap between community-based and community-driven development projects, evaluation evidence is reviewed for any project with community participation as a crucial element of its design. Using this broader definition, there is enough credible research to glean some useful insights about specific facets of these programs. That is the purpose of this review.

The article focuses on several questions for which reliable information could be found: Does community participation improve the targeting of private benefits? Are the public goods created by community participation projects better targeted to the poor? Are they of higher quality, or better managed, than similar public goods provided by the government? Does participation lead to the empowerment of marginalized groups—does it lessen exclusion, increase the capacity for collective action, or reduce the possibility that locally powerful elites will capture project benefits? Do the characteristics of external agents—donors, governments, nongovernmental organizations (NGOs), and project facilitators—affect the quality of participation or project success or failure? Finally, can community participation projects be sustainably scaled up?

The literature reviewed includes two main types of studies: impact evaluations, which use statistical or econometric techniques to assess the causal impact of specific project outcomes, and ethnographic or case studies, which use anthropological methods, such as participant observation, in-depth interviews, and focus group discussions. Although case studies cannot attribute impact, they often provide a more nuanced picture of projects in particular contexts and yield insights that can be difficult to generate with quantitative techniques.

The focus is on studies that have either undergone the test of peer-reviewed publication, or have been conducted by independent researchers. This allows the use of an exogenous rule that improves the quality and reduces the level of potential bias while casting a wide enough net to let in research from a variety of disciplinary perspectives on different types of community-based development projects.

The article looks first at the history of participatory development and the move toward community-based development as a key mechanism for channeling development assistance. The second section examines the literature on participatory development and collective action to clarify what is meant by participation and social capital, what constitutes a community, and what are the likely constraints on community participation. The third section reviews the evidence on the effectiveness of community participation and attempts to answer some of the questions raised. The fourth section focuses on the feasibility of sustainably scaling up projects, and the final section highlights the main findings and identifies the gaps in knowledge.
Participatory Development and Development Assistance

There is a long history of community-based forms of development. Clearly significant were the cooperative movement and Gandhian (Gandhi 1962) notions of village self-reliance and small-scale development, which Gandhi saw as an antidote to the corrosive effects of modernization and colonial rule. Another influential perspective was that of Paulo Freire (1970), whose *Pedagogy of the Oppressed* argues that the "oppressed" needed to unite to find a way to improve their own destinies. These ideas led to a first-wave of participatory development in the 1950s, which by 1960 had spread to more than 60 countries in Africa, Asia, and Latin America, largely through the efforts of the U.S. Agency for International Development (USAID; White 1999). Funding for these programs dried up in the early 1960s, and within a decade most were shut down. White (1999:111) notes that one important lesson from the first wave was "fadism" among development agencies, which "build up an approach—to the extent that it is virtually a requirement that a country have such a program in order to receive aid—and then lose interest, leaving the program to collapse."

Economists long remained skeptical of the approach. The early literature on development policy was strongly influenced by the work of Olson (1973) and Russell Hardin (1982) on collective action to achieve a common goal or pursue a common interest. Olson argued that without coercion or some other special device to make individuals act in their common interest, "rational self-interested individuals will not act to achieve their common or group interests." Olson was concerned with "exploitation of the great by the small," because those with smaller interests in a public good would tend to free-ride on the efforts of those with greater interests. Garrett Hardin's (1992) "tragedy of the commons" also acquired metaphorical power—particularly as environmental concerns became more important—but has much broader implications that impinge on a range of economic problems, including the domain of the public and the private, decentralization of power to local governments, and the provision and management of a host of goods and services that are to some degree public or are common-pool resources. Like Hardin, property rights theorists such as Demsetz (1970) and North (1990) argued that common property resources would be overexploited as demand rose unless the commons were enclosed or protected by strong state regulation. This view generated a great deal of pessimism in multilateral development institutions about the viability of local collective action in the provision of public goods and created a strong impetus for state provision of public goods, state regulation of common-pool resources, and an emphasis on development of private property rights.

By the mid-1980s, critics of "big development" were complaining that many large-scale, government-initiated development programs, from schooling to health to credit to irrigation systems, were performing poorly, while rapidly degrading
common-pool resources were creating significant negative environmental and poverty impacts. These complaints reawakened interest in the local management of resources and decisions. The participatory development movement led by Chambers (1983) and others was important in applying these ideas to small-scale development in ways that would allow the poor to be informed participants in development, with external agents acting mainly as facilitators and sources of funds. Further support came from the increasingly strong critique of development from academic social scientists, such as Escobar (1995) and Scott (1998), who argued that top-down perspectives were both disempowering and ineffective. Meanwhile, projects like the Self-Employed Women’s Association in India, the Orangi slum improvement project in Pakistan, and the Iringa Nutrition project in Tanzania were generating interest as highly successful instances of community-driven development (Krishna, Uphoff, and Esman 1997) that could provide important lessons for large donors.

Thinking in mainstream development circles was also significantly affected by the work of Hirschman (1970, 1984), Cernea (1985), and Ostrom (1990). Hirschman’s (1970) notions of “voice” and “exit” provided a way for development practitioners to understand how collective agency could improve well-being, views that were bolstered by Hirschman’s (1984) own attempts to apply these ideas to participatory development. Cernea (1985) showed how large organizations like the World Bank could “put people first” by working systematically at the local level.

Ostrom’s (1990) work on the management of common-pool resources shifted perceptions on the potential for collective action in poor communities. Ostrom argued that what made Olson’s and Hardin’s work most powerful also made it dangerous. She considered the use of these models as a foundation for policy troubling because their results depended on a set of constraints imposed for purposes of analysis. Their relevance in empirical settings, she argued, was an open question rather than a foregone conclusion. In the real world, after all, the capabilities of those involved can be changed, thus altering the constraints themselves. Ostrom and others assembled considerable evidence from case studies to show that endogenous institutions often managed common-pool resources—often very successfully. Thus, Hardin’s “open access” was not the universal model for managing common-pool resources and “remorseless tragedies” were not an inevitable outcome.

Sen’s (1985, 1999) influential effort to shift the focus of development from material well-being to a broad-based “capability” approach also deeply influenced the development community. Central to this approach were strategies to “empower” poor people, an agenda taken on by the World Bank and other donors as part of their response to critiques of top-down development. Arguments for “participatory development,” as advocated by Chambers (1983) and others, led to the inclusion of participation as a crucial means of allowing the poor to have control over decisions.

The inclusion of participatory elements in large-scale development assistance came quickly at the World Bank, in social investment funds (Narayan and Ebbe
1997) and other forms of assistance. Initially focusing on targeting, these projects have moved toward a more holistic attempt at inducing participation through institutions that organize the poor and build their capabilities to act collectively in their own interest (Narayan 2002). The World Bank's (2001) World Development Report 2000/2001: Attacking Poverty focused on empowerment as a key priority of development policy. This has led to a broad-based effort at the World Bank to scale up community-based development and to make it an important element of programs that seek to improve the delivery of public services (World Bank 2003).

With this second wave of interest in community-based approaches to development, a movement that originated in revolutionary goals that were anti-colonial and anti-modernization has been absorbed into the mainstream of development. This transformation has required a new vocabulary to describe its objectives and processes. The next section reviews the key terms and ideas briefly.

Participation, Community, and Social Capital

Community-based development relies on communities to use their social capital to organize themselves and participate in development processes. Thus, concepts such as participation, community, and social capital are critical to how community participation is conceptualized and implemented. Handbooks, guidelines, and terms of reference all use the concepts uncritically, assuming that they are widely and uniformly understood. What each of these concepts implies, however, is quite controversial.

Participation

The cornerstone of community-based development initiatives is the active involvement of members of a defined community in at least some aspects of project design and implementation. Although participation can occur at many levels, a key objective is the incorporation of local knowledge into the project's decision-making processes. When potential beneficiaries also make key project decisions, participation becomes self-initiated action—what has come to be known as the exercise of voice and choice or empowerment. Participation is expected to lead to better designed projects, better targeted benefits, more cost-effective and timely delivery of project inputs, and more equitably distributed project benefits with less corruption and other rent-seeking activity.

This idealized transformative capacity of participation has been challenged on several grounds. First, the exercise of voice and choice can be costly under certain conditions. At the most basic level, it may involve real or imputed financial losses due to the time commitments required for adequate participation. In addition,
participation may lead to psychological or physical duress for the most socially and economically disadvantaged, because genuine participation may require taking positions that are contrary to the interests of powerful groups. Although the premise of participatory approaches is that the potential benefits outweigh such costs, this is by no means certain.

Second, mainstreaming participation has made it an instrument for promoting pragmatic policy interests, such as cost-effective delivery or low-cost maintenance, rather than a vehicle for radical social transformation. This may simply shift some of the costs of service delivery to potential beneficiaries. Indeed, in both Asia (Bowen 1986) and Africa (Ribot 1995), participation has been described as a form of forced or corvée labor, with the poor pressured into making far more substantial contributions than the rich.

Third, the belief that exposure to participatory experiences will transform the attitudes and implementation styles of authoritarian bureaucracies (governments or donors) may be naive. The routinization of participatory planning exercises into the work of public sector implementation agencies puts new pressures on resources while leaving implementers unclear about the potential gain to themselves from this new accountability.

In an examination of several participatory projects, Mosse (2001) finds that even in projects with high levels of participation, what was labeled "local knowledge" was often a construct of the planning context and concealed the underlying politics of knowledge production and use. He identifies four aspects of this. First, participatory exercises are often public events and open-ended regarding target groups and program activities. This makes such events inherently political, and what is reflected is often strongly shaped by local relations of power, authority, and gender. Second, outside agendas get expressed as local knowledge. Project facilitators shape and direct participatory exercises, and villager "needs" are often shaped by perceptions of what the project can deliver. Third, there is local collusion in the planning exercise. People concur in the process of problem definition and planning because it creates space within which they can manipulate the program to serve their own interests. Although this can benefit both project staff and project beneficiaries, it clearly places consensus and action above detailed planning. Fourth, the concept of participation is used to legitimize the project's previously established priorities and the needs of donors to include such processes in their projects. Because such projects have little real support from the community or the project staff, operational demands eventually take over and participatory objectives and goals are sidelined.

The upshot seems to be that a project may deliver many things that both recipients and project implementers consider beneficial, but what can actually be ascribed to participation is unclear. A convincing evaluation would need to validate (or invalidate) the participatory model itself and the relationship between project process and impact.

Ghazala Mansuri and Vijayendra Rao
Community

Participatory projects are typically implemented in a unit referred to as a community. Most of the literature on development policy uses the term *community* without much qualification to denote a culturally and politically homogeneous social system or one that at least implicitly is internally cohesive and more or less harmonious, such as an administratively defined locale (tribal area or neighborhood) or a common interest group (community of weavers or potters). This notion of community is problematic at two levels. First, defining the geographic or conceptual boundaries of a community is not always straightforward. Administrative boundaries can be meaningless where settlement patterns are distinct from such boundaries or where increasing mobility or temporary migrations have transformed community boundaries. In many cases, factional, ethnic, or religious identities may further complicate the picture. Second, an unqualified use of the term often obscures local structures of economic and social power that are likely to strongly influence project outcomes.

Recent studies have shown that the uncritical adoption of the term *community* is particularly problematic for participatory projects that seek to empower people who are excluded or without voice (see, for example, Guijit and Shah 1998; Cooke and Kothari 2001; Mosse 2001). What is labeled a community is often an endogenous construct defined by the parameters of a project, by project facilitators, or by the nature of administrative or identity boundaries rather than an organic form. Also, the effectiveness of participatory strategies may hinge on an explicit understanding of local structures of power, which both limit and enhance prospects for participatory development.

Social Capital

The third key concept in the literature on implementation of community participation projects is *social capital*. This term, which entered the literature on participatory development in Robert Putnam’s (1993) work on northern Italian communities, has strongly affected mainstream thinking on community-based and -driven development (see Woolcock 1998 for a review). It has become common, for example, to refer to such projects as building social capital (Dongier and others 2003), or creating “assets for poor people” (World Bank 2000). Putnam (1993:167) describes social capital as the “features of organization, such as trust, norms, and networks, that can improve the efficiency of society by facilitating coordinated actions.” It refers to the ability of individuals to build “bonds” within their own group and “bridges” to other groups and is deeply tied to the belief that the quality and quantity of group activity are key sources of a community’s strength and its ability to work for its own betterment. Social capital is thus a stock from which people can draw to improve their incomes and which can be “built” to facilitate economic growth and development (Grootaert 1998).
This version of social capital has been criticized on many grounds, among them for not being concerned enough with issues of class distinction and power (Fine 2001; Harriss 2001); for ignoring reverse causality, with the link going from wealth to more group activity rather than the other way (Portes 1998; Durlauf 2001); and for not recognizing that it can be destructive as well as constructive (Portes 1998). The operationalization of Putnam’s ideas has recognized neither the complex strategic, informational, and relational choices that underpin the endogeneity of community formation nor the fact that community is itself an abstract social construct. Policy recommendations thus focus on building communities to facilitate processes of community participation, without recognizing that investing in social capital may not be as straightforward as investing in physical capital and may be considerably more complex than investing in human capital. Returns to investments in social capital can be realized only in concert with other members of a group and are thus likely to be subject to all the concerns about collective action.

The delinking of power and social relationships has been criticized as one of the main problems with the World Bank’s application of Putnamian social capital (Harriss 2001). A more nuanced understanding sees social capital as part of the relations of power within a social system and recognizes that different groups within a social system can have different types of social capital. It also recognizes that social capital must be viewed contextually because it is embedded within structures of power and can be used to facilitate collective action for the common good or to perpetuate symbolic or actual violence against others. The rich may have better internal and external networks than the poor and may use these networks to reproduce unequal systems of domination, as explored in the work of the sociologist Bourdieu (1984, 1990, 1998), whose ideas precede Putnam’s by at least two decades. The poor have less powerful networks that may help them cope with the vicissitudes of life but also restrict their chances for mobility. To the extent that such capital can be bequeathed, it can contribute to the perpetuation of inequality.

Lately, social scientists associated with the World Bank have argued that social capital is less an original theoretical concept and more an umbrella term that has facilitated the insertion of social relations into the thinking of development institutions dominated by economists (Bebbington and others forthcoming). However, now that social capital has made such powerful inroads into development thinking, its value as a so-called Trojan umbrella may have passed and it may be time to bring in all the complexity of thought on society and culture to inform the design of development practice—as the best community participation projects already do. Notions such as trust and norms are not generalizable, and that means that social capital has to be understood within its cultural and political context (Krishna 2002; Rao 2001), with Bourdieu supplanting Putnam as the main theorist in the area. Building the capacity for collective action cannot be divorced from a deep sense of the structures of power within which the poor attempt to
cope (Harriss 2001; Appadurai 2004; Rao and Walton 2004). The question that remains is whether large development organizations can apply these complex notions of the capacity for collective action in the everyday practice of project implementation.

In sum, precisely because community-based and -driven development turns the pyramid of development mechanisms upside down, by giving beneficiaries voice and choice, it cannot ignore the social and cultural context within which beneficiaries live and organize themselves. One possible consequence is that universalistic notions, such as social capital or community, may have to be viewed as deeply contextual and endogenous constructs. This implies that terms such as best practice should be retired to the archives of development and much greater emphasis should be placed on contextualized project design.

**Evidence on the Impact of Community-Driven Development Initiatives**

This section examines the impact of community participation projects, looking at whether they improve targeting, public service delivery, and project sustainability and reduce the risk of capture by elites. It also looks at the role of external agents and the state.

*Does Community Involvement Improve Targeting?*

*Selecting beneficiaries.* Community involvement in selecting the beneficiaries of anti-poverty programs has become increasingly common. Incorporating local knowledge, it is argued, can improve targeting, lower the informational costs of delivering anti-poverty programs, and ensure higher-quality monitoring of program implementation (Chambers 1993; Ostrom, Lam, and Lee 1994; Uphoff 1986; Narayan 1998). But such informational advantages are likely to be realized only when there are institutions and mechanisms to ensure local accountability. Some argue that such institutions are more likely to emerge in societies that are highly mobile, with a tendency toward homogeneous neighborhoods (see, for example, Seabright’s 1996 analysis in the context of political decentralization). Where mobility is low, communities are more likely to reflect social orderings with long histories and deeply entrenched power hierarchies—just where poverty programs are most needed. Consequently, local inequality in relations of power and authority may well allow program benefits to be captured by nontarget groups. In the extreme, the decentralization of poverty programs in such contexts could worsen local inequality and perpetuate local power relations.
Conning and Kevane's (2002) recent review of community-based targeting highlights this potential tradeoff. They find that although community groups are likely to have better information on who the poor are, only communities that have relatively egalitarian preferences, relatively open and transparent systems of decision-making, or clear rules for determining who is poor will tend to be more effective than outside agencies in targeting programs to the poor within those communities. Heterogeneous communities, where people have multiple and conflicting identities, may pose a challenge because of competing incentives. They also note that communities vary in their ability to mobilize information and monitor disbursements. This could affect cost-efficiency and create further opportunities for corruption and capture by elites.

What do evaluations of community-based targeting mechanisms say on this issue? Galasso and Ravallion (forthcoming) examine the targeting performance of a decentralized poverty program in Bangladesh, the Food for Education program. First participating Union Parishads (which typically have about 15 villages) were identified by the center, and then eligible households were identified at the community level. Using both household and community data, Galasso and Ravallion examine how much of the program’s performance in reaching poor families was due to the center’s efforts to reach poor communities and how much to the communities’ efforts to reach their own poor. They also look at the factors that influenced targeting at both levels. They are specifically interested in the role of village institutions and land inequality within villages as determinants of targeting performance.

Although they find a positive targeting differential (the difference in optimal spending on the poor and the nonpoor), with more poor people than nonpoor people receiving the program, the size of the effect is small and most of the targeting differential arises from targeting within villages. They find no evidence that the center is targeting villages at all. The program’s eligibility criteria for selecting participating Union Parishads, for example, cannot account for any of the variation in the targeting differential or in allocations to the poor. If anything, the intervillage component of the targeting differential tends to worsen the overall differential in participating villages. In contrast, they find that structural features of the village are significant predictors of targeting performance across villages. Villages that were more isolated or had higher levels of land inequality tended to have poorer targeting performance. The evidence thus indicates that the performance of decentralized targeting programs can be substantially constrained by local inequality.

Alderman (2002) evaluates an Albanian economic support program (the Ndihme Ekonomika) that provided social assistance to some 20 percent of the population. The program provided a block grant to communes and allowed local officials to determine eligibility and the amount of the transfer to beneficiary households. Alderman finds that local officials were able to target recipients better than the center could have done using proxy entitlement indicators. He surmises that communities were able to improve targeting by using specialized information unavailable to the center.
Other studies suggest that the center can identify poor communities reasonably well, but does much less well at identifying the poor within communities. For example, Coady (2001) finds that PROGRESA, an anti-poverty program in Mexico that selected poor households on the basis of census data without any community involvement, was more effective at targeting poor communities than at targeting poor households within them.

A series of studies on social funds, which have become a popular mechanism for public service delivery, also provide evidence on the targeting performance of anti-poverty programs. In a typical social fund, community infrastructure is built with local participation in the selection and management of facilities. In most cases, community representatives send project proposals to a central public agency, which allocates funds on the basis of such criteria as extent of community involvement, community capacity for collective action, and other factors affecting the feasibility of the proposed project. Social fund rules almost always require that funds be targeted to poor communities, in accordance with specific targeting criteria, and often require some cofinancing from fund recipients. Most social funds restrict the menu of feasible projects to a limited number of public goods, usually schools, clinics, roads, and water and sanitation facilities.

Paxson and Schady (2002) assess poverty targeting in the Peruvian social fund Foncodes using district-level data on expenditures and poverty. They find that the fund, which emphasized geographic targeting, successfully reached the poorest districts but not the poorest households in those districts—better-off households were slightly more likely than poor households to benefit. Using propensity score matching, Chase (2002) compares targeted communities in the Armenia Social Fund with communities that had not received projects but were in the pipeline for them. He finds that although the social fund was targeted to areas with the poorest infrastructure, these were not always the poorest areas. The fund was also slightly regressive in targeting households in rural areas. Like other social funds, the Armenia fund required a community contribution. Chase reports some anecdotal evidence that this may have led to a selection bias against the poorest communities, which are often unwilling or unable to contribute. Using similar techniques to evaluate the Nicaragua Social Fund, Pradhan and Rawlings (2002) find that some social fund investments were well targeted to poor communities and households. The World Bank’s Operations Evaluation Department, in reviewing these and other evaluations of social funds, concludes that although social funds’ geographic targeting has been mildly progressive, social funds were less effective in targeting poor households (World Bank 2002).

In evaluating the success of Argentina’s Trabajar 2 program, introduced in 1997 with World Bank support, Jalan and Ravallion (2003) and Ravallion (2000) find that it significantly expanded an earlier workfare program to provide short-term work to poor households and to locate socially useful projects in poor areas. The central government allocated funds to the provinces, making an effort to provide
more program funding to poorer provinces. Provincial governments then allocated funds to projects within the provinces. Projects were proposed by local governments and NGOs, which bore the nonwage costs of the projects.

Jalan and Ravallion (2003) show that program participants were overwhelmingly drawn from among the poorest households, reflecting successful self-targeting based on the low program wage. Ravallion (2000) finds a marked improvement in central government targeting of poorer provinces. He also finds some improvement in reaching poorer areas within provinces. About a third of the overall improvement in targeting poor areas came from better targeting of provinces, and the rest from better targeting of poor areas within provinces. However, despite higher provincial allocations to poorer provinces, provinces were initially less able to target their poor areas, possibly because wealthier areas were better able to propose and cofinance good projects. In response to this, project monitoring tools were set up that continuously updated targeting performance within provinces. Ravallion (2000) has shown that this simple but powerful tool, which could be put in place as part of regular project monitoring and evaluation, was able to substantially improve targeting of the poor within provinces. This project demonstrates how targeting of poor areas within a decentralized system can be improved by providing stronger incentives from the center for pro-poor targeting by local governments.

Preference targeting. Because one of the theorized advantages of participation is that it makes development demand-driven, participation should improve the match between what a community needs and what it obtains—what can be called preference targeting. There is little reliable evidence on preference targeting, which ideally requires panel data with baseline information on major problems faced by the community and postproject followup in experimental and control communities. Rao and Ibanez (2003) present evidence on the Jamaica Social Investment Fund. Sampled communities were asked to answer retrospective questions on their assessment of major problems in the community before the social fund had been introduced. These authors find that the overall match between the project and the problems identified was poor. Only in two of the five communities studied did the project match the preferences of a majority in that community. Better-educated and better-networked people were more likely to obtain projects that matched their preferences. Thus, overall preference targeting was poor, and preference targeting was worse for the most deprived within the community. Despite that, some 80 percent of respondents were satisfied with the chosen projects. This suggests that despite weak initial targeting, a broad majority of beneficiaries were satisfied with the public good provided.

Mechanisms for identifying who benefits. The literature also assumes that a community's definition of who is poor and vulnerable would be similar to that of social

Glazala Mansuri and Vijayendra Rao
planners. This may not always be the case. Harragin (2004), in a case study of
famine relief efforts in southern Sudan, finds that local notions of how food should
be distributed differed from those of aid workers, resulting in a poorly designed
project. Because what the community perceives as fair and just may be at variance with
the preferences of project supervisors, how should the effectiveness of targeting in
community-based projects be evaluated—whose preferences should count?

Political economy considerations and perverse incentives created by project
performance requirements can also result in poor targeting. Galasso and Ravallion
(forthcoming) and Ravallion (2000) note that a desire to ensure a broad geographic
spread of participants can weaken pro-poor geographic targeting by the center.
Jalan and Ravallion (2003), in their evaluation of the Trabajar 2 program, also note
that social networks were a crucial determinant of who benefited from the workfare
program. They argue that this can be corrected in the design of the program by offer-
ing a wage low enough to discourage wealthier members of the community from
participating.

Schady (2000) shows that community support for then-President Alberto
Fujimori affected the allocation of projects in the Peruvian social fund, Foncodes.
Specifically, districts with a smaller share of pro-Fujimoro votes than in the previous
election received more funding—presumably to increase Fujimoro’s popularity. This
association was strongest in wealthier districts—again suggestive of a political link.
Thus, programs that attempt to elicit community participation and meet commu-
nity demand may be no less immune to political manipulation than strictly top-
down programs.

Morris and others (1999), in their study of the Plandero anti-poverty program to
increase the incomes of the rural poor in western Honduras, find that better-off areas
were the most likely to receive program assistance, and the most deprived areas
were the least likely. They argue that the weak targeting was due to the project’s
implementation schedule, its rate of return criteria, and an evaluation strategy that
emphasized economic results for beneficiary farmers. These jointly created an incen-
tive to select areas that were easily reached—which typically tended to be better-off
areas—and to target project benefits to better-off households within these areas
because they tended to be the most creditworthy and most able to absorb project
funds.

These ethnographic and case studies suggest that the mechanisms used to identify
beneficiaries are crucial in determining how pro-poor decentralized targeting will
be, especially when community members have unequal access to project imple-
menters. The center’s ability to target benefits to poor households may be con-
strained by information gaps or political economy concerns. Although decentralized
targeting can improve outcomes, it does not automatically solve the targeting
problem. Under certain conditions, local inequality can even worsen when targeting
is decentralized.
Does Community Involvement Improve Public Service Delivery?

Most large-scale community-based development projects supported by aid organizations aim to improve access to public services in one of two ways. The project either produces the public good itself or enhances a community’s capacity to act collectively to obtain the public good from other providers, usually through social mobilization activities.

Assessing whether projects that produce public goods with community involvement are more effective than those that supply public goods produced by others, typically by governments, would focus on whether such projects were more consistent with the preferences of target groups, better designed, more sustainable, and delivered public services better. Assessments would ideally compare similar projects developed in these two ways. But few evaluations do this in a completely satisfactory way. Most of the evidence reviewed here compares community-driven development project areas with communities that are otherwise similar but have no projects or have projects of unclear method and provenance. One exception is Khwaja’s (2001) evaluation of community-driven projects funded by the Agha Khan Rural Support Program in northern Pakistan. Khwaja compares a random sample of these projects with other projects in the same village that involved no community participation. Khwaja finds that community-managed projects are better maintained than projects managed by the local government. He also finds that community participation in technical decisions reduced the quality of maintenance, whereas community participation in non-technical decisions significantly improved maintenance.

In a meta-analysis of reports on 52 USAID projects with participatory elements, Finsterbusch and Van Wincklin (1989) conclude that less technically complex and smaller projects were more effective than others. They also find that participatory projects were more effective in more developed economies than in less developed economies—suggesting that the broader institutional environment may play an important supportive role. Although this appears to suggest that institutional reform is required before participatory projects are introduced, the question is actually quite difficult to answer because institutional reform probably makes all types of development initiatives more effective.

Facilities constructed with community involvement tend to improve access to public services. Paxson and Schady (2002) find that the Peruvian social fund Foncodes increased school attendance, particularly for younger children. Chase and Sherburne Benz (2001) reach a similar conclusion for the Zambia Social Fund. They also find that household education expenditures rose where there was a school constructed by the social fund and that use of primary care services and the prevalence of child vaccinations increased where there was a health facility constructed by the social fund.

Newman and others (2002), using both panel data and random assignment approaches to evaluate the Bolivian Social Fund, find a significant reduction in
under-age-five mortality due to the provision of health clinics but little change in education outcomes with the provision of education projects. The reason, they surmise, is that investments in health went beyond providing infrastructure to providing medicine, furniture, and other necessary inputs. Newman and others (2002) also find that water projects improved water quality and access to water only when community-level training was also provided. These findings suggest that to succeed, participatory projects need to go beyond construction of facilities and may require the continuing involvement of external agencies in providing marginal inputs and training.

King and Özler (2000) evaluate the impact of school autonomy on student performance in Nicaragua, comparing participating and nonparticipating schools. Participating public schools acquired de jure responsibility for a range of functions previously managed centrally, including the right to hire and fire the director and maintain the school’s physical and academic quality. School management boards were set up to handle these tasks, drawing their membership from among school staff, parents, and students. King and Özler find that students in schools with both de jure and de facto autonomy performed better on standardized tests than students in schools that did not have autonomy or that had only de jure autonomy. Their work suggests that school effectiveness could be increased by decentralizing certain aspects of school management.

Jimenez and Sawada (1999) report that community-managed schools in El Salvador had fewer absences than comparable centrally managed schools. Jimenez and Paqueo (1996) find that schools in the Philippines that relied more on community contributions used their resources more efficiently. This literature, though not about community-driven development in the strict sense but about the impact of participation on the provision of schooling services, unambiguously supports the idea of increasing school autonomy to improve the quality of public schools.

Although these studies provide a positive impression of the impact of participation on project effectiveness, other studies are more ambiguous. Gugerty and Kremer (2000) use random assignments to evaluate the impact of a participatory program to provide agricultural inputs that was designed to build social capital among women in western Kenya. They find that the costs of inputs were greater than the benefits (increased agricultural output).

Hoddinott and others (2001) look at the relationship between participation and project employment outcomes using data on public works programs from South Africa’s Western Cape Province and attempt to isolate the impact of de jure and de facto participation on each outcome using an instrumental variable strategy. Their instruments for participation are measures of community fractionalization (percentage of adults who are divorced, a measure of racial fractionalization, an index of political fractionalization). They find that participation has no effect on any of their outcome variables. When they interact the participation variable with project type, the effects
of participation are somewhat stronger but still quite weak. However, it is unclear whether their instrument set is valid for the outcome variables they analyze, and their results are also based on a small set of projects.

In addition to impact evaluations, there are other types of studies that define the causal effects less clearly but that are nevertheless interesting. Katz and Sara (1997), who analyze the performance of water systems in a variety of countries, find that performance was markedly better in communities where households were able to make informed choices about the type of system and the level of service they required and where decision-making was democratic and inclusive. Projects constructed without community supervision or management tended to be poorly constructed by private contractors. Where construction was of poor quality despite community participation, the reason was often inadequate technical support—similar to Khwaja’s (2001) findings on communities faring worse on technical decisions. Katz and Sara also report that community members were more willing to pay for investment costs when they had control over the funds. When government staff or contractors controlled the funds, communities viewed their contribution as a tax rather than a fee for service.

Isham and Kähkönen, in analyses of water projects in Indonesia (1999b) and India and Sri Lanka (1999a), confirm that greater community participation is associated with better water supply and that well-designed community-based water services lead to improvements in health outcomes. Differences in project effectiveness are explained largely by a community’s ability to engage in collective action, with high levels of “social capital” improving participation in design and monitoring. This is also the conclusion of Rao and Ibanez (2003), who find that a community’s capacity for collective action influences its ability to successfully apply for funds from the Jamaica Social Investment Fund. Poorly organized communities are not only less likely to obtain projects but are also more likely to mismanage projects that are allocated to them.

Does Community Involvement Improve Project Sustainability or the Capacity for Collective Action?

Few studies have examined the relationship between community-based development projects and capacity for collective action. Finsterbusch and Van Wincklin (1989), in their review of USAID projects, claim without ambiguity that projects with participatory elements increased the overall effectiveness of projects particularly in building capacity for collective action. They make no attempt, however, to identify the causal direction of this claim, which they assert largely on the basis of an anecdotal assessment of project reviews. In a study of incentives for collective action in a district in south India, Wade (1987) reports that some villages had what he calls a “public realm,” with well functioning institutional arrangements for managing
common-pool resources, while others had virtually no public realm. Accounting for the difference, he argues, are economic factors that gave rise to differential collective benefits from organization. This suggests that stimulating participation by forming community organizations such as water user groups may require more than educating people about their common interests or promoting communal values. It may also require helping them to understand the collective benefits of participation.

Rao and Ibanez (2003) use propensity scores and qualitative methods in constructing a survey with a contextualized, project-specific set of variables to explain selection into the program, thereby minimizing unobserved heterogeneity. They then match on observable differences to examine whether the Jamaica Social Investment Fund increased the ability to engage in collective action and build trust. The answer is yes, but the effect was greater for more educated, better networked members of the community. Gugerty and Kremer (2000), in their random assignment study, also find that bringing in outside assistance may change the composition of beneficiary groups. The formation and training of village groups attracted wealthier and more educated men and women into leadership positions within the group because of the outside funding.

Does participation increase project sustainability? Here again the evidence is limited but instructive. Khwaja’s (2001) study suggests that projects managed by communities are more sustainable than projects managed by local governments because of better maintenance. Katz and Sara (1997) and Isham and Kähkönen (1999a, b) also find strong associations between participation and sustainability, although they do not establish causal direction. In an in-depth anthropological study of the sustainability of participatory projects, Kleemeier (2000) examines a Malawi rural piped water project and finds that half the schemes are performing poorly, with the newest ones performing best. The weak sustainability, she argues, stems largely from the lack of institutional support from external agencies—echoing the conclusions of Katz and Sara (1997) and Newman and others (2002).

For water projects in Sub-Saharan Africa, Cleaver (1999) finds that even if communities are initially successful in creating the project, they may lack the material resources and connections to sustain their efforts. Mosse (1997a) comes to similar conclusions in an in-depth examination of tank management in south India. He finds that maintenance of community infrastructure is often crucially dependent on external agents. Thus, the need for a well-functioning state apparatus does not disappear with active community involvement.

Though community participation projects have the potential to be more sustainable than top-down projects, they also appear to suffer from neglect by line ministries once they are completed. Several studies suggest that unless communities can lobby for continuing support for marginal inputs and training, their ability to sustain such projects may be limited.
Economic and Social Heterogeneity and the Risk of Capture by Elites

Previous discussion suggested that a naive understanding of the concepts of community, participation, and social capital can obscure differences that critically influence outcomes. Such differences may be driven by local structures of power; by social divides based on gender, caste, race, or ethnic identities; by uneven interest in the particular public good; or by economic inequality and uneven resource distribution. How do such differences shape outcomes? When is heterogeneity good for collective action, and when does it constrain true participation?

Heterogeneity and inequality. There is a substantial and varied theoretical literature on collective action and coordination by economists, sociologists, and anthropologists that examines the relationship between heterogeneity and the capacity for collective action. The literature identifies several constraints to collective action and the types of environments that favor or disfavor coordination, offering a broad set of hypotheses for empirical work.¹⁰

In a classic study, Olson (1973) argues that certain types of inequality might favor the provision of public goods. In particular, Olson holds that collective action is difficult to mobilize in large homogeneous groups in which no individual could make a significant difference in the level of provision of the good. In contrast, smaller and more unequal groups might do better, although there would be free-riding by those with a smaller interest in the public good.

Several other theoretical studies also suggest that inequality can be conducive to the provision of pure public goods. For example, Bergstrom, Blume, and Varian (1986) demonstrate that when individual preferences over private goods are identical and distributional changes are mean-preserving, a redistribution of wealth can leave the supply of the public good unchanged or even increase it. Itaya, Meza, and Myles (1997) show that income inequality that is so great that only the rich contribute to the public good increases welfare relative to the case in which all individuals contribute. Baland and Platteau (1998) show that inequality in entitlements to resource use is associated with greater conservation when there are decreasing returns to effort for the resource-use technology.

There have been several critiques of Olson’s group-size paradox. Oliver, Marwell, and Teixeira (1985) argue that the probability, extent, and effectiveness of collective action depend on the relation between contributions and levels of the public good and on the extent and type of heterogeneity in the population. They show that inequality of interest can increase the level of collective action. Because inequality of interest is more likely in larger groups, public goods may be provided at higher levels in large groups. However, free-riding remains a problem, and suboptimal levels of collective action and provision of public goods are possible in large groups.
Oliver and colleagues also show that a strong positive correlation between interest and resource heterogeneity increases the probability of collective action.

Heckathorn (1993) argues that the polarizing effect of heterogeneity depends on how collective action is organized. The key issue is that collective action can impose differential costs and benefits on members of a heterogeneous community. If all costs are covered by voluntary contributions, there is no polarization. If the collective action imposes negative net benefits on some members, polarization and conflict are likely.

The literature on collective action related to common-pool resources has also focused considerable attention on heterogeneity. Russell Hardin’s (1982) prisoner’s dilemma outcome for common-pool resources with open access argues that individually rational strategies can lead to collectively irrational outcomes. More recently, Dayton-Johnson and Bardhan (2002) show that the relationship between inequality and levels of collective action in conservation can be U-shaped. In theoretical work based on fisheries, they show that if fishers have earnings opportunities outside the commons that are concave functions of wealth, increased inequality has a negative effect on conservation. However, as inequality increases, levels of conservation can rise and perfect conservation is possible under perfect inequality.

The dominant economic approach to collective action on a common-property resource focuses on individual incentives to contribute, as determined by the private benefits and costs of participation. Baland and Platteau (forthcoming) show that when collective action involves conservation, users’ time preferences are also likely to be important. Specifically, people with a shorter time horizon are more likely to adopt strategies that yield immediate results and neglect longer-term considerations. However, variations in time horizons are likely to be related to the initial distribution of wealth. The poorest users may have wealth levels that are so low that collective action may violate their survival constraint. Thus, the poor may face particular barriers to collective action.

The role of economic inequality is similarly conditioned by other factors. Baland and Platteau (1999) argue that the interactions among multiple dimensions of inequality need to be considered to understand how they affect collective action. For instance, economic inequality can have very different effects when it is combined with caste or ethnic polarization than when it occurs within a more mobile social structure. When regulatory agencies have the power to enforce fines and sanctions, collect fees, and impose rules, inequality is less likely to constrain collective action (Baland and Platteau 1998).

Wade (1987) notes that cultural forces often shape collective action by providing conventions or norms that act as implicit or informal equivalents of all-or-nothing contracts (or assurance games) in resolving collective action dilemmas. He reports that effective village bodies are often composed of elites who can exercise authority with little or no pretense at representation. Thus, he argues, it may be necessary to organize around existing structures of authority, with a major role for village elites.
More recently, Chwe (1999, 2001) builds on this idea of regulatory authority. He argues that most models of collective action implicitly assume the preexistence of "common knowledge." This common knowledge permits games of strategy to be played with a common understanding of the rules of the game—everyone knows what everyone else is playing. The common knowledge assumption is arguably the core concept behind amorphous notions such as trust and social capital that figure prominently in the discourse on collective action. Rao (2003) builds on this to argue that mechanisms for generating common knowledge, which he calls symbolic public goods, are precursors to the possibility of collective action in community development because they may mitigate the effects of heterogeneity. People who belong to the community abide by the rules of the community both because they internalize its ideology and because they face sanctions if they violate the rules. The incentives work not just as exogenous constraints but as what can be called "constraining preferences" (Rao and Walton 2004). Thus, studying the ideology and beliefs underlying collective action is key to understanding how it interacts with inequality and heterogeneity.

Empirical work on heterogeneity and collective action has focused mostly on economic or social (race, caste, ethnicity) heterogeneity. Social heterogeneity is the more difficult concept to measure because it is usually more than a matter of creating an index that weighs the number of identifiable groups in the community by the size of their populations—though most of the economics literature does precisely this.

Alesina and La Ferrara (2000) examine the influence of heterogeneity on participation using survey data on group membership in the United States and data on U.S. localities. They find, after controlling for many individual characteristics, that participation in social activities is significantly lower in more economically unequal or more racially or ethnically fragmented communities. They also show that heterogeneity has the most significant impact on participation in groups where excludability is low and significant interaction among members is necessary. The study attempts to deal with the possible endogeneity of the income inequality measure. It is unclear from their study, however, what the marginal impact is of any given heterogeneity measure. In particular, it is unclear how much racial or ethnic heterogeneity matters once economic inequality has been controlled for.

In a similar study, La Ferrara (2002) looks at the influence of inequality on group participation using data from rural Tanzania. She also finds that higher levels of village inequality reduce the probability of participation in any group. She reports that groups in more unequal communities were less likely to take decisions by vote, were more likely to report misuse of funds and poor group performance, and were more likely to split up into homogeneous ethnic and income groups, interact less frequently, and to less motivated to participate.

In a random assignment evaluation of community-managed schools in Kenya, Gugerty and Miguel (2000) show that ethnic diversity results in fewer social sanctions.
against nonparticipants. This leads to low parental participation in school activities and poor attendance by teachers, which in turn reduce the funds available to the school and worsen school quality.

Bardhan and Dayton-Johnson (2001), in a survey of the evidence on community-driven water projects in India, Mexico, and Nepal, conclude that heterogeneity has a negative impact on cooperation and commons management and weakens the cohesive effect of social norms and the application of sanctions for violating cooperative behavior and collective agreements. Bardhan (2000), examining the factors affecting irrigation maintenance in south India, also finds that inequality has a negative effect. Work by Ostrom, Lam, and Lee (1994) and Ostrom (1990) shows that farmer-managed irrigation schemes had more equitable water distribution. However, these studies can say little about the impact of community heterogeneity on project choice or on the distribution of benefits since they have little to say about nonparticipants.

Not all quantitative studies, however, find that economic inequality has a monotonically negative impact. Somanathan, Prabhakar, and Mehta (2002) analyze the effect of collective action on forest conservation in India using matched-comparison difference-in-difference techniques. They find that communities with more equal land ownership conserved pine forests somewhat better than other communities. However, they find no impact of caste heterogeneity on conservation efforts. They also find that more equal land ownership does not translate into better conservation of broadleaf forests, which are of much greater importance to villagers for firewood and fodder. Some in-depth case studies of participatory projects using qualitative methods also suggest a more complex role for social heterogeneity (see, for example, Vedeld 2000).

Dayton-Johnson (2000) develops a model of cooperation in small irrigation systems and tests the model with data from a survey of Mexican irrigation societies. He finds that social heterogeneity is consistently and significantly associated with lower levels of maintenance. Landholding inequality and the proportional sharing of water in accordance with landholding size also tend to decrease maintenance, but in a nonmonotonic way. Khwaja (2001) finds empirical evidence for a U-shaped relationship between land inequality and project maintenance, with both highly equal and highly unequal communities performing better than moderately equal ones.

Capture by local elites. The frequent tendency for participatory projects to be dominated if not captured by local elites is highlighted by several case studies. Katz and Sara (1997), in a global review of water projects, find numerous cases of project benefits being appropriated by community leaders and little attempt to include households at any stage. They find that training community members was crucial for informed choice and for maintenance of projects and willingness to pay for them. Although well-trained project staff were critical for ensuring inclusiveness and
providing information on options, Katz and Sara note that even well-trained staff are not always effective in overcoming entrenched norms of exclusion. In a study of community forestry projects in India and Nepal that worked reasonably well, Agarwal (2001) reports that women were systematically excluded from the participatory process because of their weak bargaining power. Rao and Ibanez (2003) find that in the participatory projects in their Jamaican case study, wealthier and better-networked individuals dominated decision-making. In a similar case-based evaluation of social funds in Jamaica, Malawi, Nicaragua, and Zambia, the World Bank (2002) Operations Evaluation Department concludes that the process was dominated by “prime movers.”

Abraham and Platteau (2004) present evidence on community participation processes in Sub-Saharan Africa based largely on anecdotal evidence from their work in community-based development and on secondary sources. They argue that rural African communities are often dominated by dictatorial leaders who can shape the participation process to benefit themselves because of the poor flow of information. Thus, participatory development can be very difficult and may require slow, careful efforts to make communities more receptive.

The related body of literature on decentralization also provides insights. Bardhan and Mookherjee (2000) theoretically evaluate the hypothesis that local governments have better information but are less accountable and thus are more prone to capture. They show that the probability of capture increases with local inequality and that it is unlikely that local governments are universally more prone to capture.

Some degree of elite domination may be inevitable in community participation projects, particularly in rural areas, where the elites are often leaders who embody moral and political authority. Often these elites are the only ones who can effectively communicate with outsiders, read project documents, keep accounts and records, and write proposals. This domination may, however, work against the kind of broad-based democratic participation envisioned by advocates of community-driven development because an awareness that project rules have been crafted by the elite may discourage community participation in the project (Bardhan 2000).

Rao and Ibanez (2003) argue, however, that elite domination is not always elite capture. In their case study, they find a potentially more benevolent form of elite domination, with more than 80 percent of beneficiaries ultimately expressing satisfaction with the project. Substantiating this point, Khwaja (2001) finds that participation by hereditary leaders tends to improve maintenance.

Finally, social cohesion may facilitate collective action, but community participation projects do not necessarily perform better in communities with a high score on some index of social capital. Mosse (1997b), in a study of tank management in south India, shows that a traditionally cohesive village will not necessarily have good development projects, whereas communities with low levels of cohesion could have good development projects. Indigenous systems of social organization have very different moral claims, depending on the public good around which they are organized.
A festival or a temple embodies motives and constraints toward collective action that are considerably different from those of a development project, and the logic of social relations in indigenous societies may be quite contrary to the egalitarian principles that a community-based or -driven project is trying to propagate. Abraham and Platteau (2004) make a similar argument for Sub-Saharan Africa. External agencies and programs change incentives and can change the political and social dynamic. Done naively, this can disrupt the social equilibrium by introducing a new element into already contested domains of power and meaning (Mosse 1997b).

**Summary.** The social role of heterogeneity within a group is complex. Heterogeneity can increase or decrease social cooperation and can polarize or strengthen group identity. The success of community-driven development may also be affected by how well heterogeneity is “managed”—by the resources and strategies that are used to bring communities together and by how effectively differences are discussed and solutions achieved. Every time an external agent interacts with a group of people, it creates competition between different interests and incentives. The success of a project may thus depend on how those incentives are aligned—whether by persuasion, ideology, consensus, good governance, domination by greedy elites, or sheer hard work by a group of altruistic individuals.

**The Role of External Agents and the State**

Community initiatives are usually designed by a central authority that sets the basic parameters and the mechanisms for disbursing funds. Project implementers are also extremely important to the effectiveness of projects. Frontline staff who work directly with beneficiary groups are especially critical actors in building participatory processes. They are expected to mobilize communities, build the capacity for collective action, ensure adequate representation and participation, and, where necessary, break through elite domination. They must be culturally and politically sensitive, charismatic leaders, trainers, anthropologists, engineers, economists, and accountants. Despite their centrality, however, there is virtually no generalizable evidence on their role. What evidence there is comes from case studies, which present a somewhat dismal picture.

Jackson’s (1997) analysis of field worker diaries in India indicates that field staff tend to be driven by the incentives they face, which are often not well aligned with the needs of the project. In particular, Jackson notes that field workers tend to gloss over local power relations in a rush to show results. Vasan (2002) shows how facilitators in the forest management projects she examined overlooked project goals when their personal incentives were not aligned with them. Botchway (2001), studying participatory projects in Ghana, notes that project facilitators, who are typically young, inexperienced, and poorly paid, are often vulnerable to manipulation.
by locally powerful elites. Michener (1998) echoes this and points out that senior members of the implementing agency are often former bureaucrats with little experience in community empowerment. Their vision of development, guided by a lifetime of work with line ministries, is closer to "supply driven demand driven development" (Serrano-Berthet 1996; Tendler 2000; Tendler and Serrano-Berthet 1999).

Both beneficiaries and facilitators have an incentive to present the impression of a successful project to outsiders and may collude for this purpose. NGOs often avoid working in difficult communities, where quick results may be harder to demonstrate. The community may also use participation as a bargaining chip to extract resources from the outside agency—and may even construct a version of events that gets them the maximum benefit. All this adversely effects project quality and sustainability.

The state is another strategic actor in this context. Beneficiary communities, often too poor to fund their own teachers, doctors, desks, and medicine (Cleaver 1999; Kleemeier 2000), remain in need of government support for inputs, maintenance investment, and trained staff to sustain project benefits. Thus, the need for a responsive state apparatus may increase when community participation projects are implemented. But some states may manipulate project allocations to satisfy political ends (Schady 2000) or may try to use community participation programs to shift certain costs to community groups. Mosse (1997a) has an extended discussion of this motivation in the context of irrigation infrastructure.

Some studies claim that centralized bureaucracies tend to inhibit project effectiveness, although the evidence is generally weak. Finsterbusch and Van Wincklin (1989), in their review of community-managed USAID projects, hold that for such projects to be effective, the organizations that implement them also need to be decentralized and nonauthoritarian. White (1996) notes that power relations in the wider society within which participation occurs have to be taken into consideration. In some cases, the state may have to support broad-based redistributions of power for community projects to be really successful. This suggests that community-based and -driven development projects must be seen as part of a shift toward a broad-based participatory and decentralized system of governance.

It is unclear, however, how this is to be achieved. Several writers have pointed out the potential for conflict between local political interests and community organizations. Thomass-Slayter (1994) notes that as communities get stronger, they often pose a challenge to local political interests, leading to competitive relations between the state and community organizations and a withdrawal of state support. Das Gupta, Grandvoinnet, and Romani (2000) find that community development efforts can be fragile in such circumstances.

What kind of state creates the right kind of enabling environment for community development efforts? An extreme example is the Indonesian government’s appropriation of the concepts of mutual assistance (gotong royong) and self-reliance (svadaya) as central tenets of its nationalist ideology. During the long period of military rule.
participation was imposed by a strong state on the local population (Bowen 1986). Sukarno tried to use the notions of mutual assistance and self-reliance to unify the diverse groups in the new country and to provide a form of cultural legitimacy to state control. The state had to be strongly authoritarian and development had to proceed in a cooperative and collaborative manner.

Sullivan (1992) in his detailed ethnography of local development in a Javanese community, demonstrates that the combination of an autocratic state and the principles of mutual assistance and self-reliance resulted in a form of forced labor: Being a good Indonesian meant contributing labor and cash for development projects. Collective action was the norm. Grants received by the village headman assumed, in the mismatch between the size of the funds and expected cost of the project, that most funds would be locally mobilized. Contributions from the community were mobilized by the ward leaders. Everyone was expected to contribute free labor—or face social, political, material, and even physical sanctions. There was no choice but to participate.

Ribot (1999) makes a similar argument for how participatory ideologies were implemented in Sahelian forestry. Powers were devolved to village chiefs, who were accountable upward to administrative authorities rather than downward to their constituents. Participation was essentially another form of enforcing central rule. Another mode was the Indian experiment with local democracy, instituted through a series of constitutional amendments enfranchising and enabling the Panchayat system of regularly elected village councils. But although many Panchayats have strong downward accountability, they often lack the resources and the links with higher levels of the state to effectively deliver public services (Matthew and Buch 2000).

Thus, the key to making participation work is to create forms of downward accountability and simultaneously to maintain close links between the higher levels of government and the community—a kind of upward commitment. Participation in the absence of state facilitation can result in a closed village economy, which limits the possibility for improved public action.

Scaling Up Community-Based and -Driven Development

Can community-based and -driven development projects be scaled up? This question cannot be answered unequivocally. There is little generalizable evidence on the optimal pace of scaling up or on the marginal benefits of increasing community involvement. But there are stories of successful initiatives that offer some grounds for optimism. One of the most inspiring is the experience of the Self-Employed Women’s Association (SEWA), inspired by Gandhian notions of self-reliance (Rose 1993). SEWA, which provides inexpensive credit and insurance to poor women workers, has successfully unionized its members, enabling them to negotiate for minimum wages.
legal protection, and other equitable treatment. Although there has never been an impact evaluation of SEWA, it is widely regarded as successful. Similarly, the Amul movement in India has created a cooperative that has been remarkably successful in creating new markets and achieving higher incomes for dairy farmers across India (Kurien 1997).^{12}

**Is Rapid Replication Possible—or Desirable?**

One question is whether such movements can be rapidly replicated through external interventions led by large bureaucracies, as the World Bank and several other donors are attempting to do. One problem is that it is difficult to replicate the success of a highly motivated group of charismatic individuals who are able to sustain a long-term vision of structural transformation through dedication, patience, and creativity. When tasks are handed over to salaried professionals, motivated by more mundane preoccupations such as wages and promotions, incentives change.

Given the pivotal role that facilitators play in the community-based development process, the effort can be doomed from the start, especially when the facilitators work for a large bureaucracy, such as a line ministry or the World Bank (Pritchett and Woolcock 2002). They begin “seeing like a state” (Scott 1998). The complicated process of building community participation becomes routinized and subject to the imperatives of short time horizons and the need for quick results.

Failure is not inevitable. But the mechanical application of best practice guidelines, without attention to local context and adequate monitoring, can easily result in poor community-based projects. There are generalizable lessons from the successes and failures of community-based projects (Binswanger and Aiyer 2003), but the lessons need to be adapted to the historical, political, and social environment where the project is to be implemented.^{13}

Social and cultural inequities create several challenges for the poor. Appadurai (2004) argues that to break structural inequities in social relations and achieve equitable development it is important to build the “capacity [of the poor] to aspire.” This often requires organization to realize collectively what individuals cannot aspire to alone—a way out of the culture of domination and poverty. Rao and Walton (2004) describe this as building “equality of agency,” creating environments to equalize the relational and group-based structures that influence individual aspirations, capabilities, and agency. This is the kind of empowerment that advocates of community-based development envisage.

Can a structurally transforming, aspiration-building, agency-equalizing, community-empowering project result from a rapid scaling-up process? As Trawick (2001) points out, this is unlikely. Community-based development requires the transformation of a social equilibrium in which traditional systems of social organization have evolved to manage resources in a manner that serves the purposes of entrenched
elites. Breaking this up is a slow and gradual process. If external agencies try to change the political and social dynamic without fully understanding it, the social equilibrium can be severely disrupted, with nothing to take its place (Mosse 1997b). This may sometimes be necessary, but it needs to be done with full knowledge of the impact—and it cannot be done quickly. Effective community-based development requires slow, gradual, persistent learning by doing, with a project design that gradually adapts to local conditions by learning from the false starts and mistakes that are endemic to all complex interventions.

Effective learning by doing requires effective monitoring and evaluation. Phillips and Edwards (2000) show that the culture and incentives of large bureaucracies make this difficult. Task managers, facilitators, beneficiaries, and evaluation consultants all have incentives aligned to present a favorable impression, a mindset that hampers learning and encourages the replication of bad design and practice (see Pritchett 2002 for an economists’ version of this argument). The vast majority of community-based interventions, including those funded by the World Bank, have not had reliable evaluations, based on representative samples with treatment and control groups and baseline and follow-up data. If the development community is going to embark on a major change in development strategy, surely it should be concerned about how well it might work.14

Community-based approaches are particularly difficult to get right in the initial stages. Thus, successful community participation projects rely much more than other types of projects on learning by doing, which requires careful evaluations coupled with phased-in scaling up with constant adaptation. Therefore, evaluations should be a central part of the project design from the start and not merely a method to judge the effectiveness of projects after they are completed.

How to Scale Up Community Participation Projects

Community participation projects can be scaled up in several ways. One is to test different approaches and methodologies in a carefully selected set of pilots. Learning from this experiment can then lead to improved project designs applied to a much larger set of communities, in a process of piloted scaling up. Another method is to apply a design in a limited but large set of communities, say, in one province in the country. The lessons from this large test can be applied to projects expanded in a phased manner to the entire country, in a process of phased scaling up. A third possibility is to start with wide national coverage, in a process of untested scaling up. In ideal circumstances, the evidence seems to favor a mix of piloted approaches followed by phased scaling up. But sometimes, untested scaling up is called for—to deal with an economic emergency, for instance. The evidence suggests that untested expansion should be applied with extreme caution, particularly in countries with little experience with community-based development.
What, then, are the preconditions for successfully scaling up community-based projects (also see Binswanger and Aiyer 2003 on this point)? As argued, the process has to be gradual. Because of the contextual complexities involved, initial designs based on best practices are bound to be imperfect. Rapidly scaling these up, particularly in countries with little experience with community-based projects, will likely result in failed projects. A historical, political, and social analysis of the country is critical, in addition to the usual economic analysis, to carefully assess whether the government has the capacity to support community-based development in a manner consistent with both upward commitment and downward accountability.

Second, scaling up requires a strong ethic of learning by doing, with rigorous evaluations and reliable monitoring systems to provide constant feedback. Ideally, both qualitative and quantitative methods should be used to provide reliable estimates of impact and an in-depth examination of context and process (Rao and Woolcock 2003). Lessons should be incorporated into the next phases of project design to correct mistakes, which are inevitable.

Third, careful attention should be paid to the training of a core cadre of facilitators, who are pivotal to successful community-based interventions. Inexperienced facilitators should be given a chance to learn and grow, under the supervision and leadership of more experienced individuals, as part of a gradual learning by doing process.

Fourth, there should be a commitment by the country to a cultural change in the institutional environment, which has to be become more participatory, responsive, and transparent, with downward accountability. These are not just buzzwords. As Uphoff, Esman, and Krishna (1998:202) argue, “If the expansion is occurring because the government or donor sources are promoting it, this is less persuasive than if there is a spontaneous joining of the program or if local governments take over responsibility.” There also has to be a strong upward connection to higher levels of government and line ministries to facilitate the flow of resources.

Finally, changing from top-down to bottom-up development in a manner that is sensitive to local context and culture requires a long time horizon. This also means that community-based development projects should not be judged hastily. Initial evaluations may well be unfavorable. The key is to fix the problems and work toward incremental improvements.

Conclusion

One important question on the effectiveness of community-based and -driven development initiatives is the extent to which they successfully target the poor. Evidence suggests that decentralized targeting has not always been effective, especially in targeting projects to the poor within communities. In addition, the evidence, though
thin, suggests poor preference targeting—the preferences of the poor have not been adequately considered in project selection. Finally, political economy considerations and perverse incentives created by project performance requirements also constrain targeting, although evidence suggests that decentralized targeting can be made more effective by monitoring projects to improve performance incentives.

Another important question is the extent to which participatory development initiatives improve project quality and performance. Here, there is some evidence that participatory projects create effective community infrastructure and improve welfare outcomes, but the evidence does not establish that it is the participatory elements that are responsible for improving project outcomes. Few studies compare community-based projects with centralized mechanisms of service delivery, so it is difficult to tell whether alternate project designs would have produced better outcomes.

Evidence on the impact of economic and social heterogeneity on project outcomes, and on collective action capacity more broadly, suggests that the relationship is complex. Although theoretical work by economists has shown that economic inequality need not constrain collective action, empirical work has shown mixed results. The targeting of poor communities and poor households within communities is markedly worse in more unequal communities, particularly when the distribution of power is concentrated within elites. A number of studies find a U-shaped relationship between inequality and project outcomes. The role of social heterogeneity is more complex to measure. However, most econometric studies that have attempted to devise measures of social fractionalization have shown that fractionalization tends to inhibit collective activity, but there is also qualitative evidence in the opposite direction.

Even in the most egalitarian societies, however, community involvement in choosing, constructing, and managing a public good will almost always be dominated by elites, who tend to be better educated, have fewer opportunity costs on their time, and therefore have the greatest net benefit from participation. It is not clear, however, that this always represents "capture," in the sense of elites appropriating all the benefits from the public good. It may be useful to distinguish between extreme forms of capture, such as outright theft and corruption, and what might be called benevolent capture. However, when local cultures and systems of social organization result in tight control of community decisions by elites, malevolent forms of capture become likely. It is important therefore to understand what types of checks and balances are most effective in reducing capture and the systematic exclusion of the poor and of discriminated-against minorities. The problem in assessing elite capture is that there are no studies that look at an appropriate counterfactual. This remains an important area for future work.

Several case studies suggest that the success of participatory projects may also be affected by how well heterogeneity is managed, by what resources and strategies are
used to bring communities together, and by how effectively differences are debated. The involvement of external agents creates competition among different interests and incentives, and the success of projects may depend on how these incentives are aligned—whether by persuasion, ideology, consensus, good governance, domination by greedy elites, or sheer hard work by a group of altruistic individuals. This is another area where more research would be useful.

The level of community cohesion, or social capital, is also expected to improve the quality and sustainability of projects. Some studies have shown an association between the level of some index of participation and project effectiveness, but the direction of causality is unclear. Though community-based development seems likely to be more effective in more cohesive and better managed communities, evidence also indicates that better-networked or better-educated groups within a community may be better able to organize and thus benefit most from projects. There is virtually no reliable evidence on community participation projects actually increasing a community’s capacity for collective action. This is clearly an area for further research.

Several qualitative studies indicate that the sustainability of community-based initiatives depends crucially on an enabling institutional environment. Line ministries need to be responsive to the needs of communities, and national governments need to be committed to transparent, accountable, and democratic governance, through upward commitment. To avoid “supply-driven demand-driven development,” it is important that community leaders also be downwardly accountable, answerable primarily to beneficiaries rather than to political and bureaucratic superiors.

Qualitative evidence also suggests the importance of external agents, such as project facilitators, to project success. Projects often work with young, inexperienced facilitators whose incentives may not be aligned with the best interests of the community. Knowledge of their impact on the success of projects is limited and requires more investigation. This lack of evidence also relates to the question of how rapidly participatory projects can be scaled up, because rapid scaling up may rely on especially inexperienced facilitators.

Overall, because the success of community-based development is crucially conditioned by local cultural and social systems, projects are best done with careful learning by doing. Although successful projects in any context provide a tremendous learning opportunity, any wholesale application of best practices in unlikely to be useful. In a similar vein, key concepts that underpin community-based initiatives, such as participation, community, and social capital, must be adequately detailed in a context-specific manner. Case study evidence indicates that any naïve application of these notions by project implementers can lead to poor project design and to outcomes that are at odds with the stated intentions of projects.

Finally, it is important to realize that community-based development is not necessarily empowering in practice. A less fervent and more analytical approach by both...
proponents and opponents would be extremely beneficial. This requires a long time horizon and programs that are well monitored (to enable learning from mistakes) and carefully evaluated. Little is known about the impact of community-based projects, largely because most such projects lack careful evaluations with good treatment and control groups and with baseline and follow-up data. This situation urgently needs to be remedied.

Notes

Ghazala Mansuri is economist in the Development Economics Research Group at the World Bank; her e-mail address is gmansuri@worldbank.org. Vijayendra Rao is senior economist in the Development Research Group at the World Bank; his e-mail address is vrao@worldbank.org. The authors are grateful for excellent research assistance from Evan Dennis and Lupin Rahman and valuable comments from Hans Binswanger, Rob Chase, Monica Das Gupta, Gershon Feder, Emanuela Galasso, Keith McLean, David Mosse, Dan Owen, Berk Ozler, Jean-Philippe Platteau, Menno Pradhan, Martin Ravallion, Jesse Ribot, Tara Vishwanath, David Warren, Howard White, Michael Woolcock, three anonymous referees, and several colleagues at seminars and informal discussions at the International Food Policy Research Institute and the World Bank. The article was cosponsored by the Development Economics Research Group and the Social Development Anchor of the World Bank.

1. There is considerable potential for measurement error in these numbers, and there is no authoritative estimate of the World Bank's portfolio of such projects.

2. A key issue in the literature is that the goods or services produced through collective action must be nonexcludable, jointly produced, and costly.

3. White (1999) identifies a second wave in the 1970s and 1980s, initiated by the United Nations system, but it seems more a ripple than a wave because it had little influence on large lending agencies. He calls the current interest in community-driven development "a third-wave which has engulfed the World Bank."

4. In essence, the word community need not describe an object as much as it is describes a concept. It may thus be more useful to see it as an analytical rather than an empirical concept (Gusfield 1975).

5. The word community is typically used in the evaluation literature to indicate different types of target groups or the larger context in which they live. This section often uses the word in an uncritical way because it follows the literature being reviewed.

6. See also Agarwal and Gibson (1999), in the context of the literature on natural resource mobilization.

7. Ravallion and Wodon (1999) show that the program accounted for some 13 percent of primary school enrollment.

8. Note, however, that because the center identified Union Parishads while the analysis is done at the village level, variation in eligibility criteria across villages within Union Parishads may be partly responsible for this result.

9. A long report by Zyl, Sonn, and Costa (2000) on the Northeast Rural Development Program and the Rural Poverty Alleviation Program in northeast Brazil summarizes the results of several studies and finds that the combined impact of these projects was "impressive" and that the projects were well targeted and extremely beneficial to poor communities. Though their methodology and analysis are never clearly outlined, the article suggests that none of the studies had an adequate counterfactual or control group, though they sometimes had a baseline. Thus, it is difficult to judge how "impact" was calculated.

10. It is impossible to do justice here to the rich body of literature on heterogeneity and collective action. For a review of the literature in economics, in the context of environmental resources, see Baland and Platteau (2003).
11. All public goods have the property that many people can use them at once; thus exclusion is difficult. Some public goods yield infinite benefits, but not all. Common-pool resources are public goods with finite or subtractive benefits—they are thus potentially subject to congestion, depletion, or degradation (Blomquist and Ostrom 1985).


13. The World Bank, for instance, tends to borrow the social fund design from activities in Indonesia or Zambia and drop them into very different settings. This is particularly undesirable in the Indonesian case, because of the historical role of a strong state in community development.

14. Several evaluations of community-based development projects have recently begun at the World Bank, and much should be learned from them during the next few years.

References


“What Determines the Effectiveness of Community-Based Water Projects? Evidence from Central Java, Indonesia, on Demand Responsiveness, Service Rules, and Social Capital.”  


Agricultural Extension: Good Intentions and Hard Realities

Jock R. Anderson • Gershon Feder

What considerations lead policymakers to invest in agricultural extension as a key public responsibility, and what factors and agency incentives explain differences in extension system performance? To help answer these questions, this article provides a framework outlining farmers’ demand for information, the public goods character of extension services, and the organizational and political attributes affecting the performance of extension systems. This conceptual framework is used to analyze several extension modalities and their likely and actual effectiveness. The analysis highlights the efficiency gains that can come from locally decentralized delivery systems with incentive structures based on largely private provision, although in most poorer countries extension services will remain publicly funded.

The goals of agricultural extension include transferring information from the global knowledge base and from local research to farmers, enabling them to clarify their own goals and possibilities, educating them on how to make better decisions, and stimulating desirable agricultural development (van der Ban and Hawkins 1996). Thus extension services provide human capital-enhancing inputs, including information flows that can improve rural welfare—an important outcome long recognized in the development dialogue (Leonard 1977; Garforth 1982; Jarrett 1985; Feder, Just, and Zilberman 1986; Roberts 1989). That interest continues in contemporary dialogue, as evident in the workshop on public extension services convened by the World Bank, the U.S. Agency for International Development, and the Neuchatel Group to review recent approaches to revitalizing extension services (World Bank 2002).

Investments in extension services have the potential to improve agricultural productivity and increase farmers’ incomes, especially in developing economies, where more than 90 percent of the world’s nearly 1 million extension personnel are located. Yet the impact of extension on farm performance is varied, reflecting
differences in how extension services are delivered and in the circumstances of
service recipients.

Effective extension involves adequate and timely access by farmers to relevant
advice, with appropriate incentives to adopt the new technology if it suits their
socioeconomic and agroecological circumstances. Critical to adoption are the avail-
ability of improved technology, access to modern inputs and resources, and profit-
ability at an acceptable level of risk. Farmers get information from many sources.
Public extension is one source, but not necessarily the most efficient. Thus, although
extension can improve the productive efficiency of the agricultural sector, the vir-
tues and limitations of alternative mechanisms need to be considered in assessing
the cost-effectiveness of delivering information (Byerlee 1998; van den Ban 1999).

Extension usually has its greatest impact in the early stages of dissemination of
a new technology, when the information disequilibrium (and the productivity differ-
ential) is greatest. As more farmers become aware of the new technology, the impact
of extension diminishes until the need for more information-intensive technologies
arises (Byerlee 1998). The dynamic resolution of the information disequilibria
associated with specific extension messages makes observing the impact of extension
difficult.

The analysis here looks at what leads policymakers to invest in extension services
and what factors and incentives explain differences in extension system perfor-
mance. The following section provides a conceptual framework of farmers' demand
for information, the public goods character of extension services, and the organiza-
tional and political attributes that affect the performance of extension systems.
The second section analyzes several extension modalities and their effectiveness.
The third looks at the methodological issues in assessing extension outcome and
reviews the empirical literature on extension impacts. The final section notes the
political economy implications of the difficulty of attributing outcomes to extension
activities and the importance of internalizing the lessons of experience in the design
of effective extension programs.

Conceptual Framework

Extension, broadly defined, focuses on the delivery of information inputs to farmers.
Information can be of many types, from estimates of future prices for farm products
to new research products, such as improved crop cultivars and knowledge about
how to use particular inputs, such as the timing and intensity of fertilizer use (Byerlee
1998). Farmers have a demand for information and may be prepared to pay for it as
they do for other inputs according to how productive they perceive it to be (Dinar 1996).
Demand for information delivery systems supporting farming should be increasing
if, as agricultural analysts argue, farming is becoming more information-intensive
(Byerlee 1998). How that demand is met varies greatly, depending on market and institutional conditions. Gautam (2000), for instance, concludes that there is significant unmet demand in Kenya for general agricultural extension services. How different types of information are best delivered depends crucially on the nature of the information and the circumstances of the farmer.

A Welfare Economics Context: Extension Services as Public Goods

In considering whether extension services are mainly public or private goods, the usual focus has been on market imperfections, relating in particular to excludability and rivalry (Umali and Schwartz 1994). On this basis, extension services are mixed public and private goods. Some services, such as tailor-made farm management advice, are excludable in that farmers who are not willing to pay for the advice can be excluded from its benefits. Services embodied in commercial products exhibit rivalry in that one farmer’s use reduces availability to others.

Knowledge delivered by extension may be information embodied in products (improved seed, machinery) or it may be more abstract, disembodied information on agricultural practice. There are two broadly applicable types of disembodied agricultural information: general, nonexcludable information (market information or cropping patterns), which tends to be a public good, and specialized, excludable information (fertilizer recommendations for a specific field or farm operation), which tends to be a toll good, with high excludability and low rivalry—some farmers can be excluded from access, even though the value to other users is not diminished by one farmer’s use (Umali-Deininger 1996).

Various mechanisms are available for coordinating the supply of services—private sector markets, public sector hierarchies with state authority, and collective action by civil society (Wolf and Zilberman 2001). Table 1 illustrates the alternative arrangements possible in the financing and provision of extension services, from traditional public sector extension services to fully private services and public-private partnerships. Whether services are best supplied or financed by the private, public, or voluntary sectors or through joint efforts depends on the characteristics of the information service (Schwartz and Zijp 1994; Umali-Deininger 1997).

These observations have several implications (Picciotto and Anderson 1997):

- Information closely associated with market goods (purchased inputs) is generally best left to the private sector.
- Information associated with toll goods can be effectively provided by combined public and private sector efforts.
- Information on the management of common pool goods, with low excludability and high rivalry (forests, common grazing lands, water), is usually best provided by cooperative or voluntary institutions.
Table 1. Some Alternatives for Public-Private Financing and Provision of Extension Services

<table>
<thead>
<tr>
<th>Service Provision</th>
<th>Public</th>
<th>Finance Provision</th>
<th>Other Private</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Traditional extension</td>
<td>Fee-for-service extension</td>
<td>Contracts with public institutions</td>
</tr>
<tr>
<td>Private</td>
<td>Subsidies to extension service providers</td>
<td>Commercial advisory services</td>
<td>Information provided with sale of inputs</td>
</tr>
<tr>
<td></td>
<td>Publicly financed contracts for extension services</td>
<td>Sale of newspapers, magazines, other information products</td>
<td>Extension provided to contract growers</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Advertising in newspapers, radio, television, magazines</td>
</tr>
</tbody>
</table>

Source: Alex, Zijp, and Byerlee 2002.

- Only when market and participation failures are high should the public sector finance information provision—better would be public sector financing of private service delivery.

Reforms have ranged from contracting with the private sector to provide extension services in order to reduce costs and improve cost-effectiveness to drawing on private sector funding to improve the financial sustainability of extension (Beynon and others 1998). The economic rationale for farmers to pay for extension services is generally clear, and the practice is well established in high-income countries (Marsh and Pannell 2000). In developing economies, however, many producers are unable or unwilling to pay for services, in part because they have not seen examples of effective, responsive extension. Many countries have few extension service providers outside the public sector, and few public institutions have the incentives and institutional arrangements in place to encourage program cost recovery.

A Conceptual Framework for Analyzing Public Extension Organizations

Many aspects of extension work have strong public goods characteristics, and public provision of extension services (whether by central or regional governments) has been common in most countries, at least at some stage in their history. Some notable successes have been documented, but so have the many weaknesses that hamper the effectiveness of public extension. A recent worldwide review (Rivera, Qamar, and Crowder 2001:15) refers to extension systems as “failing” and “moribund,” in “disarray or barely functioning at all.” This suggests that there are some generic difficulties in the operation of public extension systems and in the typical bureaucratic-political environment within which they are budgeted and managed.
Feder, Willett, and Zijp (2001) identify eight interrelated characteristics of public extension systems that jointly result in deficient performance, low staff morale, and financial stress. These characteristics provide a framework for analyzing the performance of different levels of extension personnel, the system as a whole, and the underpinnings of different organizational forms and for predicting their likely performance.

**Scale and complexity.** In countries with large numbers of farmers working relatively small plots (as is common in most developing areas), the potential clients of extension services live in geographically dispersed communities. Underdeveloped transport links add to the cost and difficulty of reaching these farmers. High rates of illiteracy and limited connections to electronic mass media rule out reaching these clients through means that do not require face-to-face interaction (written materials, radio, television, the Internet).

Thus, the number of clients who need to be covered by extension services is large, and the cost of reaching them is high. Adding to the challenge, farmers' information needs vary even within a given geographical area because of variations in soil, elevation, microclimate, and farmers' means and capabilities. The large size of the clientele means that only a small number of farmers can interact directly with extension agents. Because direct contacts are rationed, agents often select the farmers they will interact with, preferring larger-scale, better-endowed, and more innovative farmers who can provide some in-kind payment and are likely to exhibit better performance (Feder and Slade 1993). This sort of supply-side rationing is exacerbated by self-selection by farmers. Those who attach a higher value to (larger demand for) information tend to be large-scale farmers with better opportunities for taking advantage of information.

This selectivity of contacts has ramifications for the diffusion of information through farmer-to-farmer communications. Because the farmers who tend to have more extension contacts are often not typical of the farming population, other farmers are less inclined to follow the example of contact farmers or to seek their advice (despite some positive experiences, such as in Israel: Keynan, Olin, and Dinar 1997). On the supply side, the reaction to the large clientele is to deploy large numbers of extension agents, which presents a management challenge for national organizations or organizations managed by large geographical administrative units (states or provinces). When there are large numbers of field personnel, there is a tendency to adopt a centralized, hierarchical, top-down management system. Such bureaucracies are not generally receptive to participatory approaches to information delivery and priority setting (Fleischer, Waibel, and Walter-Echols 2002), and by distancing decisionmaking from the field level, they often lead to suboptimal decisions.

**Dependence on the broader policy environment.** The effectiveness of extension work is crucially dependent on complementary policies and institutional actions over
which extension management has limited influence. These involve credit, input supplies, price incentives, marketing channels, and human resource constraints, among others. Although extension agents can adjust their advice to the overall policy and institutional climate, the value of the information is diminished when the terms of trade are tilted against agriculture, rural infrastructure investment is inadequate, and input supplies are irregular due to imperfections in input markets. Coordination between the agencies that influence these complementary factors and extension management is costly and difficult, and extension agencies generally have little leverage. Particularly detrimental are the weak linkages to the knowledge generation system, especially the national agricultural research system.

**Interaction with knowledge generation.** In contrast to the situation in the United States, where the cooperative extension service is embedded in the university system, in most developing economies the information on which extension advice is based is not generated within the extension organization itself but in separate systems (national agricultural research institutes and universities and, increasingly, private research firms). Under separate management structures and incentive structures, research systems give little weight to the extension service’s opinions and priorities. Because the performance of research systems is often assessed according to the recognition it receives within the scientific community, research priorities are not necessarily aligned with those of extension managers or the farmers they come in contact with.

Public research and extension organizations often compete for budgets. Researchers typically enjoy a higher status (they are often better educated and have greater independence), and this produces tension in interactions with extension services that is not conducive to two-way feedback or to effective extension services (see Mureithi and Anderson forthcoming on the situation in Kenya). A World Bank review (Purcell and Anderson 1997) of a large portfolio of extension projects found that research-extension links were generally weak and that neither research nor extension was sufficiently conscious of the need to understand the constraints and potentials of different farming systems as a basis for determining relevant technology and technology-development requirements. These inadequate research-extension links led to undesirable outcomes in a large proportion of the projects reviewed. More recent World Bank extension operations, building on the lessons of experience (including the importance of support for business development services for small and medium-size enterprises), have had more positive outcomes.

**Difficulty in attributing impact.** Because many factors affect the performance of agriculture in complex ways, it is difficult to attribute specific impacts at the farm level to extension services. This difficulty weakens political support and exacerbates
problems of budget allocation and staff incentives and accountability, both upward (to managers) and downward (to clients).

Evaluating the impact of extension involves measuring the relations between extension and farmers’ knowledge, adoption of better practices, and use of inputs: farm productivity and profitability; and related improvements in farmers’ welfare. But farmers’ decisions and performance are influenced by many other systematic and random effects (prices, credit constraints, weather, other sources of information), so distinguishing the impact of extension advice requires careful use of econometric and quasi-experimental methods.

The inability to attribute impact unambiguously undermines the incentives of extension staff to reach out to farmers or even update their skills and knowledge. Instead of assessing outcomes and performance, time is spent collecting and reporting on input indicators, which are easier to obtain and measure.

**Weak accountability.** Because the effectiveness of extension activities cannot be easily established and performance is measured in terms of input indicators, field staff are generally not held accountable for the quality of their extension work and are often able to shirk on quantity as well. The same impact attribution problems mean that higher-level managers, though nominally accountable for extension performance to the political level, are monitored mainly in terms of budget spent, staffing levels, and other bureaucratic indicators. Accountability to clients is only nominal, as typically there are neither mechanisms nor incentives to make extension services accountable to farmers—who are the only ones who can easily observe the quality and effectiveness of extension services. Little attention is given to systematic participation by the farming community in problem definition, problem solving, and extension programming. Without mechanisms for accountability to farmers, incentives are distorted. Extension agents divert time and energy to other activities, which earn them extra remuneration, such as promoting inputs for which they earn a commission, or helping farmers access credit.

**Weak political commitment and support.** Even in countries where agriculture is still a large economic sector, public policies and investments have traditionally favored urban areas (Binswanger and Deininger 1997). Within agriculture, extension tends to be a weak claimant on agricultural budgets. In nearly half the extension projects examined in a mid-1990s World Bank study, lack of commitment and support by senior government officials adversely affected implementation and funding (Purcell and Anderson 1997). Feder, Willett, and Zijp (2001) posit that a plausible reason for the lack of adequate support (and the resulting limited funding) by politicians and senior officials to extension investments is the absence of the kind of political payoffs that can be earned from other public outlays that have visible impacts, such as the
double cropping that follows from an irrigation investment or the reduction in transport cost following construction of a bridge or road.

**Public duties other than knowledge transfer.** Because extension services typically employ large numbers of public servants at the rural community level, governments are often inclined to assign other duties to extension staff, such as collecting statistics, administering loan paperwork and input distribution (for government-provided inputs), implementing special programs (such as erosion control), and performing regulatory duties (Feder and Slade 1993; Purcell and Anderson 1997). Many of these duties are easier than extension services for supervisors to monitor, as there are clear and quantifiable performance criteria (number of loan applications returned, number of statistical reports submitted). As mentioned, there may also be monetary incentives for performing some of these other activities (such as input distribution) that have a clear cash value to farmers. This misallocation of extension agents' time at the expense of information dissemination can go undetected because the outcomes of core extension duties are so difficult to attribute and because accountability to farmers is weak or absent.

**Fiscal sustainability.** An outcome of many of the shortcomings of public extension systems is persistent funding difficulties. The public goods nature of many extension services makes cost recovery at the individual beneficiary level difficult, whereas the dependence on public funding is problematic because of weak political commitment. When budgets shrink, fixed staff costs claim a large share of available funds, and field operations are curtailed, along with other recurrent costs (such as vehicle purchase and maintenance). Scaling down field operations reduces not only the quantity of extension inputs but also their quality, as feedback from farmers is reduced and with it timely follow-up on farmers' issues.

Fiscal inadequacy and the unsustainability of extension operations are common themes in the extension literature (see Feder, Willett, and Zijp 2001; Hanson and Just 2001). More than 70 percent of extension projects in a sample of World Bank–supported operations faced "unlikely" or "uncertain" sustainability (Purcell and Anderson 1997). More recently, this shortcoming has received critical attention in the wider agricultural development literature (Kydd and others 2001).

**Extension Modalities as Induced Institutional Innovations**

This section applies the framework developed in the previous section to analyze several extension modalities that have emerged in the past three decades. These newer approaches, which depart from the traditional public service model, reflect attempts to overcome some of the weaknesses inherent in the public extension systems of

Training and Visit Extension

The training and visit model of extension organization was promoted by the World Bank during 1975–95 in more than 70 countries (Umali and Schwartz 1994). The system stressed a single line of command, with several levels of field and supervisory staff; in-house subject matter specialists to provide training to staff and tackle technical issues reported by field staff; exclusive dedication to technical information dissemination; a strict and predetermined schedule of village visits over a two-week cycle, with contacts with selected “contact farmers”; mandatory biweekly training emphasizing the key set of messages for the forthcoming two-week cycle; a seasonal workshop with research personnel; and better remuneration and transport for extension staff. Although the training and visit design attempted to tackle some of the weaknesses of the public extension service, it also exacerbated other weaknesses. In the end, most of these new structures collapsed.

The problems of scale and complexity were tackled by a heavy reliance on formally selected contact farmers within an identifiable farming group. By working with a small number of contact farmers, who were expected to pass on what they learned to the rest of the farming group, agents were to maximize coverage. But the required staff-farmer ratios implied a significantly larger extension staff, and thus the training and visit extension systems cost some 25–40 percent more than the systems they replaced (Feder and Slade 1993; Antholt 1994). The design intended to deal with accountability by improving management’s ability to monitor staff activities, taking advantage of the strict visit schedule, identifiable contact farmers, intensive hierarchy of supervisory staff, and other quantifiable measures. The monitorable daily schedule also eliminated most activities other than information dissemination. The interaction with research was improved through seasonal meetings, but little influence was gained over research priorities.

Several features of the design could not stand up to practical realities, however. The quality of extension services remained mostly unmonitorable, and the lack of accountability to farmers was not resolved. Biases in the selection of contact farmers led to diminished diffusion as contact farmers were often replaced by “contact groups.” The strict biweekly visit schedule could not be maintained because agents often lacked new messages to convey and farmers had limited interest in frequent visits. The training and visit system appeared to have little impact over time. Although a 1986 study by Feder and Slade (1993) found a positive impact on yields in Haryana, India, three years after project initiation, studies in Pakistan (Hussain, Byerlee, and Heisey 1994) and Kenya (Gautam 2000) found no significant impact after a longer period.
Many observers, including early skeptics such as Moore (1984), agree that what eventually brought about the dismantling of the training and visit extension system was lack of financial sustainability, a general problem of large public extension systems made worse by the higher cost of the training and visit structure. As the ability to demonstrate impact was not improved, there was no significant change in the political commitment to support extension. In country after country, once the World Bank ceased funding (assuming that the new system had been "mainstreamed"), funding returned to the lower levels of the past, which could not sustain the training and visit system. Hard-pressed governments have struggled with downsizing options, in some cases supported by bilateral donors and inevitably coupled with other extension reforms (Sulaiman and Hall 2002).

**Decentralization**

Decentralization retains the public delivery and public funding characteristics of traditional centralized extension but transfers responsibility for delivery to local governments (district, county). This approach was tried by several Latin American governments in the 1980s and 1990s (Wilson 1991) and by Uganda (Crowder and Anderson 2002) and other African countries later. Decentralization is intended to improve accountability by moving services closer to the people who use them. Local governments (if democratically elected) are eager to receive positive feedback on services from the clientele-electorate. This was expected to improve extension agents' incentives and induce better service. The costs of coordination with the activities of other agencies are also generally lower for local agencies operating in smaller geographical areas. Political commitment may be stronger as well because the clientele is closer to the political leadership.

But decentralized extension agencies also face a multitude of additional problems. There is greater potential for political interference and the use of extension staff for other activities (such as election campaigns). Economies of scale in updating staff skills can be lost, and extension-research links are more difficult to organize. Analysis of Colombia's experience with the decentralization of extension confirms these concerns and documents a significant increase in the number of staff and thus in costs (Garfield, Guadagni, and Moreau 1996). Problems of financial sustainability, rather than being resolved, may merely have been transferred to the local level.

A related reform was the devolution of extension functions to farmers associations rather than to local governments, a strategy pursued in several West African countries with some notable successes (Guinea). This approach is likely to have a greater impact on accountability, because the employer is even closer to the clientele. There is also greater potential for financial sustainability, because the farmers' association that provides the public good is better able to recover costs from its members (through general membership fees, for example), although government funding is
generally also provided to the associations. Extension agents may be permanent employees of the associations or contract employees of private entities, nongovernmental organizations, or universities. Conceptually, their incentives for better service are fairly similar regardless of their standing. Remaining problems include difficulties maintaining agent quality due to loss of economies of scale in training and more difficult linkages with research.

Fee for Service and Privatized Extension

Fee for service extension programs in developing economies can reduce the fiscal burden of public extension services, though they usually entail considerable public funding even when the provider is private. Government-funded vouchers or other public support is common (Keynan, Olin, and Dinar 1997; Dinar and Keynan 2001). Small groups of farmers typically contract for extension services to address their specific information needs. Because this solves the accountability problem, the quality of service is likely to be higher. Farmers determine the type of information that is important to them, so the impact of extension advice is likely to be high (Lindner 1993). Defining the public good at the small group level and having the whole group share in the cost resolve the free-rider and nonrivalry problems. Tracing extension impact is much less of a problem than in other types of extension service provision, although issues of asymmetric knowledge of the value of information and identifiability of benefits remain and raise design challenges (Hanson and Just 2001).

Another drawback is the loss of economies of scale in agent training, because agents will generally have to update their skills individually.

An important role for public extension policy is to facilitate the development of private provision of extension services and the gradual withdrawal of the public sector (Keynan, Olin, and Dinar 1997; van den Ban 2000; Dinar and Keynan 2001; Holloway and Ehui 2001). The potential for conflict of interest in contracting arrangements may warrant public regulation and monitoring backed by public information for checking on the quality of the information supplied (Mullen, Vernon, and Fishpool 2000; Rivera and Zijp 2002).

A key drawback of fee for service modes of extension is that less commercial farmers—poorer farmers, women farmers, farmers with smaller or less favorable plots—for whom the value of information is lower, may purchase fewer extension services, because the price of the service will tend to be market-determined. This may have undesirable social implications and may also be an inefficient outcome if poor farmers undervalue information because they have less ability to prejudge its value. One way around this problem is stratification of extension systems by types of clients (Sulaiman and Sadamate 2000). Smaller-scale and poorer farmers may be served by public extension or by subsidized contracted extension services (for example, an association of small-scale farmers would receive public funds to hire extension
staff). Commercial farmers, meanwhile, would be expected to pay a higher share of extension costs in a fee for service system (Wilson 1991; Dinar and Keynan 2001). A fully privatized extension system may result in inefficiencies, however, if there are externalities, such as concerns about soil conservation (Hanson and Just 2001).

**Farmer Field Schools**

Farmer field schools were originally introduced to teach irrigated-rice farmers in Asia about integrated pest management. After being implemented in Indonesia and the Philippines, the programs were replicated in other countries and for other crops, usually with significant donor funding. A typical farmer field school educates farmers on agro-ecosystems analysis, including practical aspects of “plant health, water management, weather, weed density, disease surveillance, plus observation and collection of insect pests and beneficials” (Indonesian National IPM Program Secretariat 1991:5). The approach uses participatory training methods to educate field school participants to make farmer pest observers into “confident [integrated pest management] experts, self-teaching experimenters, and effective trainers of farmers and extension workers” (Wiebers 1993:32).

A program consists of 9–12 half-day sessions of hands-on farmer experimentation and informal training to a group of 20–25 farmers during a single crop-growing season. Initially, paid trainers lead this village-level program, delivering diagnostics and practical solutions for overall good crop management practices. Through group interactions, attendees sharpen their decisionmaking abilities and their leadership, communication, and management skills (van de Fliert 1993). Some participating farmers are selected to receive additional training that qualifies them as farmer-trainers, with official backup support, such as training materials.

The farmer field school approach seeks to rectify the problem of accountability. The trainers who conduct the field school are bound by a strict timetable of sessions and a prespecified curriculum, which can be easily verified by supervisors. Continuous interaction with a cohesive group of trainees creates accountability to the group, which is enhanced by the participatory nature of the training methods. Later, when the training is administered by farmer-trainers who are members of the community, accountability to farmers is presumed to be even greater.

A key drawback of the former field school approach is its cost, which is likely to raise problems of financial sustainability. The intense training activities are expensive per farmer trained (Quizon, Feder, and Murgai 2001a,b), so the amount of service actually delivered (the number of farmers trained) on a national level would be small. Cost-effectiveness and financial sustainability could be improved if farmer-trainers were to become the main trainers, perhaps with significant community funding, and if informal farmer-to-farmer communications were used to facilitate knowledge diffusion.
In practice, however, farmer-trainers have been a minor factor in national farmer field school initiatives in Indonesia and the Philippines (Quizon, Feder, and Murgai 2001a). A study in the Philippines found little diffusion of knowledge from trained farmers to other farmers, presumably because the content of the training is difficult to transmit in casual, nonstructured communications (Rola, Jamias, and Quizon 2002). Recent analysis of field farmer schools in Indonesia found no significant impact on yields and pesticide use by trained farmers or members of their communities (Feder, Murgai, and Quizon 2004). This suggests that both the curriculum and the training approach need to be rethought.

The Impact of Extension

Over the past four decades, extension operations have been one of the largest institutional development efforts the world has ever known. Hundreds of thousands of technicians have been trained, and hundreds of millions of farmers have had contact with extension services. As countries struggle with declining public budgets, a key question must wonder how effective these extension investments have been. Many studies have analyzed the impact of extension and reported impressive results, but the data challenges and econometric difficulties in the analyses suggest that many of the results must be interpreted with caution.

In principle, the economic analysis of extension projects is similar to that of any investment appraisal (see Belli and others 2001, for example), but inevitably challenges arise in appropriately valuing and attributing benefits. For projects that deliver agricultural knowledge products to producers, effectiveness in enhancing productivity can be quantified by estimating the economic benefits to producers (or consumers) and computing a rate of return to the investment (Maredia, Byerlee, and Anderson 2001). Rates of return can be estimated econometrically by relating productivity changes to investment in research and extension or by applying the economic surplus method, which builds benefits from the bottom up based on estimated productivity changes at the field level and adoption rates for each technology.

More comprehensive studies may also seek to trace the wider economic benefits of research and extension through factor and product markets. Economic analysts are increasingly being asked to address objectives beyond efficiency, such as equity and poverty alleviation, environmental quality, food safety, and nutrition (see, for example, Alston, Norton, and Pardey 1995). But there is still no consensus on how far research and extension organizations should depart from their traditional efficiency objective to guide policy analysts concerned with the relevance and effectiveness of investment in research and extension.

In econometric studies, a production function, cost function, or total factor productivity analysis is used to estimate the change in productivity resulting from
an investment in extension. A production function incorporates conventional inputs (land, labor), nonconventional inputs (education, infrastructure), and the stock of technical knowledge (perhaps represented by investment in extension). Recent efforts have expanded the specification to include resource quality variables (soil erosion, nutrient status) and weather variables. The estimated coefficients on extension (measuring marginal product) are then used to calculate the value of additional output attributable to the respective expenditures (holding other inputs constant) and to derive marginal rates of return to the investments.

There are many technical areas of debate in the literature on econometric methods, such as the length and shape of time-lag structures, the appropriate method of determining the rate of return from the estimations, and the quality of indices used as the dependent variable (Alston, Norton, and Pardey 1995). However, the main constraints on the wider application of econometric approaches in developing economies are data availability and quality. The econometric approach requires good-quality time-series data, which are difficult to obtain below the national or state level in most developing economies. Therefore, this approach is generally best for ex post evaluations of entire agricultural research and extension systems over a long period (say, 25–30 years), if the quantity and quality of data allow the use of statistical methods. Much of the work in this area in developing economies was pioneered by Robert Evenson (see, for example, Evenson and Pray 1991).

One good approach is to use panel data to capture both cross-sectional and time-series variability (Gautam 2000). Secondary panel data are increasingly available for many variables at the district level, especially production and input data, and some recent studies have even included districtwide data on resource quality (for a review of such studies, see Maredia, Byerlee, and Anderson 2001). Most of the studies have focused on the impact of research rather than extension. Indeed, in studies based on time-series data, it is often difficult to separate the effects of research from those of extension. As panel data become more widely available, the use of econometric analysis of extension impact will expand.

Birkhaeuser, Evenson, and Feder (1991) provided an early review of studies of extension impact and found few studies that systematically compared costs and benefits with and without an extension project. Although early evaluations of extension investments criticized the observed low levels of efficiency and frequent lack of equity in service provision, they reported relatively high benefit-cost ratios (see Perraton and others 1983). More recent studies have also found significant and positive effects (Bindlish, Evenson, and Gbetibouo 1993; Bindlish and Evenson 1993), with internal rates of return on extension investments in developing economies ranging from 5 percent to more than 50 percent (Bindlish and Evenson 1997; Evenson 1997).

The overriding lesson of Evenson’s (1997) review of 57 studies of the economic impact of agricultural extension is, however, that impacts vary widely—many programs
have been highly effective, whereas others have not. A recent meta-analysis of 289 studies of economic returns to agricultural research and extension found median rates of return of 58 percent for extension investments, 49 percent for research, and 36 percent for combined investments in research and extension (Alston and others 2000). Similar economic performance has been documented for Sub-Saharan Africa alone (Oehmke, Anandajayasekeram, and Masters 1997).

However, although economic analysis seems to provide fairly strong justification for many past extension investments, it does not tell the full story. Concern about data quality and difficult methodological issues of causality and quantification of benefits must be important qualifiers to the prevailing evidence of good economic returns from extension. In Kenya, perhaps the most closely studied case in developing areas (from Leonard 1977 to Gautam 2000), early evaluations had indicated remarkably high positive economic returns to extension investments, but a comprehensive evaluation based on new and improved data found disappointing performance—an ineffective, inefficient, and unsustainable training and visit system and no measurable impact on farmer efficiency or crop productivity (Gautam 2000). Such findings do little to dispel the skepticism of policymakers (reinforced by observations such as those of Hassan, Karanja, and Mulamula 1998) about the returns to investment in public extension. More evaluative work is clearly called for to assist policymaking and investment decisions.

Conclusion

Agricultural extension can play an important role in development. The public goods character of much extension work underpins the extensive public investment in extension services. But although public extension organizations are common in developing economies, they are often inadequately funded and their effectiveness is limited by many administrative and design deficiencies and challenges. Chief among these are the large scale and complexity of extension operations, the important influence of the broader policy environment, weak links between extension and knowledge generation institutions, difficulties tracing extension impact, problems of accountability, weak political commitment and support, the frequent encumbrance of extension agents with public duties beyond those related to knowledge transfer, and severe difficulties of fiscal unsustainability.

Among these general problems of extension organization, the difficulty of attributing impact weakens political support, leading to small budgets and problems of fiscal sustainability. Ironically, this same difficulty may explain why international development agencies continue to heavily support extension activities, financing some $10 billion in public extension projects over the past five decades. The economic justification for the project is rarely based on solid ex ante cost-benefit analysis, because
parameters are typically not available from past projects because of the difficulties of attributing impact. Attribution problems also imply that it will be difficult to establish failure once a project is completed (completion is the artificial point in time when donor funding is fully disbursed but farming and extension activities continue).

Several other factors also account for the popularity of extension projects among donors. Extension projects are relatively easy to design, typically involving a small number of recipient government agencies, often just the ministry of agriculture. This reduces bureaucratic complexity. The activities funded by the project are well-defined inputs: constructing and refurbishing extension offices, training agents, providing transport and budgets for field operations, and funding additional personnel. If the project is national in scale, it is easy for donors to build its budget to a significant size—a positive attribute for a development agency striving to maintain its own cost-effectiveness per dollar granted or loaned.

There is thus some tension between domestic decisionmakers, who are reluctant to invest heavily in extension, and development agencies, which enthusiastically promote investment in extension. The availability of external funds minimizes the need for tradeoffs between investments in extension and investments in more politically rewarding undertakings, such as irrigation systems. But it also simply postpones the day of reckoning. Once the externally funded project is over, the lack of political support resurfaces and extension budgets are again cut. The more expensive features of the foreign-funded effort are abandoned, and the size of the extension service is cut way back (Purcell and Anderson 1997).

Several lessons for future extension systems emerged from this review, including some reflections on the pros and cons of different models of extension delivery that were developed in the past few decades. Each situation calls for suitable extension provision methods, but this review emphasizes the efficiency gains that can come from locally decentralized delivery and incentive structures based on largely private provision. Most extension services will remain largely publicly funded, however, especially in impoverished developing economies.

Much remains to be done to bring appropriate extension services to poor farmers around the world. But investors need to be cautious in designing public extension systems and to draw lessons from experience. Informed by these lessons, governments should be able to increase the returns to their investment and successfully assist farmers in boosting their productivity and income, thereby contributing to stronger economic growth.

Notes

Jock R. Anderson was an adviser and is presently a consultant in the Agriculture and Rural Development Department at the World Bank; his e-mail address is janderson@worldbank.org. Gershon

Feder is research manager in the Development Research Group at the World Bank; his e-mail address is gfeder@worldbank.org. The authors have drawn on the considerable World Bank experience with extension, including the work of many of their colleagues, notably Gary Alex, Derek Byerlee, Ariel Dinar, David Nielson, Dina Umali-Deininger, and Willem Zijp. Seniority of authorship is not assigned.

1. The sample of studies was strongly oriented toward research. Only 18 of 1128 estimates of rates of return were for "extension only," whereas 598 were for "research only" and 512 for "research and extension combined."

References


Targeting Outcomes Redux

David Coady • Margaret Grosh • John Hoddinott

A newly constructed comprehensive database of 122 targeted antipoverty interventions in 48 countries is used to examine the contested issue of the efficacy of targeting interventions in developing countries. Though the median program transfers 25 percent more to poor individuals (those in the bottom two quintiles) than would universal allocation, a quarter of the interventions are regressive. Targeting is better in richer countries, in countries where governments are more likely to be held accountable, and in countries where inequality is higher. Interventions that use means testing, geographic targeting, and self-selection based on a work requirement are all associated with an increased share of benefits going to poor people. Proxy-means testing, community-based selection, and demographic targeting to children show good results on average but with wide variation. Self-selection based on consumption, demographic targeting to the elderly, and community bidding show limited potential for good targeting. The substantial variation in targeting performance within specific program types and targeting methods suggests that differences in implementation are also important factors in determining the success of targeting to poor individuals.

In the past two decades a consensus has emerged that social safety nets designed to raise and protect the consumption levels of poor households can play a crucial role in development (World Bank 1990, 1997, 2000). Implementing this agenda for reducing poverty requires methods for reaching poor people. This can be accomplished by broad targeting through programs that reach a wide swath of society, including poor people (for example, universal primary education, an extensive network of basic health care), or by narrow targeting through programs that identify the poor more specifically to confer benefits disproportionately on this group (for example, income transfer programs). The case for narrow targeting rests on the existence of a budget constraint. The overall poverty impact of a program depends on the number of poor households covered and the level of benefits they receive. With a fixed poverty alleviation budget, the opportunity cost of transfers “leaking” to nonpoor households is a lower impact on poverty reduction, reflecting less coverage...
of poor households or lower benefit levels. Targeting transfers to poor households means that more benefits can go to them.

Divergent views on the efficacy of narrowly targeted interventions are based on differing assessments of three questions: Are such methods likely to achieve better targeting outcomes? Are they cost-effective? Do they raise the living standards of poor people? This article addresses the first question. The fairly extensive literature on this topic is dominated by descriptions of individual, sometimes idiosyncratic programs. Even comparative analyses tend to cover a single region (Grosh 1994 for Latin America and the Caribbean; Braithwaite, Grootaert, and Milanovic 2000 for Eastern Europe and Central Asia), method (Bigman and Fofack 2000 on geographic targeting), or intervention (Rawlings, Sherburne-Benz, and van Domelen 2001 on social funds). This partial coverage frustrates efforts to make broader assessments about the effectiveness of different targeting methods or to draw policy-relevant lessons.

To rectify this weakness, this article draws on a newly constructed database of 122 targeted antipoverty interventions in 48 countries in Latin America and the Caribbean, Europe and Central Asia (including the former Soviet Union), the Middle East and North Africa, Sub-Saharan Africa, and South and East Asia. These data are used to address three questions: What are the targeting outcomes? Are there systematic differences in targeting performance by targeting method or other factors? What are the implications of any such systematic differences for the design and implementation of targeted interventions?

The analysis shows that the median targeted program is progressive in that it transfers 25 percent more to poor individuals—persons and households in the bottom two income quintiles—than would be the case with a universal allocation. However, for a staggering quarter of the programs, outcomes are regressive. Countries with higher income, here taken to imply better capacity for program implementation, do better at directing benefits toward poorer members of the population, as do countries where governments are more likely to be held accountable for their behavior, as suggested by better measures of voice. Targeting is also better in countries where inequality is more pronounced. This is consistent with higher welfare gains from targeting when inequality is high as well as lower costs in identifying poor beneficiaries.

Subject to several caveats, interventions that use means testing, geographic targeting, and self-selection based on a work requirement are all associated with an increased share of benefits going to the bottom two quintiles. Proxy-means testing, community-based selection of individuals, and demographic targeting to children show good results on average but with wide variation. Self-selection based on consumption, demographic targeting to the elderly, and community bidding show limited potential for good targeting. That said, examination of experiences with specific program types and specific targeting methods shows considerable variation in targeting performance. This partly reflects heterogeneity across the interventions assessed—some, like means-tested transfers, have poverty reduction as their sole objective.
whereas others, such as social funds, have multiple objectives. However, substantial variation in targeting performance by specific program types and specific targeting methods suggests that differences in implementation are also important factors in whether targeting to poor individuals is successful or not.

Data Construction and Description

The first step in the analysis was a review of the extensive literature on targeted programs and the construction of a database of targeted antipoverty interventions. To our knowledge, this work represents the most extensive attempt to construct such a database.

Database Construction

There were five criteria for inclusion in this database:

- The intervention is in a low- or middle-income country.
- A principal objective of the intervention is poverty reduction, defined in terms of income or consumption.
- The intervention involves a transfer of resources to the beneficiary.
- Documentation on the intervention contains information on the type of targeting method, its implementation, and outcomes.
- The intervention is relatively recent (generally during 1985–2003).

Included in the data are cash transfers (including welfare and social assistance payments, child benefits, and noncontributory pensions), near-cash transfers (such as quantity-rationed subsidized food and food stamps), food transfers, universal food subsidies, nonfood subsidies, public works, and social funds.

Two observations should be made on these criteria for inclusion. First, some interventions have broad objectives that may go beyond direct poverty reduction. Social funds are a good example. Although short-term poverty reduction can be an important component of these interventions, so can be the construction of physical assets valued by the poor and the development of local capacity to design, implement, and maintain infrastructure. The heterogeneity of objectives within broadly defined “antipoverty” interventions demands caution in interpreting comparisons across types of interventions.

Second, focusing the review in this way necessarily means excluding some interventions that may be targeted and may have some impact on poverty. Thus, excluded are occupationally based transfer schemes, such as formal sector unemployment insurance or occupational old age or disability pensions (the principal mechanism determining eligibility and benefit levels are employment and contributions history rather than poverty status); credit and microcredit schemes (although often targeted,
they are largely motivated by credit market failures and do not necessarily involve a transfer of resources to participants; supplementary feeding programs (mainly because the vast literature on this type of intervention did not yield studies that satisfied the criteria described); and most short-term emergency aid (because despite the clear poverty focus and targeting by need, the time scale on which it operates typically precludes an assessment of the distribution of the benefits).

Because most studies of targeting do not appear in peer-reviewed journals, searches were conducted of the gray literature using Web search engines of the World Bank, Eldis, and the International Food Policy Research Institute using the following key words: safety nets, targeting, social funds, pensions, public works, and subsidies. Additional studies were found by canvassing colleagues about work that had not yet been catalogued. Searches were also undertaken in the following academic journals for 1990–2003: *Economic Development and Cultural Change*, *Journal of Development Economics*, *Journal of Development Studies*, *Journal of Public Economics*, *World Bank Economic Review*, *World Bank Research Observer*, *World Development*, and *Economic and Political Weekly*. Existing compilations, such as Grosh (1994) and Braithwaite, Grootaert, and Milanovic (2000), were also reviewed.

The sample of interventions is not necessarily reflective of the distribution of programs around the world but rather of programs that have some measurement of targeting outcomes and that have been written up in the catalogued English language literature that was accessed in this study. A program is more likely to be written up this way if one or more of the following features apply: It is from a country with a household survey that measures consumption and participation in government programs, it is in a country with a culture of evaluation as part of decision-making, it receives funding from an international agency that requires measurement of outcomes, and it appeals to analysts and editors because of its methods or setting. This suggests, for example, that programs using community-based methods and agents are underrepresented. Community-based methods are often locally funded and chosen when data and administrative capacity are weak, features that reduce the likelihood of an evaluation being done and finding its way into the international literature. The literature on public works in Sub-Saharan Africa is probably also underrepresented. By contrast, proxy-means tests are well represented, with a large share of all such programs showing up in the sample.

**Database Description**

Based on the criteria described, information was collected on 122 interventions in 48 countries.

**Intervention type.** The distribution of these interventions shows fairly broad coverage by region and by intervention type (table 1). In some regions a particular intervention type dominates: cash transfers in Europe and Central Asia, universal food subsidies in
### Table 1. Distribution of Intervention Types by Region and Country Income Level for Sample Programs in 1985–2003

<table>
<thead>
<tr>
<th>Region and income level&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Transfers</th>
<th>Subsidies</th>
<th>Public Works</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cash</td>
<td>Near Cash&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Food</td>
</tr>
<tr>
<td>Region</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latin America and Caribbean (32)</td>
<td>14</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Europe and Central Asia (26)</td>
<td>24</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Middle East and North Africa (14)</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sub-Saharan Africa (13)</td>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>South Asia (21)</td>
<td>1</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>East Asia and Pacific (16)</td>
<td>5</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Income level&lt;sup&gt;d&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poorest (63)</td>
<td>17</td>
<td>15</td>
<td>6</td>
</tr>
<tr>
<td>Less poor (59)</td>
<td>32</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Total (122)</td>
<td>49</td>
<td>18</td>
<td>12</td>
</tr>
</tbody>
</table>

<sup>a</sup>Numbers in parentheses are total number of interventions.  
<sup>b</sup>Includes interventions such as food stamps or the right to purchase a limited quantity of food at a subsidized price.  
<sup>c</sup>For example, social funds.  
<sup>d</sup>Poorest countries have per capita income in 1995 purchasing power parity of $1,200 or less; less poor countries have per capita income of between $1,201 and $10,840.  

*Source:* Authors' compilation based on data search described in the text.

The database distinguishes three broad forms of targeting: individual and household assessments, categorical (geography, age), and self-selection, each with subcategories (table 2).

With *individual and household assessments*, eligibility is directly assessed on an individual basis using one of several methods. In a *verified means test*, information is obtained on a household’s income or wealth and compared with other sources of information, such as pay stubs and income and property tax records. Use of this...
Table 2. Distribution of Targeting Forms and Methods by Region, Country Income Levels, and Program Type for Sample Programs in 1985–2003

<table>
<thead>
<tr>
<th>Region and income level&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Individual assessment</th>
<th>Categorical</th>
<th>Self-selection</th>
<th>Community bidding</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Means tests</td>
<td>Proxy-means tests</td>
<td>Community assessment</td>
<td>Geography</td>
</tr>
<tr>
<td>Region</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latin America and Caribbean (68)</td>
<td>8</td>
<td>5</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>Europe and Central Asia (46)</td>
<td>14</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Middle East &amp; North Africa (23)</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Sub-Saharan Africa (25)</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>South Asia (49)</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>East Asia (42)</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>Income level&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poorest (147)</td>
<td>12</td>
<td>3</td>
<td>10</td>
<td>37</td>
</tr>
<tr>
<td>Less poor (106)</td>
<td>22</td>
<td>5</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>Program type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cash transfer (103)</td>
<td>24</td>
<td>4</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Near-cash transfer (36)</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>Food transfer (35)</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Food subsidy (23)</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Nonfood subsidy (9)</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Public works, job creation (29)</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Public works, program output</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>(e.g., social fund) (18)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total (253)</td>
<td>34</td>
<td>8</td>
<td>14</td>
<td>52</td>
</tr>
</tbody>
</table>

<sup>a</sup>Numbers in parentheses are total number of interventions. Many programs use more than one targeting method, so the total number of targeting methods is greater than the number of programs.

<sup>b</sup>Poorest countries have per capita GDP in 1995 purchasing power parity of $1200 or less; less poor countries have per capita GDP of between $1,201 and $10,840.

Source: Authors' compilation based on data search described in the text.
method requires the existence of verifiable records in the target population and the administrative capacity to process the information and update it regularly.

A simple means test, with no independent verification of income, can be conducted by visiting the household to verify qualitatively that visible standards of living (reflecting income or wealth) are generally consistent with reported figures. Proxy-means tests involve generating a score for applicants based on fairly easy to observe characteristics of the household, such as location and quality of dwelling, ownership of durable goods, demographic structure of the household, and the education of adult members. The indicators used in calculating this score and their weights are derived from statistical analysis of data from detailed household surveys.

An increasingly popular approach to individual assessment has been to decentralize the selection process to local communities. A group of community members or a community leader whose principal functions in the community are not related to the transfer program then decides whom in the community should benefit and whom should not—hence the term community-based targeting.

Categorical targeting, also referred to as statistical targeting, tagging, or group targeting, involves defining eligibility in terms of individual or household characteristics that are considered easy to observe, hard to falsify manipulate, and correlated with poverty. Age, gender, ethnicity, land ownership, and household demographic composition or location are common methods of targeting. Geographic targeting is often used in tandem with other methods.

Some interventions rely on self-selection. These interventions, with nominally universal eligibility, are designed with dimensions intended to encourage the poorest to use the program and the nonpoor not to do so. This is accomplished by recognizing differences in the private participation costs of poor and nonpoor households. Examples include the use of low wages on public works schemes so that only those with a low opportunity cost of time due to low wages or limited hours of employment will present themselves for work; time restrictions on transfers, requiring applicants to queue; and location of points of service delivery in areas where the poor are highly concentrated so that the nonpoor have higher (private and social) costs of access. In social fund-type interventions self-selection occurs at the community level (community bidding), with targeting occurring as a result of differences in private participation costs between poor and nonpoor communities.

Universal food subsidies (with or without quantity rationing) can be viewed as a form of self-selection because they are universally available and households receive benefits by deciding to consume the commodity. Households can often determine not just whether to participate but also the intensity of their participation. The more income elastic are the expenditures on these items, the more effective is the targeting. Food transfers often involve commodities with “inferior” characteristics (for example, low-quality wheat or rice), and households often substitute away from such commodities as their incomes rise.
Targeting methods. Within this broad taxonomy of targeting forms, interventions rely on various targeting methods and often on several methods (see table 2). The 122 interventions used 253 different targeting methods, for an average of just over two targeting methods per intervention. Although 48 interventions used a single targeting method, 42 used two methods, 21 used three methods, and 11 used four or more methods.

Thus, although certain intervention types typically rely on specific targeting methods, most use a combination of methods, presumably because there is a synergy between them that results in greater targeting efficiency. Public works programs typically use a combination of geographic targeting and self-selection based on low wages and a work requirement. If demand exceeds supply at the wage paid, public works also often require additional rationing of employment using categorical targeting. Similarly, social funds are partly demand driven and therefore have an element of community self-selection. Food subsidies are self-targeted based on consumption patterns. Cash transfers are most likely to have some form of individual assessment but are also often conditioned on other characteristics (such as age in the case of pensions or child benefit).

There are some marked differences in targeting methods by region. Most interventions using means and proxy-means testing are concentrated in Europe and Central Asia and Latin America and the Caribbean. A legacy of the central planning era in Europe and Central Asia is an extensive administrative system suited to individual assessment using some form of means or proxy-means testing. This, together with a distribution of income that at least at the time of transition was relatively equal, has meant that targeting is based either on individual assessment or some individual characteristic, such as age. A reliance on food subsidies explains why self-targeting based on consumption patterns is the dominant targeting method in the Middle East and North Africa. South Asia is notable for extensive use of geographic targeting and for relatively high reliance on self-selection based on work or consumption. Latin American and Caribbean countries also use geographic targeting extensively, along with means or proxy-means testing or targeting to children. Interventions in Sub-Saharan Africa and East Asia show more mixed patterns. There are also broad differences across income levels. Generally, poorer countries rely more on self-selection and categorical targeting, whereas individual assessments are relatively more common in less poor countries. The one exception to these general patterns is categorical targeting by age, which is used relatively less frequently in poor countries.

Assessing Targeting Effectiveness

This section outlines the methodology used to compare targeting performance across interventions by creating an indicator of targeting performance. It also identifies some important caveats to be kept in mind when interpreting this indicator and briefly describes targeting outcomes in terms of this indicator of targeting performance.
Measures of Targeting Effectiveness

There is a growing body of literature on measuring targeting effectiveness (Ravallion and Chao 1989; Ravallion 1993; Cornia and Stewart 1995; Coady and Skoufias 2001). Comparing the performance of the targeting methods used in the range of programs considered in this meta-analysis requires a comparable performance indicator for each program. The definitions, methods, and presentations in the original studies vary in ways that make it difficult to assemble a single summary performance indicator. Incidence and participation rates may be reported over the full welfare distribution; for the poorest 10, 20, or 40 percent of the population; or for a poor and nonpoor classification that differs by country. Other studies report none of these measures but use other, less common ones. Of course, the measures of welfare used are not always strictly comparable. The problem is how best to compare targeting performance outcomes using data that are not strictly comparable.

Most studies catalogued in the database provide information on at least one of the following indices:

- The proportion of total transfers received by individuals or households falling within the bottom 40, 20, or 10 percent of the national income distribution.
- The proportion of beneficiaries falling within the bottom 40, 20, or 10 percent of the national income distribution.
- The proportion of total transfers going to poor households or individuals, defined in terms of some specified part of the welfare distribution (for example, falling in the bottom 35 percent of the income distribution).

Ideally, information would be available on the proportion of total transfers received by households falling within different deciles (40th, 20th, 10th, and so on) of the national income distribution. This is a better measure than the proportion of beneficiaries by decile, which says nothing about variations in the levels of transfers. These two measures—proportion of total transfers and proportion of beneficiaries—are only equivalent when transfer levels are uniform across beneficiaries.

Because the studies did not use a single common measure of targeting performance, a uniform measure of targeting effectiveness was constructed based on a comparison of actual performance and a common reference outcome: the outcome that would result from neutral (as opposed to progressive or regressive) targeting. A neutral targeting outcome means that each decile receives 10 percent of the transfer budget or that each decile accounts for 10 percent of the program beneficiaries. Neutral targeting can be thought to arise either from the random allocation of benefits across the population or from a universal intervention in which all individuals receive identical benefits.

The indicator is constructed by dividing the actual outcome by the appropriate neutral outcome. For example, if people or households in the bottom 40 percent of...
the income distribution receive 60 percent of the benefits, the indicator of performance is calculated as $60/40 = 1.5$, meaning that targeting has resulted in the target group (here, people in the bottom two quintiles) receiving 50 percent more than they would have received under a universal intervention. A value greater than one indicates progressive targeting, and a value less than one regressive targeting (unity denotes neutral targeting).

The performance indicator used in the analysis that follows is based on a lexicographic selection process among the available incidence indicators, as follows. Preference is given to a measure of the share of benefits accruing to a target group rather than to a measure based on a portion of beneficiaries in the target group. Within that, preference is giving to the proportion accruing to the bottom two quintiles. If that information is not available, then to the bottom quintile, and if that is not available, then to the poorest decile. Such a performance indicator could be calculated for 85 of the 122 programs (table 3).

**Descriptive Results**

There is enormous variation in targeting performance, with scores ranging from 4 for the Trabajar public works program in Argentina to 0.28 for value-added tax exemptions on fresh milk in South Africa. The median value is 1.25, so that the “typical” program transfers 25 percent more to households in the bottom quintiles or below the poverty line than would be the case with a universal allocation. However, a staggering 21 of the 85 programs—25 percent—are regressive, with a performance index of less than 1, suggesting that in these cases a random selection of beneficiaries would provide greater benefits to the poor.

It is instructive to focus on the 10 worst and 10 best programs. The 10 worst have scores ranging from 0.28 to 0.78 with a median score of 0.60. Five are food subsidy programs and three involve cash transfers. Median performance rises from 1.25 to 1.3 if interventions using self-selection based on consumption are withdrawn from the sample, and the proportion of regressive interventions drops from 25 percent to 16 percent. Only one of the poorly performing programs uses either means or proxy-means targeting methods, none use geographic targeting, and the countries come from across the income spectrum. Scores for the top 10 programs range from 2.02 to 4.00, with a median score of 2.15, and all but one are in either Latin America and the Caribbean or Europe and Central Asia. Of these 10, 9 involve cash transfers; 9 use means, proxy-means, or geographic targeting; and 7 are in less poor countries.

Cash transfers show up in both the 10 best and 10 worst programs, highlighting the possibility that variations in targeting performance may reflect poor implementation rather than poor program potential. However, public works programs are all in the top half of the performance table, and social funds are nearly all in the bottom half. This is consistent with there being a tradeoff between the objective of reducing
<table>
<thead>
<tr>
<th>Country</th>
<th>Program</th>
<th>Individual assessment</th>
<th>Categorical assessment</th>
<th>Self-selection</th>
<th>Share of Transfers going to</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Means test</td>
<td>Proxy-means test</td>
<td>Community test</td>
<td>Geographic Age.</td>
</tr>
<tr>
<td>Argentina</td>
<td>Trabajar public works</td>
<td>4.90</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Estonia</td>
<td>Cash social assistance</td>
<td>3.47</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Dominica</td>
<td>Cash transfers</td>
<td>1.00</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Hungary</td>
<td>Cash social assistance</td>
<td>2.72</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Albania</td>
<td>Nghime Ekonomika cash</td>
<td>2.65</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Yemen</td>
<td>Social welfare fund cash</td>
<td>2.15</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Poland</td>
<td>Social assistance cash</td>
<td>2.10</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Romania</td>
<td>Minimum income guarantee</td>
<td>2.08</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Chile</td>
<td>ylieF cash transfers</td>
<td>2.02</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Nicaragua</td>
<td>RPS conditional cash transfer</td>
<td>1.99</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Honduras</td>
<td>PRAF cash transfer</td>
<td>1.98</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Chile</td>
<td>School lunches</td>
<td>1.95</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Slovenia</td>
<td>Otroski Doadatek child benefit</td>
<td>1.95</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>Cash social assistance</td>
<td>1.95</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Belgium</td>
<td>ESF public works</td>
<td>1.93</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Kyrgyz Republic</td>
<td>Unified monthly cash benefit</td>
<td>1.85</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Chile</td>
<td>PASIS cash to poor elderly</td>
<td>1.83</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Republic of Korea</td>
<td>LPS cash for home care</td>
<td>1.81</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Peru</td>
<td>&quot;Glass of milk&quot; program</td>
<td>1.80</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Chile</td>
<td>PEAU/POJH public works</td>
<td>1.78</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>Noncontributory pensions</td>
<td>1.73</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Indonesia</td>
<td>JPS-BK health subsidy</td>
<td>1.68</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>India</td>
<td>Rural AP State PDS rice</td>
<td>1.63</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>School Lunches</td>
<td>1.63</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
|              |                                              |                      |                       |                |                            |         |         |       |       |             |                      |             |             | (Continued)
<table>
<thead>
<tr>
<th>Country</th>
<th>Program</th>
<th>Performance</th>
<th>Means test</th>
<th>Proxy-means test</th>
<th>Community assessment</th>
<th>Geographical Age</th>
<th>Self-selection</th>
<th>Share of Transfers going to</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mexico</td>
<td>LIConSA milk subsidy</td>
<td>1.60</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>64.0</td>
</tr>
<tr>
<td>India</td>
<td>PDS-subsidized jowar</td>
<td>1.58</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>63.0</td>
</tr>
<tr>
<td>Hungary</td>
<td>Child cash allowance</td>
<td>1.57</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>23.5 39.5 62.6</td>
</tr>
<tr>
<td>Mexico</td>
<td>PROGRESA conditional cash</td>
<td>1.56</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>22.0 39.5 62.4</td>
</tr>
<tr>
<td>Colombia</td>
<td>Housing subsidy</td>
<td>1.50</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>35.0 60.0</td>
</tr>
<tr>
<td>Colombia</td>
<td>ICBF day care/school food</td>
<td>1.50</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>28.0 60.0</td>
</tr>
<tr>
<td>Indonesia</td>
<td>JPS-PDR public works</td>
<td>1.48</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>29.6</td>
</tr>
<tr>
<td>Jamaica</td>
<td>Food stamps – MCH</td>
<td>1.45</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>29.0 58.0</td>
</tr>
<tr>
<td>Indonesia</td>
<td>JPS education subsidy</td>
<td>1.44</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>28.8</td>
</tr>
<tr>
<td>India</td>
<td>Rural Mh. State PDS</td>
<td>1.36</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10.4 26.3 54.5</td>
</tr>
<tr>
<td>Zambia</td>
<td>Housing subsidy</td>
<td>1.35</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>28.0 54.0</td>
</tr>
<tr>
<td>Uzbekistan</td>
<td>Child allowance</td>
<td>1.35</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>28.4 54.0</td>
</tr>
<tr>
<td>Latvia</td>
<td>Family child allowance</td>
<td>1.33</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15.0 29.0 53.0</td>
</tr>
<tr>
<td>India</td>
<td>PDS-subsidized rice</td>
<td>1.33</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>53.0</td>
</tr>
<tr>
<td>Indonesia</td>
<td>JPS-OPK rice subsidy</td>
<td>1.32</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>26.4</td>
</tr>
<tr>
<td>Bolivia</td>
<td>Social investment fund</td>
<td>1.30</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11.0 24.5 52.0</td>
</tr>
<tr>
<td>Jamaica</td>
<td>Food stamps program</td>
<td>1.30</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>31.0 52.0</td>
</tr>
<tr>
<td>Honduras</td>
<td>Social fund program</td>
<td>1.25</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>19.0 32.0 50.0</td>
</tr>
<tr>
<td>Chile</td>
<td>Viviendas Basicas mortgage</td>
<td>1.25</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>28.0 50.0</td>
</tr>
<tr>
<td>India</td>
<td>PDS-subsidized oil</td>
<td>1.25</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>50.0</td>
</tr>
<tr>
<td>Sri Lanka</td>
<td>Food stamps</td>
<td>1.25</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>28.0 50.0</td>
</tr>
<tr>
<td>South Africa</td>
<td>Maize VAT exemption</td>
<td>1.23</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>28.0 50.0</td>
</tr>
<tr>
<td>Vietnam</td>
<td>SGF food transfers</td>
<td>1.22</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>24.4</td>
</tr>
<tr>
<td>India</td>
<td>PDS-subsidized kerosene</td>
<td>1.20</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>48.0</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>Food for education</td>
<td>1.20</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>48.0</td>
</tr>
<tr>
<td>Morocco</td>
<td>Food subsidies, flour</td>
<td>1.18</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>23.0 47.0</td>
</tr>
<tr>
<td>India</td>
<td>PDS-subsidized sugar</td>
<td>1.13</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>45.0</td>
</tr>
<tr>
<td>Armenia</td>
<td>Family cash benefit</td>
<td>1.13</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14.0 25.0 45.0</td>
</tr>
<tr>
<td>Peru</td>
<td>Social fund program</td>
<td>1.10</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>23.0 44.0</td>
</tr>
<tr>
<td>Country</td>
<td>Program/Category</td>
<td>Fiscal Year 19</td>
<td>Fiscal Year 2</td>
<td>Fiscal Year 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------------------------</td>
<td>----------------</td>
<td>---------------</td>
<td>---------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bulgaria</td>
<td>Cash pensions</td>
<td>1.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nicaragua</td>
<td>Social fund program</td>
<td>1.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>India</td>
<td>Urban AP State PDS rice</td>
<td>1.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zambia</td>
<td>Social fund program</td>
<td>1.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mozambique</td>
<td>GAPFU cash transfers</td>
<td>1.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>India</td>
<td>Urban Mh. State PDS</td>
<td>1.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tunisia</td>
<td>Universal food subsidies</td>
<td>1.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uzbekistan</td>
<td>Child/low-income benefit</td>
<td>1.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Egypt</td>
<td>Universal flour subsidy</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>India</td>
<td>PDS-subsidized wheat</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latvia</td>
<td>Social assistance program</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Romania</td>
<td>Family child allowance</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Egypt</td>
<td>Universal broad subsidy</td>
<td>0.98</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bulgaria</td>
<td>Child/family allowance</td>
<td>0.95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Egypt</td>
<td>Universal sugar subsidy</td>
<td>0.95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Egypt</td>
<td>Universal oil subsidy</td>
<td>0.95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Armenia</td>
<td>Social fund program</td>
<td>0.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tunisia</td>
<td>Food subsidies</td>
<td>0.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poland</td>
<td>Family child allowance</td>
<td>0.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Romania</td>
<td>Discretionary assistance</td>
<td>0.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morocco</td>
<td>Food subsidies, sugar</td>
<td>0.85</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colombia</td>
<td>SSB utility subsidies</td>
<td>0.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South Africa</td>
<td>Beans VAT exemption</td>
<td>0.79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latvia</td>
<td>Cash pensions</td>
<td>0.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Algeria</td>
<td>Universal food subsidies</td>
<td>0.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South Africa</td>
<td>Oil VAT exemption</td>
<td>0.68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morocco</td>
<td>Food subsidies, oil</td>
<td>0.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Armenia</td>
<td>Humanitarian assistance</td>
<td>0.58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colombia</td>
<td>Subsidio Familiar cash</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yemen</td>
<td>Universal food subsidies</td>
<td>0.45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vietnam</td>
<td>Transfers to old/infirm</td>
<td>0.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South Africa</td>
<td>Milk VAT exemption</td>
<td>0.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Republic of</td>
<td>Old age pension</td>
<td>0.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Author's compilation based on data search described in the text.
current poverty (through public works wage transfers) and the objective of reducing future poverty through developmental public investments (through the assets created by social fund programs). Also, the predominance of less poor countries among the top half of the table suggests that characteristics correlated with income, such as administrative capacity, are important to targeting performance.

Summary statistics on targeting performance—sample size, median, inter-quartile range, and the inter-quartile range as a percentage of the median—by targeting form and method initially suggest a clear hierarchy of targeting performance (table 4). As might be expected, interventions using individual assessments outperform interventions relying on categorical targeting, which in turn outperform interventions relying on self-selection.

Closer inspection, however, reveals that such impressions are too general to be useful. First, there is considerable heterogeneity within these broad methods of targeting. Most notably, self-selection includes interventions that have the highest median performance (those using a work requirement) and those that have the lowest median performance (self-selection based on consumption). Second, self-selection interventions based on consumption and community bidding for interventions have lower median values than other interventions and relatively low variations in these values as measured by the inter-quartile range as a percentage of the median. This

<table>
<thead>
<tr>
<th>Targeting form and method</th>
<th>Sample size</th>
<th>Median targeting performance</th>
<th>Interquartile range</th>
<th>Interquartile range as % of median</th>
</tr>
</thead>
<tbody>
<tr>
<td>All methods</td>
<td>85</td>
<td>1.25</td>
<td>0.68</td>
<td>54.4</td>
</tr>
<tr>
<td>Any individual assessment</td>
<td>37</td>
<td>1.50</td>
<td>0.75</td>
<td>50.0</td>
</tr>
<tr>
<td>Means testing</td>
<td>26</td>
<td>1.55</td>
<td>0.90</td>
<td>58.1</td>
</tr>
<tr>
<td>Proxy-means testing</td>
<td>7</td>
<td>1.50</td>
<td>0.58</td>
<td>38.7</td>
</tr>
<tr>
<td>Community assessment</td>
<td>6</td>
<td>1.40</td>
<td>0.78</td>
<td>55.7</td>
</tr>
<tr>
<td>Any categorical method</td>
<td>58</td>
<td>1.32</td>
<td>0.64</td>
<td>48.5</td>
</tr>
<tr>
<td>Geographic</td>
<td>33</td>
<td>1.33</td>
<td>0.51</td>
<td>36.9</td>
</tr>
<tr>
<td>Age, elderly</td>
<td>12</td>
<td>1.16</td>
<td>0.81</td>
<td>69.8</td>
</tr>
<tr>
<td>Age, young</td>
<td>26</td>
<td>1.53</td>
<td>0.65</td>
<td>42.5</td>
</tr>
<tr>
<td>Other categorical</td>
<td>17</td>
<td>1.35</td>
<td>0.48</td>
<td>35.6</td>
</tr>
<tr>
<td>Any self-selection method</td>
<td>38</td>
<td>1.10</td>
<td>0.41</td>
<td>37.2</td>
</tr>
<tr>
<td>Work</td>
<td>6</td>
<td>1.89</td>
<td>0.30</td>
<td>15.9</td>
</tr>
<tr>
<td>Consumption</td>
<td>25</td>
<td>1.00</td>
<td>0.35</td>
<td>35.0</td>
</tr>
<tr>
<td>Community bidding</td>
<td>7</td>
<td>1.10</td>
<td>0.22</td>
<td>20.0</td>
</tr>
</tbody>
</table>

Source: Authors' compilation based on data search described in the text.
suggests that other things being equal, even the best examples of these targeting methods produce relatively small targeting gains. By contrast, whereas other methods report higher median values, they also tend to have proportionately higher variations in targeting effectiveness. So although these methods offer potentially large gains, there is no guarantee they will improve targeting performance.

One way to explore the source of variation in targeting outcomes is with a Theil inequality index. An attractive feature of the Theil index is that it is subgroup decomposable. When the data are grouped by some characteristic, variations in targeting can be allocated across these programs into two categories: targeting variations due to variations within each group and targeting variations due to variations across groups. When programs are grouped by region, variations in average performance across continents explain only about 28 percent of total variation. When programs are grouped by type, variations in average performance between programs explain 36 percent of the total variation. Grouping by targeting method (geographic, means or proxy-means, both, or other targeting methods) explains only 20 percent of the total variation.

One way of interpreting these large variations is to consider implementation effectiveness. No matter how good the choice among methods or programs, effectiveness of implementation is a key factor determining targeting performance. This point is further strengthened by noting that raising the performance of all programs with the same targeting method and with performance below the method median to the median for that method increases mean targeting performance from 1.38 to 1.55, a return of 17 percentage points. This issue is picked up again in the section on the regression analysis.

Caveats and Limitations

Before reporting on the regression analysis, it is important to make explicit several caveats for interpreting the performance measure and, thus, the analysis based on it.

First, the performance measure is a combination of various measures as already discussed, although the measure used for the vast majority of the interventions (80 percent) is the percentage of benefits accruing to the bottom 40 percent or 20 percent of the national income distribution. This raises concerns about comparability. For example, it may be more difficult to target the poorest 20 percent than the poorest 40 percent, so programs assessed using the 20 percent measure may appear relatively ineffective solely because of the performance indicator used.

This issue was addressed in a number of ways. A second performance measure was calculated that gives, through its lexicographic ordering, priority to the proportion of resources flowing to the bottom 10 percent, then the bottom 20 percent, then the bottom 40 percent. There are no meaningful changes to the results reported in tables 3 and 4 using this performance measure. All regressions were also run...
(reported later) using both measures of targeting performance, and again there were no meaningful changes to the results. This is not completely surprising given that the performance measure and the alternative have correlation coefficients of 0.94 (for levels) and 0.97 (for ranks). As a further check, the multivariate regressions always include variables that control for the performance measure used.

Second, focusing on the benefits accruing to the bottom of the income distribution ignores where in the remaining parts of the distribution the leaked benefits are going. For example, finding that a program is very ineffectively targeted at the bottom 20 percent is less worrying if the leaked benefits accrue mostly to people just above this income cut-off. This is partly why priority is given to the 40 percent measure of performance in constructing the performance index. It is also arguable that a focus of the bottom 40 percent coincides more closely with the objectives of most targeted programs. In any case, the fact that the results are extremely insensitive to the ordering is at least suggestive that where the cut-off point is drawn between 20 percent and 40 percent is somewhat inconsequential.

Third, recall that the data collated are only a sample of the hundreds of antipoverty interventions and that the performance indicator could be calculated only for two-thirds of this sample. These observations point to the possibility of sample selection bias—the possibility that certain characteristics of these programs (such as the fact that they were evaluated and documented) are themselves associated with the measures of targeting performance. A good example of this possibility relates to community targeting. The sample is only a fraction of the studies in Conning and Kevane (2001); it could well be that only successful interventions using community targeting have been well documented.

Fourth, some of the mistargeting observed here arises because households that were poor when the program admission decision was made were better off at the time of assessment or vice versa. This has implications for the design of targeted interventions. When there is considerable movement of households in and out of poverty and no mechanism for updating eligibility, methods that rely on static indicators of living standards (such as proxy-means tests) are likely to perform less well than methods that rely on self-selection.

Finally, the performance index focuses solely on the benefit side of the equation and ignores cost, which may be extremely important in selecting targeting methods or programs. For example, it is often argued that well-designed public works programs can be effective at concentrating benefits in the hands of the poor. But the high nontransfer costs associated with such programs (including nonwage costs and forgone income) substantially reduce their cost effectiveness.

Neglect of the cost side largely reflects data restrictions. Evidence on administrative costs was scant. Some cost data were available for 32 programs, but both cost and performance indicator data were available for only 20. Moreover, the cost data suffer from a severe lack of comparability. Most of the data for Latin America are
taken from Grosh (1994) and give administrative costs as a share of the program budget, based on official records. No attempt is made to determine how much of program benefits are siphoned off through corruption. Much of the cost data on South Asian programs are constructed from survey data on the value of benefits received by sample households. Based on the total program cost and appropriate grossing up, the total cost per dollar of benefit received is calculated. Corruption and theft appear to contribute more to total program expenses than legitimate administrative expenses, though little is said about these expenses. In any case, even when cost data are available, focusing on benefit incidence is extremely important in its own right.

It is worth reemphasizing that the objective of effectively targeting transfers, though always important, is often only one of the objectives of interventions. To the extent that there are tradeoffs between these other objectives and that of effective targeting, this needs to be taken into account in the overall evaluation of any program. It may be that these other objectives impinge more on program design and how the program is “sold” and delivered. Nonetheless, most policy analysts would accept that monitoring the targeting performance of programs dedicated mainly to poverty alleviation is always desirable, especially in developing economies, where poverty is high, budgets are tight, and other policy instruments (such as a comprehensive income tax system) are less developed, less sophisticated, and less progressive.

Regression Analysis

Although factors other than choice of method or program may be relatively large, this does not mean that these choices are unimportant. How important they are is shown by the results of a series of regressions that identify how performance varies systematically across these choices as well as across country characteristics (table 5). Targeting methods are themselves choices; they are not exogenous or predetermined. Consequently, it is incorrect to treat these results as causal relations. Rather, they are measures of partial correlation or association.

The first specification explores how country characteristics such as income, government accountability, and inequality are associated with (log) incidence. Income is measured as log gross domestic product (GDP) per capita (in purchasing power parity U.S. dollars) as of 1995. The hypothesis is that as a country becomes wealthier, it acquires the institutional capacity needed to design a well-targeted intervention.

Government accountability is based on work by Kaufmann, Kraay, and Zoido-Lobaton (1999), who define voice and accountability as a composite measure based on aspects of political processes, civil liberties, and political rights, capturing the
Table 5. Multivariate Analysis of Targeting Performance for Sample Programs in 1985–2003

<table>
<thead>
<tr>
<th>Variable</th>
<th>Basic results</th>
<th>Dependent Variable in Levels</th>
<th>Uses median regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log GDP per capita</td>
<td>0.120</td>
<td>0.181</td>
<td>0.150</td>
</tr>
<tr>
<td></td>
<td>(2.10)</td>
<td>(3.27)</td>
<td>(2.95)</td>
</tr>
<tr>
<td>Log voice</td>
<td>0.161</td>
<td>0.176</td>
<td>0.217</td>
</tr>
<tr>
<td></td>
<td>(2.06)</td>
<td>(2.88)</td>
<td>(2.52)</td>
</tr>
<tr>
<td>Log Gini coefficient</td>
<td>0.363</td>
<td>0.110</td>
<td>0.439</td>
</tr>
<tr>
<td></td>
<td>(1.75)</td>
<td>(0.39)</td>
<td>(1.17)</td>
</tr>
<tr>
<td>Means testing</td>
<td>0.284</td>
<td>0.285</td>
<td>0.300</td>
</tr>
<tr>
<td></td>
<td>(2.43)</td>
<td>(2.59)</td>
<td>(2.90)</td>
</tr>
<tr>
<td>Proxy-means testing</td>
<td>0.252</td>
<td>0.108</td>
<td>0.078</td>
</tr>
<tr>
<td></td>
<td>(1.43)</td>
<td>(0.74)</td>
<td>(0.45)</td>
</tr>
<tr>
<td>Community assessment</td>
<td>0.198</td>
<td>0.119</td>
<td>-0.012</td>
</tr>
<tr>
<td></td>
<td>(0.97)</td>
<td>(0.59)</td>
<td>(0.09)</td>
</tr>
<tr>
<td>Geographic</td>
<td>0.260</td>
<td>0.336</td>
<td>0.341</td>
</tr>
<tr>
<td></td>
<td>(2.72)</td>
<td>(3.25)</td>
<td>(3.33)</td>
</tr>
<tr>
<td>Age, elderly</td>
<td>-0.055</td>
<td>-0.089</td>
<td>-0.117</td>
</tr>
<tr>
<td></td>
<td>(0.37)</td>
<td>(0.68)</td>
<td>(0.091)</td>
</tr>
<tr>
<td>Age, young</td>
<td>0.198</td>
<td>0.128</td>
<td>0.140</td>
</tr>
<tr>
<td></td>
<td>(1.98)</td>
<td>(1.34)</td>
<td>(1.53)</td>
</tr>
<tr>
<td>Other categorical</td>
<td>-0.132</td>
<td>0.028</td>
<td>0.137</td>
</tr>
<tr>
<td></td>
<td>(0.88)</td>
<td>(0.22)</td>
<td>(1.23)</td>
</tr>
<tr>
<td>Work</td>
<td>0.511</td>
<td>0.404</td>
<td>0.404</td>
</tr>
<tr>
<td></td>
<td>(3.66)</td>
<td>(3.74)</td>
<td>(3.73)</td>
</tr>
<tr>
<td>Community bidding</td>
<td>-0.018</td>
<td>-0.012</td>
<td>-0.029</td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.11)</td>
<td>(0.24)</td>
</tr>
<tr>
<td>F-statistic</td>
<td>3.46</td>
<td>3.56</td>
<td>5.41</td>
</tr>
<tr>
<td>R²</td>
<td>0.318</td>
<td>0.406</td>
<td>0.484</td>
</tr>
</tbody>
</table>

Note: The numbers in parentheses are the absolute values of t-statistics. All specifications contain controls, not reported, indicating whether the performance measure is based on proportion of benefits going to the poorest 20 percent, poorest 10 percent, the "poor," or the proportion of poor found in the population. Specifications 1–5 estimate standard errors using the methods proposed by Huber (1967) and White (1981). Specification 6 calculates standard errors using the bootstrap with 50 repetitions (uses median regression). Specifications 1–4 and 6 express the dependent variable in logs; specification 5 uses levels.

Source: Authors' compilation based on data search described in the text.

extent to which citizens participate in the selection of their governments and citizens and the media can hold governments accountable for their actions. They compiled subjective perceptions regarding the quality of governance in different countries, using such sources as polls of experts, commercial risk rating agencies, and cross-country
surveys. The analysis here used countries' rankings, which provide an easier way of interpreting the estimated coefficients. In the sample, Vietnam has the lowest percentile rank for voice (6) and Costa Rica the highest (88).\textsuperscript{11}

Country-specific Gini coefficients are included on the grounds that it might be easier to identify potential beneficiaries when income or consumption differences across individuals are greater. Also included (but not reported) are controls indicating whether the performance measure is based on the proportion of benefits going to the poorest 20 percent, the poorest 10 percent, the "poor" defined with reference to a poverty line, or the proportion of poor in the population. Doing so takes into account confounding effects arising from the use of different measures of poverty incidence in the studies on which this analysis is based. Standard errors are computed using the methods proposed by Huber (1967) and White (1980).\textsuperscript{12}

The results shown in specification 1 show that as country income rises and as inequality rises, so does the targeting performance of antipoverty interventions (see table 5). Targeting is also better in countries where government accountability is better.

Specification 2 looks solely at the impact of choice of targeting method. Dummy variables are included for the nine targeting methods described: three kinds of individual assessment (means testing, proxy-means testing, community selection of individual beneficiaries), four kinds of categorical targeting (geographic, the elderly, the young, others), and two kinds of selection (work requirement, community bidding for projects). The omitted category is self-selection based on consumption. This was chosen as the base category for two reasons. Some argue that this form of targeting is a transition tool to be used only until the capacity for more precise mechanisms—such as means testing—is developed.\textsuperscript{13} Others see self-selection based on the consumption of food as the preferred targeting mechanism and have expressed skepticism about the comparative ability of alternative targeting methods to reach the poor.\textsuperscript{14} Hence, an attractive feature of this specification is that the coefficients on these methods should be interpreted relative to self-selection based on consumption.

Specification 2 shows that means testing, geographic targeting, and self-selection based on a work requirement are all associated with an increased share of program resources going to the poorest 20 percent relative to self-selection based on consumption. Proxy-means testing and targeting the young are also associated with improved incidence, though these are measured with larger standard errors. Targeting the elderly, community assessment, other methods of categorical targeting, and selection based on community bidding are not associated with better incidence relative to the base category of self-targeting based on consumption.

Countries with greater capacity for program implementation may do better at directing benefits toward poorer members of the population either by choosing finer targeting methods or by implementing their choices more effectively. In such cases, the associations in specification 2 may be misleading, merely reflecting correlation
between unobserved implementation capacity and observed targeting methods. This possibility is explored in specifications 3 and 4. Controlling for country income level (specification 3) or income, voice, and inequality (specification 4) does not appear to eliminate the positive association—relative to self-selection based on consumption—between means testing, geographic targeting, and self-selection based on a work requirement and targeting performance. Targeting performance is better in countries with higher levels of income and where governments are held accountable for their actions. Specifically, a 10 percent increase in income is associated with a 1.8 percent increase in targeting performance. Raising the voice rank from 37 (Pakistan’s rank) to 67 (India’s rank) would be associated with about a 30 percent improvement in targeting performance. It is possible, however, that geographic targeting will also be more effective in countries with marked inequalities. Indeed when geographic targeting is dropped from the specification (but all other methods are retained), the parameter estimate for the log Gini coefficient is almost identical to that reported in specification 1.

Several additional specific checks were performed to investigate the robustness of this result. Specification 5 uses the same sample and regressors as specification 4, but the dependent variable is expressed in levels instead of logs. Basic results remain unchanged: means testing, geographic targeting, and targeting based on a work requirement improve targeting performance relative to the omitted category, self-selection based on consumption. There is no meaningful change in any of the other results. Specification 6 estimates median regressions, which express differences in performance in terms of differences in medians. This is an attractive check on robustness because the median is considerably less sensitive to outliers, an especially important consideration when working with small sample sizes. The results are broadly similar to those for specification 4—which uses an identical set of regressors, sample, and dependent variable—with the one exception being a markedly larger coefficient on the log Gini coefficient.

The discussion has focused largely on the association between different targeting methods and targeting performance relative to self-selection based on consumption and conditioning on country characteristics. Also explored is the association between combinations of targeting methods and targeting performance (table 6). In addition to controls for income, voice, governance, inequality, and how the performance measure is constructed, the number of targeting methods used is added to specification 1. The results show that use of more methods is associated with improved targeting, each additional method improving performance by 15 percent. When a series of dummy variables is used in specification 2 to represent the number of targeting methods, the findings are similar. The sample size is too small to explore the association between specific groups of methods and targeting performance, but these results suggest that combining methods improves targeting.
Table 6. Association between Targeting Performance and Number of Methods Used for Sample Programs in 1985–2003

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of methods used</td>
<td>0.165</td>
<td>0.182</td>
</tr>
<tr>
<td></td>
<td>(3.97)</td>
<td>(1.66)</td>
</tr>
<tr>
<td>Used two methods</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Used three methods</td>
<td>0.300</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.88)</td>
<td></td>
</tr>
<tr>
<td>Used four or five methods</td>
<td></td>
<td>0.533</td>
</tr>
<tr>
<td></td>
<td>(3.11)</td>
<td></td>
</tr>
<tr>
<td>Log cap per capita</td>
<td>0.141</td>
<td>0.141</td>
</tr>
<tr>
<td></td>
<td>(3.06)</td>
<td>(2.93)</td>
</tr>
<tr>
<td>Log voice</td>
<td>0.229</td>
<td>0.232</td>
</tr>
<tr>
<td></td>
<td>(3.49)</td>
<td>(3.07)</td>
</tr>
<tr>
<td>Log Gini</td>
<td>0.280</td>
<td>0.278</td>
</tr>
<tr>
<td></td>
<td>(1.35)</td>
<td>(1.31)</td>
</tr>
<tr>
<td>F statistic</td>
<td>6.42</td>
<td>5.08</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.426</td>
<td>0.419</td>
</tr>
<tr>
<td>Sample size</td>
<td>84</td>
<td>84</td>
</tr>
</tbody>
</table>

Note: Numbers in parentheses are absolute values of t-statistics. Specifications 1 and 2 contain controls, not reported, indicating whether the performance measure is based on proportion of benefits going to the poorest 20 percent, poorest 10 percent, the "poor," or the proportion of poor found in population. Specifications 1 and 2 estimate standard errors using the methods proposed by Huber (1967) and White (1980).

Source: Authors' compilation based on data search described in the text.

Conclusion

This article addresses the contested issue of the efficacy of targeting interventions in developing countries using a newly constructed database of 122 targeted anti-poverty interventions in 48 countries. These data are used to address three questions: What are the targeting outcomes? Are there systematic differences in targeting performance by targeting method and other factors? What are the implications of any such systematic differences for the design and implementation of targeted interventions?

The median value of the measure of targeting performance is 1.25, meaning that the median program transfers 25 percent more to poor individuals—persons and households in the bottom income groups—than would be the case with universal allocation. In this sense, targeting works. However, a staggering 21 of the 85 programs for which the performance measure can be built (25 percent) are regressive, with a performance index of less than 1. In these cases, random selection of beneficiaries would provide greater benefits to the poor. Some of this regressivity is driven by the inclusion of food subsidy interventions that use self-selection based on consumption as a targeting method. But even when these are dropped from the sample, 14 percent of targeted antipoverty interventions are still found to be regressive.

Countries with better capacity for program implementation, as measured by GDP per capita, do better at directing benefits toward poorer members of the population.

David Coady, Margaret Grosh, and John Hoddinott
Countries where governments are more likely to be held accountable for their behavior—where "voice" is stronger—also appear to implement interventions with better targeting performance. Targeting is also better in countries where inequality is more pronounced and differences in economic well-being are presumably easier to identify.

Mindful of caveats, interventions that use means testing, geographic targeting, and self-selection based on a work requirement are all associated with an increased share of benefits going to the poorest 40 percent. Proxy-means testing, community-based selection of individuals, and demographic targeting to children show good results on average, but with wide variation. Self-selection based on consumption, demographic targeting to the elderly, and community bidding show limited potential for good targeting.

That said, there is considerable variation in targeting performance when experiences with specific program types and targeting methods are examined. This partly reflects heterogeneity across the interventions assessed because some have multiple objectives whereas others are focused solely on poverty reduction. But when targeting by specific method or program type is considered, there is considerable intra-group variability. Thus, it is not surprising that although community assessment generally performs no better than self-targeting based on consumption, Alderman's (2002) study of community targeting in Albania describes a highly successful example of this form of targeting. Similarly, Case and Deaton (1998) and Duflo (2000) show that in South Africa targeting the elderly is an effective method for reaching poor children, even though targeting the elderly generally performs relatively poorly when compared with other methods for reaching poor people.

Thus, although the patterns observed are instructive, they should not be interpreted as a lexicographic ranking of methods. Differences in individual country characteristics and implementation are also important determinants of outcomes and must be considered carefully in making targeting decisions. For example, in countries where illiteracy is high, it makes little sense to require potential beneficiaries to fill out forms as part of a means test. Such requirements discourage otherwise eligible beneficiaries from applying, as Munro (2003) illustrates for Zimbabwe. This suggests that work on targeting should extend beyond simple quantitative comparisons of methods to consider more detailed and often qualitative issues of comparisons within methods. How does and should implementation differ in different settings, and how can constraints of political economy, poor information, or low administrative capacity best be accommodated or reduced? In a companion article, Coady, Grosh, and Hoddinott (forthcoming) provide a more detailed discussion of the merits, limitations, and options for implementing individual targeting methods in an attempt to move in this direction.
Notes

David Coady is research fellow in the Food Consumption and Nutrition Division at the International Food Policy Research Institute; his e-mail address is d.coady@cgiar.org. Margaret Grosh is lead economist in the Social Protection Anchor Unit at the World Bank; her e-mail address is mgrosh@worldbank.org. John Hoddinott is senior research fellow in the Food Consumption and Nutrition Division at the International Food Policy Research Institute; his e-mail address is j.hoddinott@cgiar.org. The authors thank Akhter Ahmed, Harold Alderman, John Blomquist, Stephen Devereux, Lant Pritchett, Kalanidhi Subbarao, and Salman Zaidi for comments and reference materials; two anonymous referees for helpful comments; Yisgedu Amde and Sanjukta Mukherjee for research assistance; and Manorama Rani for document processing. They also thank participants in seminars at the International Food Policy Research Institute and the World Bank, a conference on Chronic Poverty and Development Policy at the University of Manchester in April 2003, and the North Eastern Universities Development Consortium Conference 2004 at Yale University for useful remarks. The study was commissioned by the World Bank’s Social Protection Anchor Unit for the Safety Nets Primer series. Other papers are available online at www.worldbank.org/safetynets.

1. Van de Walle (1998) contains a useful discussion of these approaches.


3. This focus is conditioned by three factors. First, if targeting is largely ineffective, the answers to the remaining questions are moot. Second, there are simply not enough studies with cost data, and the cost data that are available suffer severely from lack of comparability. Third, assessment of impact requires careful attention to the counterfactual or what beneficiares would have done in the absence of these interventions. Few studies of transfer programs in developing countries do that with any care; exceptions are Datt and Ravallion (1994), Ravallion and Datt (1995), Jalan and Ravallion (1999), and Skoufias (2001).

4. The database is available in the form of an annotated bibliography (Coady, Grosh, and Hoddinott 2003) that includes details on the study itself (title, authors, reference details, year of publication, study objective), background information on the intervention (program name, year implemented, program description, type of benefit, program coverage and budget, transfer levels), targeting method (what criteria were used to determine eligibility, targeting mechanism), how the intervention operated, targeting performance (who benefited), and descriptions of impact on welfare and costs of targeting.

5. Also, the microcredit literature is vast. Attempting to construct a database comparable to that developed here is a major undertaking beyond the scope of this analysis.

6. Strictly speaking, all programs are self-targeted to some degree, because there are always some actions (and therefore costs) required of beneficiaries to register for and collect a benefit.

7. Social funds also use other mechanisms, such as geographical targeting. Differences in access to information or capacity for demanding social funds also account for differential access to these interventions.


9. How these rankings are calculated varies by study. In general, they are based on nationally representative household survey data that estimate the share of transfers that accrue to individuals in the bottom deciles of the consumption distribution. Such estimates do not take into account changes in private transfer behavior or labor supply. Coady, Grosh, and Hoddinott (2003) provide a description of these calculations by intervention as well as references to the original source material.

10. Income and inequality data are from the World Bank’s World Development Indicators database.

11. Preliminary work included a measure of government effectiveness also drawn from Kaufmann, Kraay, and Zoido-Lobaton (1999). However, it is never statistically significant, quite possibly because it is highly correlated with log per capita income.

12. Following the suggestion of an anonymous referee, we explored whether the studies from which data were drawn for this analysis had been published in a refereed journal, a book, or was unpublished work undertaken by World Bank staff, International Food Policy Research Institute (IFPRI)
staff, or individuals based in other institutions. These controls can be thought of in two ways. Arguably, work published in journals (and possibly books) has been more rigorously reviewed, so those studies will be less dogged by measurement error. On the other hand, there may be publishing biases in that only studies with "significant results" are submitted and accepted by journals. Of the 85 estimates of targeting performance used here, 17 come from refereed journals, 17 from books, 37 from unpublished work by World Bank staff, 8 from unpublished work by IFPRI staff, and 6 from unpublished work by researchers based at other institutions. When dummy variables for type of publication are included as additional controls, they are not found to be jointly statistically significant and hence are not reported here. But coefficients on dummy variables for unpublished studies by World Bank staff or by individuals based in other institutions tend to be negative, consistent with the publishing bias hypothesis.


14. Such implicit concern is found, for example, in Cornia and Stewart (1995).

15. Kaufmann, Kraay, and Zoido-Lobaton (1999) caution that these composite measures are likely to be measured with error. As such, they are likely to provide lower-bound estimates of the impact of these characteristics.

16. More precisely, a quartile regression was estimated centered at the median with standard errors obtained via bootstrap resampling with 50 repetitions to correct for heteroscedasticity. Increasing the number of repetitions does not appreciably alter the standard errors.

17. As a further specification check, specification 4 was reestimated, but the sample was restricted in three ways: including only studies reporting the share of benefits accruing to the bottom 40 percent, including only studies reporting the share of benefits accruing to the poorest 40 percent or the poorest 20 percent, and including only studies that report the share of benefits accruing to the poorest 40 percent, the poorest 20 percent, or the poorest 10 percent. The results, available on request, are comparable to those reported in specification 4.

References


An Assessment of Privatization

Sunita Kikeri • John Nellis

Mounting empirical evidence of privatization's benefits coincides with increasing dissatisfaction and opposition among citizens and policymakers. This dissatisfaction reflects the growing questioning of the benefits of privatization, the general downturn of global markets in the past few years and the resulting swing of the pendulum back toward increased governmental supervision, the overselling of privatization as a panacea for all economic problems, and the concern that privatization does not produce macroeconomic and distributional gains equivalent to its microeconomic benefits. This article takes stock of the empirical evidence and shows that in competitive sectors privatization has been a resounding success in improving firm performance. In infrastructure sectors, privatization improves welfare, a broader and crucial objective, when it is accompanied by proper policy and regulatory frameworks. The article argues that despite the growing concerns privatization should be neither abandoned nor reversed. Rather, there should be a strengthening of efforts to privatize correctly; by better tailoring privatization to local conditions, deepening efforts to promote competition and regulatory frameworks, enforcing transparency in sales processes, and introducing mechanisms to ensure that the poor have access to affordable essential services.

Almost every country is divesting some or all of its state enterprises to the private sector or involving the private sector in managing and financing activities previously owned and operated by the state. The reasons for privatization are well established. Especially in developing economies and in infrastructure and network industries, state enterprises have proved wasteful and inefficient, producing low-quality goods and services at high cost. Sheltered from competition, state enterprises were often overstaffed and required to set prices below costs, resulting in financial losses that in acute cases amounted to as much as 5 to 6 percent of gross domestic product (GDP) annually. Bailouts and fiscal strains resulted. Covering state enterprise losses through fiscal transfers required governments to finance larger fiscal deficits and increase tax revenues or reduce public spending in other areas, or both. Financing
losses through the state banking system reduced the private sector’s access to credit and threatened the viability of the financial sector. Many governments became incapable of providing capital to their state enterprises, even profitable ones, for maintenance and repair, much less for badly needed expansion and renewal.

Attempts through the 1970s and early 1980s to impose hard budget constraints, expose state enterprises to competition, and introduce institutional and managerial changes yielded meager improvements in performance. Some effects looked promising (for example, in New Zealand) but proved unsustainable. Backsliding was common, and the poorer the country, the quicker and deeper the reversal. By the middle of the 1980s, following the powerful lead of the Thatcherite revolution in the United Kingdom, government ownership itself came to be seen as a principal reason for the inability to effect major and enduring state enterprise reform. In industrial countries, the shift to private ownership was motivated in equal parts by the failures of reforms short of ownership change, a sea change in ideology, and the short-term fiscal attractions of selling state assets. In developing economies, the impetus was much more financial and fiscal. With the international financial institutions leading the way, privatization was vigorously promoted as a tool to reduce the budgetary burden caused by state enterprise inefficiencies and, in the case of infrastructure, to improve performance and access to investment capital for modernization and expansion of networks.

Despite the widespread adoption of privatization and the positive economic assessments of what privatization could do, a number of critics—sometimes including the general public—have expressed strong reservations about privatization’s fairness and sometimes its efficiency impact. Fueled by some recent problematic and highly visible cases, this skepticism continues, despite the growing number of empirical assessments concluding that in the main privatization improves profitability and efficiency and increases returns to shareholders, particularly for firms in competitive or potentially competitive markets.

What are the strands of the antiprivatization argument? First, opponents contest that privatization has produced financial and operational benefits, or at least enough to offset the social dislocation it causes. Some who acknowledge performance improvements attribute them to increased competition rather than change of ownership, with the implication that less painful instruments could effect needed financial and efficiency gains. Second, there is fear that privatization leads to layoffs and a worsening in labor conditions, in the short term in the divested firms and in the longer run in the economy as a whole. Third, some argue that even if privatization enhances enterprise efficiency, the bulk of the benefits accrue to a privileged few—shareholders, managers, domestic or foreign investors, those connected to the political elite—whereas the costs are borne by the many, particularly taxpayers, consumers, and workers, thus reducing overall welfare. In addition, many are concerned that the often perceived corruption and lack of transparency in privatization transactions have
minimized gains and increased broader problems of governance. Underlying all of these arguments is the fundamental concern that privatization has been applied without proper regard to a country’s economic and social conditions, often at the behest of external actors.

This article takes stock of the empirical evidence on privatization outcomes and explores why privatization provokes such opposition. The first section summarizes privatization trends in the past 15 years. The second section reviews the literature on the impact of privatization at the enterprise and social welfare levels, taking into account the differences between competitive and infrastructure firms, and examines the employment and broader macroeconomic and fiscal impacts. The third section discusses the growing concerns with privatization, focusing on how to design and implement privatization reforms that achieve their underlying economic objectives.

**Trends in Privatization**

Privatization started slowly.¹ Through most of the 1980s, there were only a few divestiture transactions a year. The number of transactions peaked in the mid-1990s and then declined after 1997. Between 1990 and 1999, global proceeds totaled US$850 billion, growing from $30 billion in 1990 to $145 billion in 1999 (figure 1). Organisation for Economic Co-operation and Development (OECD) countries, along with Brazil, account for the overwhelming bulk of the proceeds, mainly from public offerings of large firms in countries of the European Union (Mahboobi 2000).

**Figure 1. Global Privatization Proceeds 1990–99 (US$ billions)**

In non-OECD countries, privatization activity grew rapidly through the mid-1990s, with tens of thousands of enterprises sold and roughly $250 billion in revenues raised during 1990–99. Proceeds peaked at $66 billion in 1997 and then fell following the Asian and Russian financial crises and the ensuing general economic downturn. Revenues during 1990–99 were accounted for largely by infrastructure privatizations, mainly in the telecommunications and power industries, followed by petroleum, mining, agriculture, and forestry. Manufacturing sales accounted for about 16 percent of developing economy privatization proceeds, mainly from sales in Eastern and Central Europe and Latin America (figure 2). By the end of the 1990s, privatization revenues were concentrated in the oil and gas sectors in Argentina, Brazil, India, Poland, and Russia.

By region, Latin America and the Caribbean accounted for the largest share of privatization proceeds (figure 3), with the largest contributions coming from the sale of infrastructure and energy firms in Argentina, Mexico, and Brazil. Eastern Europe and Central Asia sold the largest number of firms, mainly through mass privatization voucher programs before 1995 in Russia, the Czech Republic, Slovakia, Kazakhstan, Lithuania, Ukraine, and Moldova. Sales proceeds were low under the giveaway voucher schemes, but revenues grew after 1995, as countries such as the Czech Republic, Estonia, Hungary, and Russia began or expanded case-by-case sales, including large firms in banking, transport, oil and gas, and infrastructure.

Before the financial crisis of 1997, East Asian countries generally concentrated on opening up their economies to new private entry rather than on privatizing enterprises. This approach was workable, given the region’s smaller reliance on state enterprises as agents of economic policy (except in the People’s Republic of China and a few other Asian socialist states), the success of China’s evolutionary

---

**Figure 2.** Privatization Proceeds by Sector 1990–99 (US$ billions)

![Figure 2](image)

approach to property reform, and the relatively sound financial and fiscal position of most Asian states.

China stands out. Over a 25-year period, Chinese governments created or allowed forms of industrial ownership, particularly at the subnational level, that successfully combined elements of collective and private property. Later, new private entry and foreign direct investment were permitted and encouraged. By 1998, the nonstate sector (including private agriculture) accounted for 62 percent of GDP, whereas state enterprises' share in industrial output declined from 78 percent in 1978 to 28 percent in 1999 (Zhang 2001).

Since 1987, there have been numerous privatizations of small and medium-size collectively owned enterprises in China and further experimentation with ownership forms at the local level. Government has repeatedly announced its intent to clean up and formally privatize even large state enterprises. Many firms carved out their better assets to set up new companies for initial public offerings on the stock market and then became the largest controlling shareholder. There has thus been some dilution of shares over time but—until very recently—mostly to other state entities. Although less concentrated ownership has led to private takeovers in a few cases, full privatization continues to move slowly, mostly because of the social welfare
responsibilities of state enterprises. But the need to face the burgeoning financial problems of the largest state enterprises has led to a renewed emphasis on privatization. Initial reviews indicate a positive impact on firm performance, with higher wages for workers but starkly lower overall employment (Jefferson and others 2003).

In Sub-Saharan Africa, sales rose from 175 in 1990 to more than 400 in 1996 and then jumped to more than 2,200 by 1998,² with sales concentrated in Mozambique, Angola, Ghana, Zambia, Kenya, Tanzania, and Guinea (Campbell White and Bhatia 1998; Nellis 2003). But the small size of most divested firms limited the financial impact: African sales accounted for 3 percent of total developing country proceeds during 1990–99. Estimated revenues over the decade for 37 Sub-Saharan states totaled $9 billion, less than the amount raised by sales in New Zealand alone (Nellis 2003) and about a third of the value of two Brazilian telecommunications auctions in the mid-1990s.

In the Middle East and North Africa, privatization revenues have been modest, below even those of Sub-Saharan Africa. Revenues grew in the late 1990s, largely as a result of Morocco’s telecommunications sale and the privatization of cement and other large and medium-size companies in Egypt.

In South Asia, Sri Lanka has had an active privatization program covering virtually all sectors (including infrastructure), but India has accounted for the bulk of regional proceeds through sales of minority shares in large companies and the recent sale of controlling stakes in a few large firms.

Impact of Privatization

In many ways privatization in the early years was a leap of faith. Few dispute that state enterprise performance had not lived up to expectations and that reforms short of ownership change had had modest effects. Still, there was neither great theoretical justification nor hard evidence at the beginning of the 1980s that the performance problems of state enterprises could be altered by change in ownership. Thus, the bulk of privatizations before 1992–93 took place in the absence of empirical support.

Then, during the 1990s, the privatization assessment industry grew rapidly. Assessments analyzed types of impacts and differences between enterprises in competitive markets and those in monopoly sectors. Most of the 100-plus assessments now available cover competitive enterprises, focusing on performance and comparing productivity and profitability and changes in output, investments, and capacity utilization before and after sale. The studies conclude that privatization improves performance and increases returns to new owners and shareholders, with more robust findings in high- and middle-income countries than in low-income countries.

In infrastructure sectors, the more crucial issue is the broader welfare effects of privatization—the net addition to or subtraction from societal wealth produced by
privatization, based on an assessment of the gains and losses to stakeholders. Given the daunting data and methodological demands of this approach, the few studies of the strict welfare type that have been done tend to cover middle-income or developed economies. They generally indicate that in the proper policy and regulatory settings, privatization substantially improves welfare compared with what would have happened had the enterprise remained under public ownership. As will be discussed, these findings are supported by more recent partial or limited welfare-type work, concentrated largely in Latin America.

There has been much concern about the employment and broader distributional impacts of privatization. Studies show large-scale job reductions in highly protected infrastructure sectors, though the number of workers dismissed as a result of divestiture is generally small relative to the total labor force. Analysts conclude that privatization is not a prime contributor to the large recent increases in general unemployment rates in developing economies. Wider income distribution issues are only beginning to receive analytical attention. Finally, the few available analyses of the fiscal and macroeconomic effects of privatization show fiscal benefits and a positive correlation between privatization and growth.

Enterprise Performance

Studies for a wide range of countries show that privatization improves enterprise performance. For firms in competitive markets (infrastructure firms are treated in the next section on welfare effects), profitability usually increases, often substantially, as do efficiency (measured by real sales per employee), output, and investment. These outcomes are seen in cross-cutting studies covering both developed and developing economies, as well as in case studies of developing and transition economies, but they are generally more robust for high- and middle-income countries than for low- or lower-middle-income countries.

Cross-cutting studies. Megginson and Netter (2001) survey cross-cutting studies that evaluate the impact of privatization on firm performance. The most comprehensive studies use a similar methodology to compare performance measures before and after privatization (over at least three-year periods) for large numbers of companies in developed and developing countries, privatized mainly through public share offerings (Megginson, Nash, and van Randeboorh 1994; Boubakri and Cosset 1998; D'Souza, Nash, and Megginson 2000). These studies find similar results. Weighted averages of the mean values show that profitability, defined as net income divided by sales, increases from an average of 8.6 percent before privatization to 12.6 percent afterward. Efficiency rises from an average of 96.9 percent in the year of privatization to 123.3 percent in the period after privatization. Between 79 and 86 percent of firms see increases in output per worker. Most firms achieve economically and statistically
significant increases in output (real sales) after privatization and significant decreases in leverage. Capital investment spending increases slightly, whereas employment changes are ambiguous (see later discussion). Accounting for most of the performance improvements are changes in the incentive and management structure, along with improved corporate governance.

One concern is that such cross-cutting studies suffer from selection bias. Firms sold by public offering might be the cream of the crop, because to meet stock exchange listing requirements they will have to have been profitable for some time, possess up-to-date and accurate accounts, and in general be among the largest and best-performing firms privatized. That might mean that performance improves not because of privatization but because firms with the highest potential are privatized. Selection bias may be intensified by the overrepresentation of developed economies in the samples of early studies, because of data limitations. Comparing accounting information across countries and at different points in time is also risky. Most of the studies do not account for exogenous changes in the macroeconomic or business environment, which can influence the outcomes of privatization. Finally, there is concern that using profitability as an indicator of improved performance is flawed because the private sector is by definition profit maximizing, whereas profit is not as salient a goal in the public sector.

Some studies address these methodological drawbacks. Boubakri and Cossett (1998) analyze the performance of 79 newly privatized firms in 21 developing economies between 1980 and 1992. Their sample is well diversified, with wide geographical coverage, countries of different levels of development, and firms of different sizes and in different industries and market structures. They find significant increases in profitability (124 percent higher on average after privatization), operating efficiency (real sales per employee up 25 percent on average and net income per employee up by 63 percent), capital investment spending (up 126 percent), and employment (up 1.3 percent) and a decline in leverage. The changes in profitability and efficiency were larger in middle-income countries than in low-income countries. One study that looks at the incidence and importance of selection bias finds no evidence of it in a set of privatizations in Central Europe (Frydman and others 1997).

There is also the possibility that performance improvements would have happened without a change of ownership—if general economic conditions improved and boosted all firms, or if public sector managers and owners were able to put in place and sustain reform measures. Omran (2001) reviews indicators in privatized and remaining state-owned firms in Egypt in the 1990s and finds that all firms improved, regardless of ownership type or change. He concluded that general liberalization was more important than privatization in explaining firm behavior. Few studies evaluate the counterfactual in any systematic way, however, the most notable being Galal and others (1994; discussed later).
Another issue is timing. Dewenter and Malatesta (1997) and Hodge (2000), although agreeing that ownership change produces positive effects, argue that many performance improvements occur well before privatization, while enterprises are still under state ownership. But these improvements are generally motivated by the “announcement” effects of divestiture. From Mexico to the United Kingdom, many long-avoided reforms were made in the run-up to privatization, including change of management, layoffs and other cost-cutting measures, and enhanced competition through changes in trade regime and pricing. The issue is whether these improvements could have been initiated and sustained had they not been precursors to divestiture. Proponents of privatization view it as necessary to “lock-in” the gains and prevent backsliding (World Bank 1995).

**Developing country studies.** In developing economies, most of the growing body of work assessing performance before and after privatization concludes that privatization improves enterprise performance. La Porta and López-de-Silanes (1997), in a study of 218 nonfinancial firms privatized in Mexico during 1983–91, conclude that state enterprises went from being highly unprofitable before privatization to being profitable thereafter, closing the performance gap with control groups of similar firms in the private sector. Output (inflation-adjusted) increased 54.3 percent, sales per employee roughly doubled, and profitability increased 24 percent. Controlling for changes in the macroeconomic environment, they find that improvements were due mainly to productivity gains resulting from better incentives and management associated with private ownership and partly to lower employment costs resulting from labor reductions.

In Brazil, privatization also improved the efficiency and profitability of state enterprises (Macedo 2000). During 1981–94, before privatization, the ratio of profits to net assets was negative, averaging −2.5 percent and falling to −5.4 percent toward the end. Significant gains were achieved after privatization. The large steel mill, which had been incurring heavy losses, became profitable, and investments increased dramatically. Higher profits brought more tax revenues to the government, and the company began paying dividends. Using a similar methodology to analyze the performance of 50 Brazilian state enterprises before and after privatization, Pinheiro (1996) concludes that privatization significantly improved performance, particularly when there was a change of control rather than a sale of only a minority stake. Results were stronger for companies that had been recently sold, indicating that privatization works better when combined with liberalization measures that remove barriers to entry and exit, result in positive interest rates, and reduce access to budget resources.

More—and more robust—success stories come from high- or middle-income countries than from low-income countries faced with difficult market conditions and wary investors. A study of 16 African firms (10 from North Africa) privatized through
public share offering during 1989–96 finds a significant increase in capital spending by privatized firms—but only insignificant changes in profitability, efficiency, output, and leverage (Boubakri and Cossett 1999). Still, a number of privatized companies in Sub-Saharan Africa increased capacity utilization through new investments, introduced new technology, and expanded markets (Campbell White and Bhatia 1998). Several recent examinations of country privatization programs—in Ghana (Appiah-Kubi 2001), Mozambique (Andreasson 1998), and Tanzania (Due and Temu 2002; Temu and Due 1998)—report strong performance improvements in privatized manufacturing, industrial, and service firms.

Transition economies. Assessing privatization’s impact in transition economies is more difficult. Concurrent sweeping economic and social changes compound the problem of separating privatization’s effects from other factors. Information and analytical shortcomings are particularly acute, especially for economies that were formerly part of the Soviet Union. Djankov and Murrell (2002), in reviewing empirical studies of enterprise restructuring and ownership change, conclude that private ownership produces more restructuring than state ownership in most transition economies. But regional differences are acute: The privatization effect in Central and Eastern European countries is more than twice that in former Soviet countries. In the enormously important case of Russia, they find little if any difference in performance between privatized and state firms.

Different sales methods and the types of owners they produce seem to account for much of this variance in outcomes. The Central and Eastern European countries that privatized largely through trade sales on a case-by-case basis (Estonia, Hungary, Poland) ended up with concentrated strategic owners, often foreigners, who tend to be more productive than diffused domestic shareholders. Firms in former Soviet countries, which relied mainly on mass privatization through vouchers, tended to have less positive results. Mass privatizations led to insider ownership (by managers and employees) and widely diffused shareholding among small, first-time equity holders—as evidenced most acutely in Russia, where 70 percent of the 13,000 enterprises privatized by vouchers in 1992–94 became insider-owned.

The hope was that these inside owners, supported by newly formed investment funds, would soon open their firms to outsiders with money and expertise. But insiders proved reluctant to give up control, and outside investors were wary of the unsettled circumstances of early transition Russia. The upshot of this failure to concentrate ownership through the secondary market was, for a time, limited restructuring. Subsequent nontransparent cash sales of the larger enterprises—exemplified by the notorious loans-for-shares program—helped create kleptocracy, as many high-potential firms were transferred to a small group of investors at very low prices (Black, Kraakman, and Tarassova 2000). Although some prominent economists (most notably Stiglitz 1999) condemned the process, others concluded that privatization,
unfair as it was, still led to performance improvements in Russia and other in transition economies, at least in firms where outsiders succeeded in securing control (Barberis and others 1996; Earle 1998; Earle and Estrin 1998; Aslund 2001). The argument is one of political economy: “Privatization in Russia worked considerably better than its politically feasible alternative: doing nothing” (Shleifer and Tresiman 2000: 38).

Method of sale and concentration of ownership do not explain all the variation in performance. Another important part of the explanation is differences in levels of institutional development and in policy approaches to new entry and hard budget constraints (discussed further later). In transition economies everywhere, the best performers were new private entrants—firms that were never in state hands.

Welfare Effects

Privatized infrastructure firms also recorded performance improvements. A recent survey shows that the introduction of incentives helped reduce costs and improve revenue collection in infrastructure firms (Harris 2003). Although gains were most dramatic in the telecommunications sector, in large part because of increased competition, substantial improvements took place in less competitive sectors such as power and water as well. Losses in the Chilean electricity sector, for example, more than halved after privatization and similar gains took place under more difficult circumstances in Georgia and Namibia.

But for infrastructure sectors, with their monopoly or network characteristics, the ultimate test for assessing privatization's impact is not simply firm performance, as it is for competitive firms, but the difference in economic welfare relative to what would have happened had the enterprise remained state-owned. Better financial and operational performance at the firm level is part of the story, but a transaction can benefit a firm and its owners without benefiting other stakeholders and society at large.

Selling an inefficient public sector monopoly to an unregulated private owner will almost certainly result in increased firm profitability and higher returns to the new shareholders and perhaps in higher salaries and expanded job opportunities for workers and greater returns to government. But these gains can easily be outweighed by the welfare losses imposed on consumers and the economy as a whole from inadequate access to products and services, their suboptimal supply, or their excessively high price.

Despite the recognized importance, only a few studies of the broader welfare effects have been undertaken. One problem is the heavy data demands. To estimate the counterfactual, analysts need detailed information on firm performance before and after the sale and equally detailed information on the policy climate, social outcomes, and myriad other factors. Even in the most data-rich settings, such as the United Kingdom, analysts readily admit that construction of the counterfactual, the
"what if." is inevitably based on a more than normal amount of "crystal ball gazing" (Newbery and Pollitt 1997).

A seminal study by Galal and others (1994) estimates the welfare consequences of privatization in 12 (mostly) infrastructure enterprises in one developed and three middle-income countries, looking at the effects on enterprise efficiency, subsequent investment, and consumer welfare. Employing a modified form of cost-benefit analysis, the study examines the impact on all actors, compares performance before and after privatization, and contrasts performance after privatization with a hypothetical scenario of reformed state ownership, with new technology and more rational procedures. The conclusion: Divestiture substantially improved economic welfare in 11 of the 12 cases, mainly due to a dramatic increase in investment, improved productivity, more rational pricing policies, and increased competition and effective regulation (figure 4).

In a study of the welfare effects of privatization of the electricity sector in the United Kingdom, Newbery and Pollitt (1997) argue that efficiency improved significantly in the first years following privatization. But from their counterfactual analysis, they conclude that the new private shareholders captured the bulk of the financial gains at the expense of government and taxpayers. Consumers/taxpayers did reap some benefits: It was not a case of winners and losers, but of huge winners

![Figure 4. Welfare Effects of Divesting 12 State Enterprises](image-url)

*Note: Welfare gains are presented as a percentage of annual sales in the last year before privatization. Source: Galal and others (1994).*

---

*The World Bank Research Observer, vol. 19, no. 1 (Spring 2004)*
and very small winners. Domah and Pollitt (2000) look at the welfare consequences of privatizing regional electricity supply and distribution in England and Wales and estimate large welfare gains, first to the seller and then to consumers.

There has been just one application of the strict welfare methodology in low-income countries, Jones, Jammal, and Gokgur's (1998) study of 81 privatizations in Côte d'Ivoire. The analysis covers not just infrastructure firms but also a range of firms already operating in competitive markets (in agriculture, agro-industries, and tradable and nontradable sectors). For the entire privatized sector, they conclude that there were substantial benefits: Firms performed better after privatization than before, they performed better than they would have had they remained under public ownership, and the set of transactions as a whole contributed positively to economic welfare, with annual net welfare benefits equivalent to about 25 percent of predivestiture sales. These results stemmed from a number of effects, including increases in output, investment, labor productivity, and intermediate input productivity.

In recent years, several partial or limited welfare analyses of privatization have been conducted, most focusing on Latin America. One set of studies sponsored by the World Institute for Development Economics Research of the United Nations University developed less elaborate counterfactual constructions in an effort to measure the social effects of privatization of infrastructure and network industries. The Bolivia study (Barja and Urquiola 2001) simply compares performance before and after privatization, looking for breaks in trend lines at the time of sale and for differences in the economic and political frameworks in the two periods, and estimates the possible effects on performance after privatization. In Peru, Törero and Pascó-Font (2001) limit themselves to estimating the effects of infrastructure privatization on different income groups, using a rudimentary counterfactual that assumes no change of ownership and the continuation of preprivatization pricing policies. A similar study for Argentina (Delfino and Casarin 2001) finds varying social outcomes depending on which of three price elasticity of demand figures is used, that all three show increases in production, quality, and access following divestiture. Almost all the studies conclude that on the whole, consumer surplus expanded after privatization, even though in most cases prices also increased.

Another recent set of Latin American studies—covering Argentina, Bolivia, Mexico, and Nicaragua—examines the distributive impact of privatization. Based on the country studies, McKenzie and Mookherjee (2002) calculate the welfare impact on consumers from changes in price, access, and quality of service, taking into account changes for different expenditure groups. They find that in all cases privatization resulted in increased access to services, especially for poorer consumers, who had less access to begin with. Prices fell in half the cases and rose in half, though the positive distributional gains from access outweighed the impact of increased prices. Privatization was generally followed by an improvement in service quality. In
almost no cases was there evidence that changes in price or access led to increased poverty.

Regarding access, Harris (2003) confirms a positive impact in such sectors as water and power in countries at various levels of development, as increased investment leads to expanded coverage and access and as sales contracts often require that much of the expansion benefit previously unfavored groups or regions. Clarke and Wallsten (2002) use household data from around the world to examine the performance of public utilities in meeting universal service obligations and the impacts of reform. They find that state monopolies everywhere except in Eastern Europe failed to provide service to poor and rural households and that privatization reforms did not harm poor and rural consumers and in many cases improved their access to utility services. Prices, as noted, often rise following privatization to offset prices that have long been well below cost. In many instances this has a negative distributional impact—but in addition to being outweighed by access, these effects can be muted by regulatory frameworks or subsidies aimed at protecting the less favored. Chile, for example, subsidized telephone costs in rural regions. (For a general discussion of mitigating measures, see Estache, Foster, and Wodon 2002; Clarke and Wallsten 2002; and Estache, Gomez-Lobo, and Leipziger 2000.)

The welfare studies conclude that privatization generally increases the resources available in the economy, including those available to governments. The studies also conclude that although few privatizations result in gains for all stakeholders (sellers, buyers, consumers, workers, and competitors), most produce gains for some and losses for others, depending on how the transaction is structured, the policy framework, and the institutional development and competence in the economy. Although the distribution of gains and losses varies in the studies, in almost no case do the new private owners come out on the losing side. The variance is larger for other stakeholders, including consumers, workers, and sellers. Even in cases of increased consumer surplus, the distribution of gains varies by income decile.

What is clear from the studies is that the aggregate gains are greatest when privatization is combined with proper competition policies and regulatory frameworks. The welfare effects have been shown to depend crucially on the fairness and capacity of the regulatory system. Chisari, Estache, and Romero (1999:376) conclude that “how serious governments are about the fair distribution of gains from privatization reform is revealed by how serious they are about regulation.” (Regulatory issues are treated in greater detail later.)

Employment and Distributional Effects

State enterprises tend to be overstaffed. Consider these examples. In Sri Lanka in 1992, estimated redundancy in eight of the largest firms (in electricity, railways, shipping, sugar, cement, and petroleum) averaged 53 percent (Salih 2000). Prior to
privatization. Argentine railways, with more than 90,000 employees, had a wage bill equivalent to 160 percent of the firm's total revenues (Ramamurti 1997). Such levels of overstaffing contributed to the financial weakness of state enterprises.

Excess labor is one of the first cost areas addressed by reforming governments or new private owners. A recent survey of 308 privatized firms shows worker reductions in close to 80 percent of firms after privatization (Chong and López-de-Silanes 2002). An earlier review of 17 privatizations found job increases in 4 (averaging 23 percent), no change in 6, and job losses in 7 (averaging a substantial 44 percent of the workforce before privatization), predominantly in tobacco, water supply, and electricity (Van der Hoeven and Sziracki 1997).

A number of highly protected and deeply politicized enterprises have seen huge declines in net employment, often before but also after privatization: 80 percent in Argentina's railways, 72 percent in petroleum, and 50 percent in electricity enterprises; 82 percent in Brazil's railroads; 42 percent in Manila water; and 50 percent in a study of Mexican firms. Moreover, although D'Souza and Megginson's (1999) study of 78 privatized firms in 25 countries finds insignificant employment declines for the group as a whole, reductions were substantial for a subgroup of noncompetitive firms. Much of this labor shedding was required to bring employment and labor costs in line with that in competitor or similar firms.

Some firms in competitive sectors, with relatively efficient staffing levels before privatization, and firms in high-demand sectors, such as telecommunications, experienced little decline in employment (Galal and others 1994; Megginson, Nash, and van Randenborgh 1994; Boubakri and Cosset 1998; Kikeri 1998). The general expectation is that downsizing will be temporary and that growth in restructured private firms will rebound. In some cases, this has happened. In a set of Eastern European countries, employment declined just before and during privatization, but subsequently increased (Estrin and Svejnar 1998). Jones, Jammal, and Gokgur (1998) confirm labor shedding before privatization in their study of Côte d'Ivoire. But they find that firms slated for privatization shed less than the economy as a whole, suggesting that layoffs were a response to weak economic conditions rather than to privatization itself. Indeed, they argue that privatization may have resulted in less labor shedding, possibly to minimize workers' opposition to privatization. Moreover, the privatized sector did significantly better than enterprises as a whole in terms of job generation.

Several studies report that employees who retain their jobs in privatized firms receive the same or higher wages than they did before. In Brazil, for example, employment reductions were sizable in large firms privatized in the 1990s (48 percent on average), but productivity improvements resulting from restructuring led to higher wages and performance-based incentives for workers who remained (Macedo 2000). Similar evidence is found in Argentina (Ramamurti 1997), Côte d'Ivoire
(Jones, Jammal, and Gokgur 1998), Malaysia (Galal and others 1994), and Mexico (La Porta and López-de-Silanes 1997).

There is a widespread public perception that privatization is the main cause of the large increases in unemployment in many regions in the last decade and thus a prime contributor to a number of social ills, including increased poverty and inequality. Recent research suggests a more nuanced picture. Privatization has contributed to general unemployment levels, but only slightly. It is rarely a principal or even important cause of rising unemployment, poverty, or inequality. The evidence comes mainly from Latin American studies that argue that while privatization has resulted in job losses, the aggregate numbers are small relative to the national workforce (Barja and Urquiola 2002; López-Calva and Rosellon 2002; Ennis and Pinto 2002; Freije and Rivas 2002, with findings summarized in McKenzie and Mookherjee 2002). They conclude that privatization is not a main cause of overall increases in unemployment and wage differentials, even where both have risen dramatically.

Other studies support this view. Data from Argentina suggest that privatization was not a major contributor to the rise in unemployment between 1993 and 1995 but that the interest rate shock from regional instability was (Chisari, Estache, Romero 1999). In the early 1990s in Hungary and Poland, despite the slow pace of privatization, official unemployment grew rapidly, reaching 14.1 percent in Hungary in 1993 and 16.7 percent in Poland in 1993–94 (Nellis 1999). Behrman, Birdsall, and Szekely (2000), in an econometric study of the impact of economic liberalization on wage differentials in Latin America, conclude that privatization was negatively correlated with the growing wage inequality in reforming Latin American economies. Rather, privatization was mitigating the “disequalizing effects” of liberalizing reforms in the financial sector, the tax regime, and capital markets.

Finally, what of the broader distributional effects? Despite innovations aimed at spreading equity holdings—such as voucher schemes in transition economies and “capitalization” in Bolivia—there is a widespread perception, even among observers sympathetic to privatization, that privatization has had negative effects on wealth distribution, with upper-income groups gaining far more in equity shares than lower-income groups, at least in the short run.

Birdsall and Nellis (2002) review the available studies and conclude that most privatization programs have, at least in the short run, worsened the distribution of assets (very likely) and incomes (likely). This is far more evident in transition economies than in Latin America, and less clear for utilities—where the poor have tended to benefit from greater access—than for banks and oil companies and other natural resource producers. In the best-studied cases from Latin America, the conclusion is that privatization has contributed little or nothing to the growing inequality in the region and that it either reduces poverty or has no effect (McKenzie and Mookherjee 2002).
Macroeconomic and Fiscal Effects

The macroeconomic effects of privatization have been less studied than other aspects. One of the few such studies, by Davis and others (2000), calculates significant and positive benefits. Governments tended to be financially better off after privatization than before. Gross proceeds from privatization were substantial, amounting to 2 percent of GDP in a sample of 18 countries, and the fiscal situation of governments that saved rather than spent privatization proceeds improved over time. Privatization produced other positive impacts on government revenue. Not surprisingly, government transfers to state enterprises declined substantially following privatization (figure 5), and broader indicators of consolidated state enterprise accounts for a number of countries indicated much smaller deficits.

The study also finds a positive correlation between privatization and overall rates of growth. The authors argue that although privatization is not the sole cause of subsequent increases in growth rates, it is a good proxy for the range of structural reform measures that contribute to the overall result. Investors and markets view privatization as an indicator of reform credibility, a less tangible but important macroeconomic effect.

Sheshinski and López-Calva (1998) also find that privatization improves the public sector's financial health. Budget deficits decline during the reform period. Low-income countries that are less aggressive privatizers have a larger deficit, on average. In high- and middle-income countries, privatization reduces net transfers

---

**Figure 5.** Gross Budgetary Transfers and Subsidies to State Enterprises for Selected Countries (Percent of GDP)

Excludes Petróleos Mexicanos and includes some decentralized government agencies.  
*Source:* Davis and others (2000).
to state enterprises, and transfers become positive when the government starts collecting taxes from privatized firms—another contributor to positive macro-economic effects. Despite concerns about the difficulties of tax collection, there is evidence of increased downstream revenues to governments through higher taxes in Africa (Campbell White and Bhatia 1988), Argentina (Shaikh and others 1996), Brazil (Macedo 2000), and Mexico (La Porta and López-de-Silanes 1997), among others.

On the use of net proceeds from privatization, the conventional wisdom is that the more they are devoted to retiring debt, the better. Debt reduction lowers interest rates, reduces further borrowing and inflation, and boosts overall growth. Four countries in the study by Davis and others (2000)—Argentina, Egypt, Hungary, and Mexico—had an initial stock of registered public debt ranging from 40 percent to 130 percent of GDP. Privatization proceeds contributed to a sharp decline in public debt between the year before the period of most active privatization and the last year of active privatization, helping strengthen and stabilize the economy. One problem is that there are strong claimants to the proceeds other than debt relief, and politicians often are obliged to balance the economic ideal with the politically feasible. Many governments (Argentina, Bolivia, Estonia, Hungary) have devoted a portion of privatization proceeds to covering pension costs (a form of debt retirement, because they are obligations of the state) or—a much more risky use—to the restructuring of some key state enterprises before or even instead of privatization (in Poland and the Czech Republic). Experience cautions against the use of proceeds to finance current expenditures, given the one-time nature of the proceeds and the risk that spending may become entrenched at unsustainable levels.

**Emerging Issues**

Despite the largely positive economic assessments, privatization is increasingly disliked—in general and in network industries in particular. Public opinion surveys from Western Europe, Latin America, South Asia, and Russia reveal that large and growing percentages of citizens view privatization as a harmful policy, often imposed by external agencies without due consideration for the economic and social context. Protests of higher prices, corrupt transactions, alienation of national assets to foreign investors, and job losses have become common—and in several cases, deadly—in Latin America, Sub-Saharan Africa, and India. Even in the United States, troubled partial privatizations (water in Atlanta) and flawed liberalization (associated in the public mind with privatization, such as electricity in California) have led many to oppose further reductions of state involvement, particularly in sectors such as water and power. Journalists in both developing and industrial countries often portray privatization in a negative light (see, for example, Finnegan 2002). Nongovernmental organization activists and analysts, supported by such noted economists...
as Joseph Stiglitz, question both the manner in which privatization has been implemented (without adequate regard for local conditions) and, at times, its essential rationale.

The causes of this discontent are several. First, when done badly, privatizations have simply gone wrong. In Russia and other economies in transition, hundreds of firms were more or less given away to a small group of agile or well-connected insiders. Many of the highest potential assets were sold at very low prices. Even when ordinary citizens obtained shares (in exchange for vouchers), this equity quickly lost most of its value. The bad outcomes may be few compared with the total number of privatization transactions, but they are highly visible and have helped mobilize popular opposition—failed privatized toll roads in Mexico; troubled or canceled infrastructure concessions in Bolivia, Costa Rica, Hungary, India, Indonesia, Senegal, Ukraine, and half a dozen other countries; a privatized mine in Zambia that failed to obey the law on severance payments.

Second, and related, the process is political dynamite. The benefits of privatization tend to be diffused among consumers or citizens as a whole, small for each affected individual, and slow to materialize. To illustrate, a 10 percent reduction in electricity tariffs amounts to a large sum in the aggregate but might be of little significance to individual consumers. It normally takes some time before the other possible benefits of privatized electricity generation—less pollution, more reliable service that might stimulate private investment elsewhere in the economy—are discernible. The costs of privatization, however, are likely to be concentrated (dismissed workers, increased charges), large for each individual affected, and rapidly apparent. Politically, losses are felt more intensely than gains, with proreform consumers generally unorganized, silent, and nearly invisible politically, and antireform workers, civil servants, and antimarket intellectuals frequently organized and vocal. The political process naturally responds to the intense and those with “voice.”

Third, privatization may have been oversold, particularly in countries with weak institutional capacity. Proponents underplayed the costs, predicted quick and widely shared gains, insisted on speed, and often claimed that ownership change was critical to a general economic turnaround. With the benefits overplayed, it is not surprising that privatization is blamed when general economic conditions fail to improve or when competitive markets and regulatory frameworks are not sufficiently developed to support privatization.

Going forward, what can be done to address these concerns? There is still much debate about ownership and competition that needs to be addressed. Equally important is improving the privatization process itself, taking measures to promote competition, putting in place proper regulatory frameworks, ensuring transparency, mitigating social costs, and tailoring privatization to local conditions.
Ownership or Competition?

Two decades of experience have not settled the debate over how much ownership matters. Neoclassical economic theory is somewhat agnostic on the effects of ownership. It regards market structure and the degree of competition to be of equal or greater importance. Empirically, some analysts conclude that increased exposure to competition accounts for most of the positive changes seen in privatized firms. Indeed, some studies suggest strongly that competition has been more important than ownership change in bringing about efficiency gains.

But the question remains: If enhanced competition and market restructuring are so efficacious on their own, why do the efficiency effects rarely occur or persist in the absence of ownership change? The poor experience with state enterprise reform in the 1980s and before shows the difficulties and limited results of reforms short of ownership change (see Shirley 1983; Nellis and Shirley 1991). Governments found it difficult to apply the full package of reforms needed (exposing state enterprises to competition, requiring them to access private capital markets for investment funds, creating a market for managers, isolating the process from political interference) and to leave the package in place long enough to change incentives and behaviors. Even where performance improved, reforms did not endure, usually because of renewed political interference (World Bank 1995; Majumdar 1998; Shirley and Xu 2000). Governments almost never allowed insolvent state enterprises to fail and go out of business. Soft budgets continued, and in the absence of exit options, there was little pressure on government officials, managers, and workers to reform.

The difficulty of reforming state enterprises without privatization and the generally improved performance after privatization support the importance of ownership but do not conclusively prove it. Those who look at the issue statistically rather than causally argue that ownership change is associated with effective and enduring competition. For example, Shirley and Walsh (2000) sum up the ownership or reform debate based on a review of some 50 empirical studies covering a variety of countries and sectors. They find greater ambiguity about ownership in the theoretical literature than in the empirical literature. The clear majority of empirical studies concluded that privatized—and private—firms perform better than state enterprises, a finding that is robust across sectors and market structures and across developed and developing countries. Although a few studies find better performance by state enterprises in infrastructure sectors in developed economies, no studies find better performance by state enterprise in any sector or market situation in developing economies. Shirley and Walsh find that private firms do better in fully competitive markets. This advantage persists but is less pronounced in monopolistic markets, and the evidence is less conclusive.

The issue is less one of privatization versus competition. Rather, privatization and procompetition policies appear to be mutually reinforcing. Sachs, Zinnes, and Ellat
(2000) examine the empirical evidence across 24 transition economies and conclude that ownership change is not enough to generate improvements in economic performance. But when ownership change is combined with institutional reforms—aimed at removing barriers to entry and exit, improving prudential regulation and corporate governance, hardening budget constraints, and developing capital markets—progress is much greater. Maximum impact is achieved when market competitiveness, hardened budget constraints, and improved regulatory frameworks coincide with privatization. The higher the level of institutional reforms, the more positive the economic performance impact from a change of ownership. But institutional reforms do not guarantee performance improvements unless there is a minimum level of ownership change: The key finding is that economies must have private ownership and pro-competition policies to progress. So, again: Ownership matters, but policies and institutions matter just as much.

**Promoting Competition**

Privatized firms perform better than state enterprises, but new private firms perform better than both. Promoting competition by removing entry and exit barriers and by linking privatization with financial sector reforms is crucial for the development of a dynamic and competitive private sector and thus for successful privatization.

*Entry and exit.* Particularly in infrastructure firms, economic benefits are maximized when privatization is combined with new entry, the break-up of large entities, price deregulation, and the development of effective regulatory frameworks—with the last being critical for equity as well as efficiency. Competitive markets and good regulation reinforce the benefits of private ownership. Divesting into competitive markets may reduce the revenues from sale, but efficiency, not revenues, should be the primary objective of privatization. A recent report (World Bank 2004) shows that earlier privatizations that granted long exclusivity periods to investors increased the sales price but led to problems as exclusivity reduced competition and thereby incentives for investment. In telecommunications, for instance, exclusivity reduced network expansion by 10–40 percent and the annual growth rate of the network by more than 2 percent. As countries have learned from experiences, exclusivity periods have grown shorter and some privatizations have taken place without any exclusivity.

Competition is equally important in tradable sectors. It requires eliminating import restrictions, deregulating prices, and simplifying procedures for starting a business. Removing entry barriers is particularly important in transition economies, where state enterprises dominated all markets and where restrictions on private participation and entry were powerful. For example, privatization of large state enterprises in Poland initially proceeded slowly, though most small firms were sold off quickly. Entry was permitted and vigorously encouraged, and harder budgets were
at least temporarily imposed on most state enterprises that remained. Competition increased in all sectors of the economy, and asset stripping was minimized. However, the long and increasingly expensive delay in dealing with some large, overstaffed, loss-making firms, which are still in state hands, is a major contributor to Poland's recent economic difficulties. China's success is also in good part due to opening entry to domestic quasi-private enterprises and to foreign investment. In changing the public-private mix, privatization was for a long time less important than the emergence of new private businesses, although here, too, privatization of the larger firms is gaining ground and made easier by the presence of a competitive private sector.

Ultimately, competition means freedom to fail. Closure signals mismanagement, so governments have seldom allowed even obviously nonviable firms to go under. Privatization facilitates the liquidation or exit of nonviable firms. When purchasers have incorrectly estimated the market or their ability to restructure firms, closures have resulted: private owners have been able to do what public owners could not. Critics point to high liquidation rates after privatization (in places as different as Armenia and Guinea) as evidence that privatization is a failure. But closures do not necessarily indicate that privatization was misguided. Had these firms remained in state hands, they would likely have continued to receive subsidies, using scarce resources with high opportunity costs. Given the political difficulties associated with shutting down state enterprises, privatization may allow the liberation and transfer of assets from problematic management in the public sector to better management and more productive use in the private sector.

Linkages to financial sector reforms. Competition is engendered by the quality, pace, and scope of financial sector reforms, which in turn affect privatization outcomes. For example, the poor performance of privatized firms in many transition countries, including the Czech Republic, Slovakia, and Russia, resulted partly because insiders obtained control of the assets, but also because of the absence of financial sector reforms that would have forced even bad managers to take the right steps—or leave the way clear for other owners.

In the Czech Republic, the state continued to dominate commercial banking throughout the 1990s, maintaining essentially a controlling interest. Pressure from government allowed weak firms to borrow to stay in business. Little of the credit was applied to restructuring and much of the debt was nonperforming, with a fair percentage of it stolen (Cull, Matesova, and Shirley 2001). This severely weakened commercial banks and resulted in an enormous bailout. The Czech economy has largely recovered, but at a great cost. By contrast, Estonia and Hungary had less problematic privatizations because they implemented their bank restructuring and privatization programs early; dealt rapidly with the bad debt problems; tackled difficult legal and institutional reforms, such as bankruptcy and protection of minority shareholders; and partly as a consequence, received more inflows of foreign direct
investment in the 1990s than their regional competitors. These factors, added to sales methods favoring concentrated ownership, produced more positive privatization outcomes.

A World Bank (2001a) report on finance and growth finds that the lower the income of a country, the higher the proportion of its bank assets that are state-owned. The theory was that state-owned banks would distribute capital to more productive investments, provide greater access to credit for deserving sections of society, and be less prone to crises. In practice, it has proven difficult to design incentives to guide either public or private banks toward efficient resource allocation. With state-owned banks, incentives are often especially weak, for political and other reasons, and the results are usually worse. State-owned banks have incurred some of the largest losses of recent times. State ownership tends to reduce competition through higher spreads on interest rates, leads to less stock exchange activity and nonbank credit, and results in greater concentration of credit, usually to the largest 20 firms, often inefficient state enterprises.

Privatization of banks is thus part of the solution. But experience in Argentina, Chile, and Mexico, among others, shows that bank privatization is a special case, more akin to infrastructure transactions than to privatizations in competitive sectors. As in infrastructure, good policy, monitoring, and enforcement are key. In weak regulatory environments, poorly designed and implemented bank privatizations have provoked major financial crises (Chile in the late 1970s, Mexico in the early 1990s). Although these rapid, insufficiently thought-out privatizations were mistakes, extended state ownership or control can have an equally detrimental impact, as the Czech case illustrates. Just as hasty privatization in weak environments can lead to problems, so can excessive delays, which undermine real sector reforms and impose high costs. Bank privatization remains the preferred strategy, but it requires caution and far more preparation than privatizations of ordinary commercial firms. Experience shows the need to sequence the phasing out of state ownership with the opening up of entry for private banks and improvements in the regulatory environment (World Bank 2001a).

Regulation

Successful privatization of natural monopolies requires development of regulatory frameworks and institutions that are independent, accountable, and resistant to capture by the private provider or the state. Such frameworks are essential to protect consumers against abuses of monopoly power, assure investors that they will be fairly treated, and address broader equity concerns.

Alexander and Estache (1999) find that Latin American countries, such as Chile, that devised regulatory frameworks up front and developed reasonable capacity to implement and enforce regulations had better success with privatization. Guasch.
Laffont and Straub (2003) also find that having well-defined regulations and a regulatory agency at the time of privatization lessens the need for subsequent renegotiations of contracts. The regulatory framework reduces the scope for error and the need for subsequent modifications that are time-consuming, disruptive, and signal wavering commitment. Wallsten (2002), in one of the few empirical studies on sequencing regulation and privatization, finds that establishing a regulatory authority before privatization of the telecommunications sector in Latin America and Africa was correlated with increased telephone penetration, investment, and mobile cellular subscriptions. Attention to the regulatory framework before privatization also increases the price investors are willing to pay for the firm.

There is widespread acceptance of the key elements of a good regulatory framework, including the need for coherent policies, transparency and public disclosure, predictability in the rules of the game, a proper balance between autonomy and accountability, and adequate institutional capacity. During the 1990s, developing economies created some 200 regulatory agencies with these elements in mind as part of infrastructure restructuring (Estache and others 2003). How have these regulatory frameworks and agencies worked in practice?

A recent assessment concludes that experience has been mixed (World Bank 2004). The report acknowledges progress in establishing independent regulatory agencies and notes that some agencies work well. But it argues that the technical and political complexities, with often conflicting objectives and tradeoffs, have made regulation one of the most challenging aspects of privatization. Political interference occurred at many levels—most critically in the sensitive area of service charges, but also in funding, the appointment of civil servants to regulatory boards, and the staffing of regulatory agencies. These factors weakened the independence and effectiveness of regulatory agencies, which in extreme cases became mere extensions of their ministries.

Lack of transparency is another problem. Citing surveys of regulators in sectors such as telecommunications and power, the report suggests that some agencies have no statutory obligation to explain their decisions, and some that do still fail to open meetings to the public. Also contributing to poor regulatory performance are inadequate data collection processes and the absence of quantitative models to estimate the impact of regulatory decisions on key financial and economic indicators affecting operators, consumers, and the government (Estache and others 2003).

Regulation has proven particularly difficult in low-income countries with weak overall institutional capacity. Governments in these countries have found it particularly difficult to regulate powerful economic actors fairly and effectively, to the detriment of privatization efforts. Among the problems they face are inexperience dealing with complex technical, legal, and financial issues; few if any precedents to build on; little or no reliable information on cost and performance; no watchdog groups; and no experience with independent agencies (Smith 2000).
Repeated efforts by governments and donors to build regulatory capacity, particularly in very poor countries, have produced few successes (urban water supply in Côte d’Ivoire is one). One reason is the attempt to transfer models and approaches from developed economies into developing economies without taking into account their differing political, legal, and institutional circumstances. Another is that changing ownership takes much less time than developing regulatory capacity, and the lower the income level, the slower the developmental process. Newbery (2001) points out that it took regulators in the United Kingdom five years to distribute to consumers some of the substantial efficiency gains produced by privatization of the electricity industry. Quick results should not be expected in developing economies, particularly in low-income countries.

The strongest lesson of experience is that there is no universal model: Regulatory frameworks need to take into account each country’s unique political, legal, and institutional context (World Bank 2004). Better design and sustainability of regulatory frameworks are equally important. Alexander and Estache (1999) note that this generally involves creating clear and reasonable incentives; establishing competition, which drives incentives for efficiency and simplifies regulation; addressing the details of regulation (average changes, quality levels and penalties for not meeting them, interconnection rules) early in the process; and putting in place clear rules to ensure that all information is available in a timely and consistent manner. Estache and others (2003) highlight two additional steps to improve regulatory performance. One is to develop quantitative models that take into account the initial condition of the service, the objectives, and the regulatory instruments, thus allowing regulators to analyze sensitivities and simulate scenarios while minimizing subjectivity. The second is to do a better job of educating the general public, particularly on the tradeoffs between efficiency and fairness of service charges, which are the most acute and politically visible issues.

Transparency

Along with fears of increased unemployment and concerns about selling national assets to foreigners, a leading political concern in privatization is fear of nontransparent and corrupt transactions. Lack of transparency leads to allegations—and documented cases—of corruption, provides ammunition to opponents, creates backlash from investors and the public at large, and threats to halt or even reverse privatization and liberalizing reform in general. In Latin America and elsewhere, surveys documenting the continuing decline of support for privatization reveal deep dissatisfaction with the perceived incidence and severity of corruption (Lora and Panizza 2002).

Strengthening transparency requires a host of measures. Especially important is ensuring that transactions occur without special privileges for insiders or other
favored purchasers. Among other measures, that requires adhering strictly to standard, well-publicized procedures: vetting actions by the press or other outsiders (such as Transparency International); opening bids on TV or in public sessions; and publicizing the terms of the transaction or the privatization contracts. Promoting competition in the privatization transaction—from the selection of advisers to the selection of the final buyer—may be the most effective way to support transparency—and it also yields the maximum economic and financial benefits.

Although in some cases negotiated sales may be the only option, in general the greater the openness and competition in the selection process, the greater the number of bidders, the higher the price paid—and the higher the level of public acceptance and satisfaction. In Mexico, La Porta and López-de-Silanes (1997) found that an additional bidder participating in a tender increases the net revenues to government by 12 percent. Public offerings are widely regarded as the most transparent sales method, but most developing economies do not possess the capital markets, quality firms, or business environments required to apply this method.

Although enhancing transparency takes time, delays in privatization entail costs. Evidence from Bulgaria, Mexico, South Africa, and elsewhere indicates that once a firm is slated for privatization, delays in completing the transaction lead to declining operations, asset stripping, and lower sales price. La Porta and López-de-Silanes (1997) found in Mexico that net revenues to the government dropped 24 percent for each year that privatization was delayed. Ultimately, few bidders may come forth to buy the diminished assets, creating pressure on government to make special and costly concessions. But selling firms, particularly infrastructure firms, without first enhancing competition and regulation, and thus transparency, has proven even more costly. The general rule should be to move swiftly in privatizing firms operating in competitive or potentially competitive markets, but to take the time to get the market and regulatory structures right when privatizing infrastructure firms or banks.

Social Safety Nets

The extensive labor force reductions that usually accompany the restructuring of large state enterprises often lead to political backlash. One way to reduce tensions is to engage in dialogue with employees early on and to jointly work out acceptable solutions. Generally, this involves compensation payments, free or low-priced shares in the privatized firm, and augmented retirement and severance benefits to encourage voluntary departures in place of layoffs. Voluntary departures are often considered more politically and socially acceptable, and the financial and economic returns can be high. But such programs can be quite costly in the short run and can result in adverse selection (the best, most mobile workers apply to leave) and, in the case of early retirement payments, heavy financial burdens on the social security system.
(Kikeri 1998; Rama 1999). For these reasons, Chong and López-de-Silanes (2002) argue that the optimal economic approach for governments to follow in any labor restructuring is simply not to intervene.

Compensation packages have often been combined with retraining to help workers reintegrate into the labor market. Although retraining programs are popular with governments and donors alike, the few available evaluations question their cost effectiveness (Dar and Tzannatos 1999). Targeting training to those who request it and who possess characteristics that increase the likelihood of putting the training to use improves the chance of success. Counseling and job search assistance have been found to be more cost effective than training (Dar and Tzannatos 1999). Contracting or outsourcing arrangements are another option. In Argentina, 5,000 surplus workers in the privatized oil company started 200 private businesses contracting with the oil company (Kikeri 1998). This approach has been tried in a number of other countries and sectors as well.

Concluding Comments

Mounting empirical evidence of privatization's benefits coincides with increasing dissatisfaction and opposition among citizens and policymakers. This dissatisfaction reflects the growing questioning of the benefits of privatization, particularly of large infrastructure and network industries, among large segments of affected populations. It reflects the general downturn of global markets in the past few years—and the collapse of several iconic firms—and the resulting swing of the economic pendulum back toward stability and accountability, and thus increased governmental supervision. It reflects the overselling of privatization as a panacea for all of a country's economic problems. It reflects the concern that privatization does not produce macroeconomic and distributional gains equivalent to its microeconomic benefits and that the transactions will be handled corruptly, with the proceeds lost or stolen.

On the other side of the ledger, we know that ownership change in productive firms, as well as private involvement in a less than full ownership capacity, usually improves the financial situation of the firm and the fiscal position of the selling government, increases returns to shareholders, and in the right policy circumstances, generates significant welfare benefits as well. These are major achievements. But they have not been able to offset the negative views of privatization, the fears that it creates or adds to injustice, inequity, and instability.

The evidence suggests that renewed efforts to reform state enterprises by methods short of ownership change or private involvement will not prove effective. The costs of no or slow privatization can be high. Thus, despite the problems, privatization should be neither abandoned nor reversed. Rather, there should be a strengthening and redoubling of efforts to privatize correctly. This means more advance analyses

Suinita Kikeri and John Nellis
and better tailoring of privatization to local conditions. In countries with weak capacities, it also means emphasizing the policy and institutional underpinnings of market operations rather than focusing solely on privatization transactions. This involves developing and protecting competitive forces, creating proper regulatory frameworks (essential for both efficiency and equity) before privatization, introducing and enforcing transparency in sales processes, developing social safety nets for the adversely affected, and introducing innovative pricing and subsidy mechanisms to ensure that the poor have access to affordable essential services.

Notes

Sunita Kikeri is advisor in the Investment Climate Department at the World Bank; her e-mail address is skikeri@worldbank.org. John Nellis is senior fellow at the Center for Global Development, Washington, D.C.; his e-mail address is john.nellis@verizon.net. The authors would like to thank Michael Klein and Neil Roger for their comments on an earlier draft on which this article is based. Anonymous reviewers provided valuable comments.

1. This section uses data on global privatization proceeds from the World Bank’s Global Development Finance 2001 (World Bank 2001b). Data on privatization revenues must be viewed cautiously because they can be concentrated in a few large public offerings and thus do not reflect the scope or progress of a country’s overall program.

2. Those numbers include small retail units and sales of minority shares.

3. However, previous liberalization reforms that did not include privatization had accomplished little in Egypt. It could be argued that only when privatization was a realistic option and credible threat did the remaining state enterprise managers take seriously the calls for reform.

4. The project, titled “Impact of Infrastructure Reforms in Latin America,” produced papers on Argentina, Bolivia, Chile, Peru, and Spain and on various aspects of measurement and assessment of the social impact of utility privatization.

5. Most of the countries covered in the study had an International Monetary Fund program in place, with limitations on the deficit that may have influenced this finding.

6. Less than 2 percent of infrastructure concessions have been canceled (Harris and others 2003).

7. Tandon (1995) argues that there are many cases where privatization appears to have “resulted” in efficiency improvement, but that in most of the cases privatization was contemporaneous with deregulation or other types of competition-enhancing measures.

8. In reviewing the liberalization and privatization of the British electricity sector, Newberry and Pollitt (1997) argue that competition rather than privatization improved performance. They show that the efficiency gains were considerably less in the parts of the sector that were privatized but not liberalized than in other parts.

9. China is a rare case of evolutionary success, and even in China the problem of poorly performing core state enterprises has been postponed, not resolved.

References


Managing Oil Price Risk in Developing Countries

Julia Devlin • Sheridan Titman

This article presents a simple framework for understanding the impact of oil dependence on growth in terms of an optimal savings and investment strategy. Among the more important factors underlying this strategy is the extent to which oil price changes are temporary or permanent. This in turn determines whether a country should rely on stabilization and savings funds or the use of financial instruments to manage oil revenues—or both. Country experiences with stabilization and savings funds are surveyed, and the case is presented for using financial instruments to manage oil price risk. Policy implications for enhancing the use of financial instruments are explored, including an expanded role for international financial institutions.

Beginning with Hotelling's seminal paper in 1931, a large body of literature in economics has focused on exhaustible resources and the special problems they pose for development. Between 1960 and 1990, resource-rich countries grew three to four times more slowly than resource-poor countries, and this gap in growth rates has widened considerably since the 1970s (Ranis 1991; Sachs and Warner 1995). Among the factors underlying this phenomenon is the impact of commodity price volatility on macroeconomic aggregates, particularly investment spending and the adjustment costs associated with rapid changes in expenditure (Ramey and Ramey 1995).

The problems of oil dependence are no less severe for oil importers, but they are dealt with less explicitly in this article. Among African oil-importing countries, in particular, unexpected increases in oil prices adversely affect government budgets and domestic currency and contribute to a deterioration in rural-urban terms of trade, particularly in countries that export primary agricultural commodities (Rutten 2001).

This article surveys country strategies in managing oil dependence, particularly the strategies of oil-exporting developing countries, rather than analyzing the problems associated with oil dependence, on which there is already a large body of literature.
The article presents a simple framework for understanding the impact of oil dependence on growth and surveys countries' use of stabilization and savings funds, fiscal rules, and hedging programs to respond to the problems of the volatility and exhaustibility of oil revenues. Although the country examples look at issues of managing oil price risk, many of the lessons apply to other minerals.

The first section briefly discusses an optimal savings and investment strategy for oil-exporting countries and the rationale for implementing savings and stabilization funds. The second section surveys country experiences with stabilization and savings funds. The third section explores the potential benefits of using financial instruments to manage oil price risk. The fourth section examines why so few developing economies use financial instruments to manage price risk and looks at some problems and mechanisms to mitigate the shortcomings of this approach. The final section presents some policy implications.

The Optimal Savings and Investment Policy

The main problems of oil dependence are associated with volatility in oil prices and exhaustibility of oil wealth. For a country where oil represents roughly 20 percent of gross domestic product (GDP), a one-standard-deviation shock to the price of oil represents an income shock equivalent to 6 percent of GDP (Hausmann and Rigobon 2003). Both volatility and exhaustibility can be incorporated into a simple framework for dealing with oil dependence in a developing country context.

Consider a hypothetical economy with substantial oil resources but lacking basic infrastructure, such as roads, schools, and telecommunications equipment. Suppose that the country has relatively benevolent leaders who would like to embark on a sustainable level of infrastructure investment funded by commodity revenues. In this context, sustainability implies that the expected present value of investment expenditures cannot exceed the expected present value of oil revenues. Because of borrowing constraints, investment must be funded from current oil revenues or from past oil revenues that have accumulated in a savings or stabilization fund. A savings fund is designed to create a store of wealth for future generations by converting a depletable revenue stream into a perpetual income flow. A stabilization fund is designed to stabilize revenue flows and, implicitly, expenditure.

An important component of this argument is that fluctuations in the rate of infrastructure investment reduce the efficiency of the investment program for several reasons. Building roads, schools, and hospitals requires a reasonably skilled workforce, and it is costly to train new workers when investment levels unexpectedly increase and to lay off workers when investment levels are unexpectedly cut. These training and lay-off costs are considered investment adjustment costs.
If commodity prices were certain and there were no political constraints, determining a country’s optimal level of infrastructure investment would be straightforward. Infrastructure investment should increase smoothly over time in a way that minimizes investment adjustment costs. In this case, savings and stabilization funds serve simply as a place to bank excess commodity revenues until they can be efficiently invested.

If a country’s revenues are generated from a commodity with a very volatile price, implementing an investment plan that increases slowly and smoothly over time becomes much more challenging. One of the roles of savings and stabilization funds is to allow the country to smooth out investment expenditures and thus increase the efficiency of investment by minimizing adjustment costs. As will be discussed, the extent to which savings and stabilization funds can smooth out investment depends on the random process generating the commodity prices.

An Example

Consider an economy that has a substantial amount of oil that can be extracted relatively quickly, say, over 10 years. Because of the costs associated with quickly increasing the level of investment, the economy should optimally invest its oil revenues over a time horizon that is substantially longer than the time it takes to extract the oil. Thus, the economy will need some sort of savings plan that allows it to bank the excess oil revenues, enabling it to continue to invest after the oil is depleted.

When oil prices are certain, the optimal investment and savings choice can be solved using dynamic programming. The optimal investment and savings levels at each date depend on current and future oil prices, current investment levels, marginal productivity of capital, and costs associated with increasing and decreasing the rate of investment. In general, the optimal solution involves an investment level that increases slowly until it reaches a steady-state level and then tapers off as oil revenues are exhausted.

Now consider the case of uncertain oil prices. Assuming that the oil price risk is not hedged, the optimal level of investment in each period will fluctuate as oil prices fluctuate. Intuitively, when the price of oil increases, the economy becomes richer, allowing a higher investment level. However, because investment fluctuations due to oil price fluctuations diminish the average efficiency of the investment program (because of the adjustment costs), the optimal rate of investment should fluctuate substantially less than the price of oil.

To understand this, consider an economy that is investing US$2 billion a year on infrastructure. Now suppose that because of an increase in oil prices, the country has $3 billion a year in revenues that can be invested. It would not make economic sense to immediately increase investment to this higher level because this would require that the country quickly hire and train a large number of construction workers.
and find supplies of building materials, a process that is not likely to result in the use of the best workers and suppliers. Moreover, if the price of oil subsequently falls to its previous level, infrastructure investments would need to be cut, forcing the country to lay off the new construction workers and perhaps bankrupt its new suppliers.

These arguments suggest that savings and stabilization funds can contribute substantially to economic development by dampening the effect of oil price volatility on investment volatility. The next subsection looks deeper into the economics of oil price uncertainty to provide insights about conditions in which savings and stabilization funds provide the greatest benefits and conditions in which they can only partially reduce investment volatility. Under these second conditions, hedging offers substantial benefits.

**The Random Process Generating Oil Prices**

The choice of a country's stabilization or savings plan depends on the temporary and permanent factors that influence the random process that determines oil prices (Deaton and Laroque 1992; Arrau and Claessens 2001; Cashin and McDermott 2002). Oil prices fluctuate from month to month because of temporary changes in global economic and political conditions that affect the supply and demand for oil. For example, political problems in Venezuela may temporarily disrupt oil supplies, causing prices to rise. Prices subsequently fall again as the problem is resolved or as other producers step up their production to offset the shortfall. Similarly, a recession may lead to a temporary oversupply that generates temporary price declines. Long-term or permanent factors can also affect oil prices. Permanent changes can arise because of longer-lasting changes in demand arising from the emergence of substitute fuels or stronger incentives for conservation as well as technological changes and new discoveries that increase the supply of oil.

The distinction between the temporary and permanent components of price changes is important for understanding the optimal investment and savings strategy in an economy with a savings or stabilization fund. When prices are strongly mean-reverting, as when price changes are largely temporary, the present value of future revenues is not very sensitive to changes in spot prices. This implies that the long-term sustainable level of investment is not likely to be substantially affected by a change in spot prices. Savings and stabilization funds serve a useful purpose in this case, because they allow a country to maintain an investment strategy with very little variation in an environment where revenues vary substantially.

When price changes are mainly permanent (less mean-reverting and resembling a random walk), the present value of future revenues is strongly affected by changes in spot prices. Savings and stabilization funds cannot be used to smooth out investment expenditures, and oil price volatility will significantly affect the efficiency of the country's infrastructure investment program.
A savings or stabilization fund can serve a useful purpose when commodity prices are strongly mean-reverting, but it cannot substantially smooth the effect of price changes over time when prices follow a random walk. In this second case, substantial gains are associated with reducing the variance of revenues by using financial instruments to deal with price risk.

The evidence on the statistical properties of oil prices tends to show more support for mean-reversion (Pindyck 1999; Barnett and Vivanco 2003) than for persistence (Cashin, Liang, and McDermott 2000; Engel and Valdes 2000), using unit root tests, although much depends on the time interval under scrutiny. This is underscored by the behavior of futures markets, which suggests that markets expect price shocks to be largely temporary. One way to evaluate the relative importance of the temporary and permanent components is to compare the fluctuations in futures and forward prices of various maturities with spot price changes. If price changes are largely temporary, long-term futures prices will have substantially less volatility than either short-term futures prices or spot prices. Evidence indicates that the long-term futures prices of oil are about half as volatile as the short-term futures prices (Barnett and Vivanco 2003; Schwartz and Smith 2000).

**How Effective Are Savings and Stabilization Funds in Theory?**

The effectiveness of a stabilization or savings fund depends on the relation between a country’s initial oil revenues and its investment needs and whether oil price changes are temporary or permanent.

Consider the case of an oil-exporting developing country that expects to generate enough revenue on average to fund its investment plans but may have a surplus or a shortfall in some years because of price volatility. This is a fairly typical scenario for small-population, high oil-resource countries, such as East Timor, Kuwait, and São Tomé and Principe. Suppose the country plans to invest $2 billion a year in health care and education and expects oil revenues to be sufficient to fund this investment. However, in years when oil prices are high, revenues greatly exceed $2 billion, and in years when oil prices are low, there is a shortfall.

When oil revenues greatly exceed the planned investment allocation and most of the price volatility is temporary, a savings or stabilization fund can eliminate most of the negative effects of fluctuating oil prices. Because most of the volatility in this case represents temporary changes, the present value of future oil revenues is not substantially affected by month-to-month changes in oil prices. Moreover, when oil revenues exceed investment plans, the surplus can be used to offset future shortfalls resulting from temporary price declines. However, when either of these conditions is violated, a savings or stabilization fund will not be sufficient to offset the potential negative effect on investment of fluctuations in oil prices. In this case, there will be gains to using market-based financial instruments.
Now consider the case in which planned investment is close to expected oil revenues and there is a smaller expected surplus from oil revenues. This is more typical in countries with large populations and moderate to declining oil revenues, such as Iran, Russia, and Venezuela. If prices are volatile and if the country saves very little, the savings fund may be fully drained before the first decline in prices.

Issues relating to the importance of temporary and permanent components of oil price volatility are more subtle and require additional explanation but can be easily understood through examination of the polar cases of those that are 100 percent temporary and those that are 100 percent permanent. When price changes are mean-reverting, the present value of all future oil revenues is not very sensitive to changing current oil prices, implying that the country’s long-term ability to meet its investment plans is not likely to be impaired as long as the savings or stabilization fund is sufficiently funded to meet the temporary shortfall. Changes in current oil prices should have little influence on the level of investment. In contrast, when oil prices follow a random walk, changes in spot prices are permanent, implying large changes in the present value of future oil revenues. The sustainable level of future investment will also change when oil prices change, so volatile prices are likely to result in substantial year-to-year changes in the level of investment, even with a savings or stabilization fund.

**Optimal Investment and Savings Plans When Oil Prices Are Uncertain**

The optimal funding of investment and savings funds thus depends on the following factors:

- The rate at which oil reserves can be extracted. If oil reserves are extracted quickly, the country should amass a substantial fund that allows it to continue to build up infrastructure when oil revenues start to decline.
- The costs associated with increasing and decreasing the level of investment. When there are substantial costs associated with altering the level of infrastructure investment, the benefits from smoothing investment are greater, implying a greater need for a fund that acts as a buffer when oil prices decline.
- The productivity of the investment. If the investment contributes substantially to productivity, the country should invest more initially and take the risks associated with subsequently having less of a buffer in the event of a decline in oil revenues.
- The random process that determines oil price changes. When oil prices are strongly mean-reverting, a relatively modest fund will be an adequate buffer against oil price declines. However, as the tendency to mean-revert declines, savings and stabilization funds will become less effective.
How Effective Are Stabilization and Savings Funds in Practice?

During the 1990s, the number of stabilization and savings funds proliferated, with the addition of Algeria, Azerbaijan, Ecuador, Iran, Kazakhstan, Mexico, Nigeria, Norway, Venezuela, and others based on the successes and failures of such pioneers as Alaska, Alberta, Kiribati, and Papua New Guinea in the 1960s and 1970s. Empirically, the effectiveness of funds appears to be limited, although data limitations have prevented rigorous analysis in this regard. Time-series analysis suggests that in countries with such funds, fiscal spending is less correlated with changes in the price of the resource, although it is difficult to separate the effects of the fund from the impact of prudent expenditure policies (Davis and others 2001).

More recent analysis using pooled cross-section and time-series data for 71 countries for 1970–2000 suggests that although funds appear to have a dampening effect on government spending as a percentage of GDP, this effect is offset by a deterioration in the fiscal balance as the size of the fund increases. However, implementing a fund appears to raise fixed capital investments as a share of GDP by nearly 3 percentage points, and there is a positive relationship between the balances held in the fund and fixed capital investment, suggesting that funds may have some impact on adjustment costs related to investment. In some countries, however—Chile, Norway, and Oman, for example—funds appear to deliver a number of favorable outcomes: less volatility in government spending, lower government spending, and higher shares of gross fixed capital investment. This suggests that country-specific circumstances matter a lot, in particular the use of fiscal rules and targets to guide spending decisions over a longer time horizon.

Funds and Fiscal Rules

Funds are clearly not a guarantor of fiscal stability. Many countries have erred by assuming that expenditures will automatically be stabilized and fiscal restraint encouraged by using a stringent accumulation rule to stabilize revenue flows. Accumulation rules have been price contingent (accumulation of revenues greater than a target price), as in the case of the Chile Copper Stabilization Fund; revenue contingent (a set percentage of oil revenues), as in the case of the Alaska Permanent Fund; and both (a set percentage of oil revenue above a reference price, as in the case of the Venezuela Stabilization Fund). The crucial link is the one between oil prices and fiscal expenditure (Devlin and Lewin 2002).

In countries where funds work well, there tend to be strong mechanisms that break the link between oil price behavior and fiscal expenditure, generally in the form of a fiscal rule. For example, Norway’s Norges Fund income is treated as central government net cash flow and transferred to the treasury to finance the nonoil deficit. Linking fund accumulation to fiscal surpluses avoids the problem of an overall deficit.
deterioration in the government’s net asset position as a result of fund accumulation. In the case of Chile, withdrawals from the fund are subject to the fiscal rule, with the structural balance calculated by factoring out the cyclical component of the copper price and other cyclical factors (Perry 2002; Fiess 2002).

A related issue is how to determine a benchmark price for forecasting fiscal needs. Many countries (such as Chile and Russia) have engaged in prolonged discussion and dispute over forecasting methods, a process that is vulnerable to political pressure. In Chile, the copper benchmark price was traditionally determined annually by a group of experts in a largely nontransparent fashion. An alternative is to use widely available and generally unbiased estimates provided by futures markets, such as the price of a five-year swap, for longer term fiscal planning.

There are two important caveats about the effectiveness of fiscal rules and operational guidelines in developing economies. First, in a perfect Barro world, private borrowing by domestic firms could potentially offset government savings with a fiscal rule. However, the empirical evidence is not conclusive, and most oil-exporting countries tend to be more affected by private capital flight than by private borrowing abroad. This suggests that expectations are more likely influenced by fundamental issues of institutional weaknesses characteristic of oil-exporting economies, particularly with respect to property rights and the functioning of the judicial system.

A second caveat is that informal norms and practices tend to have a greater effect on fiscal behavior in developing economies than formal rules and medium-term expenditure frameworks (Schick 1998), especially in oil-exporting countries where the “rentier” nature of oil revenues tends to weaken formal mechanisms of revenue accountability (Eifert, Gelb, and Tallroth 2003). Thus, in weak fiscal environments, the effectiveness of funds may also be influenced by the level of financial and income incentives in the government overall.

**Funds and Managing Windfalls**

Prudent use of windfalls requires appropriate governance structures based on transparency and accountability, and funds can provide an opportunity for developing economies to import better governance mechanisms to deal with windfalls. In Botswana during the 1970s and 1980s, for example, recurrent and investment spending were based on estimates of long-run diamond revenues and during windfalls reserves were accumulated in the Central Bank. This provided leeway for using windfall revenues to target investments in areas of major bottlenecks to economic development by the second half of the 1980s (Hill 1991).

Although there are no guarantees against the possibility that by allowing better saving of windfalls, funds will transfer resources from good governments to bad (Collier 2002), the experience of relatively effective funds suggests that a system of checks
and balances both internal and external to the fund can provide a mechanism for “padlocking” the fund’s resources. Ideally, funds should have vertical accountability to an oversight board of representatives of the central bank or ministry of finance and horizontal accountability to agencies in the state and outside observers, such as the media and civil society, particularly environmental groups and traditional exporters (Karl 2000; Bates 1997). Specifically, there must be built-in mechanisms for control, reporting, and evaluation of fund resources and operations (Heilbrunn 2002). In addition, funds should be professionally managed, with oversight by the ministry of finance or central bank. In Norway, the Ministry of Finance supervises the activities of the fund and sets guidelines for investments and reporting requirements. Transfers to and from the fund require parliamentary approval.

**Transparent Design of Funds**

The institutional design of the funds can also make the earnings and use of resource revenues more visible (Bjerkholt 2002). In the case of the Alaska Permanent Fund, for example, which uses invested oil revenues to distribute annual dividends to all Alaskan citizens, the population can check the monthly earnings and expenditures of the fund to determine the exact amount of dividend checks (table 1). One proposal along these lines is to deposit surplus commodity revenues into pension accounts to encourage more public scrutiny of fund resources (Hannesson 2001).

Information on the fund’s activities should be publicly available and widely disseminated. Where private oil companies produce most of the oil and natural gas, public disclosure of tax records and revenue transfers to the government is another alternative for improving the transparency of commodity revenues. Electronic tracking of oil market sales worldwide may help increase the transparency of these revenue flows.

**Optimal Size of Funds**

Another lesson is that the size of the fund matters. For economic and political reasons, larger funds tend to be more inefficient and to create more distortions. The optimal size of a fund tends to be much smaller than expected, with the determining factor being the statistical properties of oil prices rather than rules for accumulation and withdrawal of funds (Arrau and Claessens 2001; Crain and Devlin 2002). Larger funds (or fiscal surpluses) are also more vulnerable to political economy pressures—as evidenced by the Indonesian government’s decision during the 1970s to hide fiscal surpluses.
Table 1. Features of Selected Savings and Stabilization Funds

<table>
<thead>
<tr>
<th>Fund and date founded</th>
<th>Rules</th>
<th>Governance</th>
<th>Asset management</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alaska Permanent Fund</td>
<td>• Accumulates 50 percent of mineral revenues.</td>
<td>• Principal protected by constitution.</td>
<td>• Alaska Permanent Fund Corporation, independent from state.</td>
</tr>
<tr>
<td>1976</td>
<td>• Uses earnings for dividends, inflation proofing, adding to principal.</td>
<td>• Governed by board of six trustees—members of public, commissioner of revenue, and one governor’s appointee.</td>
<td>• Investments largely out of state.</td>
</tr>
<tr>
<td></td>
<td>• Calculates dividends as net income of fund for last five years, times 21 percent, divided by 50 percent.</td>
<td>• Executive director employed to manage the fund.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Pays dividends to each Alaskan resident.</td>
<td>• Expenditures limited to 1.2 percent of net income.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Links oil income and spending through the dividend program, since there is no state income tax.</td>
<td>• Regular reporting and auditing, with all information publicly available.</td>
<td></td>
</tr>
<tr>
<td>Norway</td>
<td>• Fully integrated into budget process under the control of the treasury.</td>
<td>• Parliamentary approval for transfers to and from the fund.</td>
<td>• Ministry of Finance sets guidelines for asset strategy.</td>
</tr>
<tr>
<td>1990, activated 1995</td>
<td>• Finances nonoil deficit.</td>
<td>• Regular reporting to Parliament on fund status.</td>
<td>• Managed by Norges Bank.</td>
</tr>
<tr>
<td></td>
<td>• Receives fund earnings, nonoil surplus.</td>
<td>• Supervision by Ministry of Finance.</td>
<td>• Parliamentary approval for changes in guidelines.</td>
</tr>
<tr>
<td></td>
<td>• Higher government spending or lower taxes result in smaller allocations to the fund.</td>
<td>• Quarterly and annual reporting.</td>
<td>• Investment in overseas assets.</td>
</tr>
<tr>
<td></td>
<td>• Operations incorporated into fiscal accounts.</td>
<td>• Regular, public audits.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Legal requirement to provide information on fund management to public.</td>
<td>• Managed by Ministry of Finance</td>
<td></td>
</tr>
<tr>
<td>Oman</td>
<td>• Accumulating when price is above State General Reserve Fund reference price ($15 a barrel).</td>
<td>• Supervised by the Financial Affairs and Energy Research Council.</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
The Case for Using Market-Based Instruments to Manage Oil Price Risk

When oil revenues lag behind initial investment expenditures and prices are mean-reverting, there may be strong efficiency gains from using financial instruments to manage oil price risk. The use of market-based financial instruments has long been proposed as the first-best solution for dealing with oil price volatility, but risk management programs are rarely implemented (Engel and Mellor 1993; Larson and Varangis 1996; Claessens and Duncan 1993; World Bank 1999).

Benefits of Using Financial Market Instruments

Instruments such as swaps, futures, and options make it possible to lock in a known price for a given period, thus eliminating price uncertainty, and they allow a country to transfer commodity price risk to the market, rather than self-insuring. A risk management strategy can also make credible the promise not to spend windfall income (Hausmann 1999). To mitigate the effects of permanent price components, small savings or stabilization funds can be combined with financial instruments to avoid paying risk premiums or depleting the fund (Larson and Varangis 1996; Shimko 1995; Humphreys 2000).

Instruments for managing commodity price risk are becoming more sophisticated with the development of derivatives and are being traded on capital markets.
For example, futures and options for crude oil and natural gas are traded on the New York Mercantile Exchange, whereas forward and swap contracts are traded on over-the-counter markets.

Exchange-traded instruments are usually more transparent, easier to monitor, and more liquid than over-the-counter instruments, such as swaps and forwards. Options are one way to protect against volatility in oil prices. With a put option, a government buys the right to sell a certain quantity of oil at a specific price for a specific period, effectively establishing a price floor. If prices fall below this floor, the value of the strike price increases and the option can be exercised or sold for a profit. To help offset the cost of the premium on a put option, the government can sell call options or the right to buy oil at a specific price over a given period, effectively establishing a price ceiling. Various combinations of puts and calls enable a government to benefit from the upside potential as well—for example, with a “costless collar,” a government can sell a call option and use the premiums to purchase a put option. The state of Texas is using this strategy to hedge the 25–30 percent of state revenues tied to oil revenues.

An adequate assessment of basis risk—the difference between the benchmark price used to hedge and the realized price in the exporting country—is important for using futures and options. Although in practice the basis risk for oil tends to be much smaller than the overall price risk, the benefits and costs of using risk management instruments are clearly affected by basis risk. For Alaskan oil, for example, prices have ranged from $9 to $41, whereas the basis difference has ranged from $0.80 to $4.20 (Lindahl 1996).

Over-the-counter instruments, such as swaps and forward contracts, are more flexible than exchange-traded instruments because they are customized transactions between the government and the financial intermediary. They can mitigate basis risk, are available in large volumes for single transactions, and frequently cover longer periods. Initial deposits and margin calls are also avoided. In a swap transaction, the oil-producing government contracts with a private bank to lock in a price (say, $20 a barrel) for a long period, say, two to three years. The government then sells the oil on the open market, and both parties calculate an average sales price every six months. If the average price that the country receives is less than $20 a barrel, the bank pays the difference. If oil prices rise, the payment flows in the reverse direction. There are several variations on this approach. A country could also use a swap to establish a price floor and could sell a cap (say, $25 a barrel) to help cover the cost of the floor. The country would pay the bank any oil revenues received over the $25-a-barrel cap.

A forward contract is like a futures contract except that it is an agreement between a buyer and a seller directly (not on an exchange), and forward contracts generally give rise to physical deliveries.

Using both exchange-traded commodities and over-the-counter instruments, oil-producing governments can sell their production forward or buy insurance against
large price declines. A good time to start risk management is when prices are high relative to trend. During a high price period, a producing country can lock in revenues at the high price through swaps and protect against low price scenarios with put options.

**Global Externalities from Using Financial Markets to Manage Oil Price Risk**

There are likely to be favorable externalities for the global economy from greater use of financial markets to manage oil price risk. When prices fall, a constrained producer may want to increase (rather than reduce, as microeconomic theory would predict) its production to generate sufficient revenues to meet its investment needs. Thus, because of the tendency of some producers to increase production and contribute to an oil glut, what would otherwise be a temporary price decline can be both amplified and extended.

Although consumers benefit from the decline in oil prices, the increased volatility created by this perverse incentive to increase production as prices fall creates substantial inefficiencies in consuming countries as well as in producing countries. When oil prices are very volatile, consuming countries are likely to find it difficult to sustain fuel efficiency and conservation efforts, whereas producing countries are likely to experience greater political instability.

When a country is able to hedge a large portion of its oil revenues, it no longer tends to increase production as prices decline. Indeed, the country will be able to benefit by timing its oil production to coincide with higher oil prices. Oil producers will generate higher revenues on average, and oil prices will be less volatile.

**Why Few Developing Country Governments Use Financial Instruments to Manage Oil Price Risk**

Despite the expansion of markets to absorb commodity risk, there is little evidence that developing economies are using market-based instruments to manage oil price risk. This appears to result mainly from two causes: the status quo problem and the lack of coordination in international financial markets.

**The Status Quo Problem and the State of Texas Hedging Program**

Although the use of financial instruments reduces risk and increases efficiency, there may be some risk for government officials who implement such a policy. Policymakers are likely to be evaluated by how well market-based instruments work relative to the status quo with no use of hedging.

Officials must be concerned not only with the volatility of prices but also with the possibility that the country will end up worse off as a result of the risk management
decisions. Although policymakers will benefit politically from the use of risk management instruments if commodity prices decline, the political costs may outweigh the benefits if prices rise. Because there is close to a 50 percent chance that any given risk management decision will lose money relative to the status quo of no risk management, it is not surprising that government officials do not generally consider risk management a viable alternative.

Finally, the effectiveness of risk management strategies needs to be evaluated from the vantage point of fiscal stability, an important area for further research. Recent evidence suggestions that when market-based instruments are not used, government will realize the highest portion of expected revenue but also will experience the highest level of revenue volatility and more frequent budget deficits. Exchange-based risk management can effectively manage extreme downside risk and stop large budget deficits from occurring (Buttimer, Shaw, and Swidler 1999).

Many of these issues came into play in the oil price risk management program of the state of Texas, one of the few examples of using market-based instruments to hedge a state budget. Texas collects about half a billion dollars from a 4.6 percent production tax on crude oil, and in 1991 the Texas Senate introduced a bill authorizing a two-year pilot risk management program, with expanded legislation passed in 1993. To avoid opposition, the program was initiated in a quiet, unobtrusive way as part of existing treasury operations with program support financed by unclaimed royalties. Political support for the risk management program was created and maintained through continuing efforts to educate legislators and their staffs and to address fears about risk management operations. In addition, the program had well-defined safety mechanisms, including board oversight, use of exchange-traded options, monthly and quarterly reporting, and a system of internal and external checks and balances. The State Depository Board, which had a strong reputation for technical competence, was selected to supervise the program. Guidelines and operating procedures, drawn up in consultation with an advisory committee, covered authorization to trade, decision-making responsibilities, position limitations, broker constraints, daily monitoring, separation of responsibilities, and internal reporting to both fiscal authorities and the oversight board.

In addition, a system of checks and balances was established to monitor the activities of the risk management program, incorporating treasury supervision, oversight board supervision, Risk Management Group trading, banking functions, accounting functions, broker and clearing agents, and external audit. Treasury staff familiar with exchange and fiscal spending transactions were included in the auditing of the risk management program—staff checked the daily activities of the program and wired the money to be used in purchasing options on behalf of the Risk Management Group. Other controls included stop-loss limits on the amount of money that could be placed at risk in one day ($500,000) and a total amount that could be lost throughout the program ($2.5 million—the program would stop immediately if that
limit was reached). Limiting risk management instruments to exchange-traded options (no futures or swaps) offered additional security because no open-ended losses could be incurred (as they could be with futures), options are more liquid and easier to understand, and trades would be subject to greater scrutiny in a well-regulated market. At the end of two years, the program would be assessed and either closed down or continued.

This experience shows that to implement an effective risk management program, governments need to identify a clear objective (such as risk management) to deal with fears of speculation. Ideally, risk management programs would be the responsibility of existing treasury operations, not a new unit, and would rely on existing staff with high levels of technical expertise in commodity markets. Clear operational guidelines should be developed, including authorization to trade, decision-making responsibilities, position (stop-loss) limitations, broker constraints, daily monitoring, separation of responsibilities, and internal reporting. Alternative risk management strategies (options, swaps, futures) should be evaluated on paper first, and the selected strategy should be introduced as a pilot program with a sunset provision if the program loses too much money. The program should have clear opportunities for upside potential.

Coordination Failures in International Financial Markets

Low use of financial instruments by developing country governments also stems from a coordination failure in international financial markets. Suppose that major oil corporations and major sovereign producers simultaneously decide to hedge their oil price exposure. This would of course solve the status quo problem. But there is a second problem that would need to be overcome before these risk management strategies can be implemented. Selling the equivalent of billions of barrels of oil in financial markets requires buyers who are willing to purchase these amounts.

Who are such potential counter-parties? In general, they come from two groups. The first group consists of institutions and individuals with an incentive to take the other side of these transactions to hedge their own risks. For example, power producers with gas-fired generators have an incentive to lock in gas prices to hedge the costs of generating power. Similarly, large purchasers of oil, such as airlines, plastics, and chemical companies, can serve as natural counter-parties for oil producers. The amount of oil such companies consume is likely to be relatively small, however. Most oil is consumed by individuals, in the form of gasoline and heating oil. But few individuals find it practical to buy oil futures and forward contracts to hedge future expenses.

The second major counter-party group is made up of large pension funds, individual investors, and hedge funds. Because individuals in developed economies, who are the largest oil consumers, are unlikely to hedge, there is likely to be a substantial
gap between the amount of oil that producers should in theory sell forward and the amount of oil that consumers can be expected to buy forward. That means that there should be a positive risk premium associated with holding oil price risk that should make such investments attractive to institutional investors. But most pension funds and other institutional investors lack the expertise to invest in these commodities, and learning that skill will likely take many years. Although there are hedge funds that can take the other side in these transactions, these funds tend to be thinly capitalized and can likely take on only a small part of the necessary exposure.

With the amount of capital that can plausibly be used to take the other side in risk management transactions so small relative to the aggregate oil price risk, prices on futures and forward markets would decline substantially if producers tried to hedge a significant amount of the oil price risk. Markets are liquid enough to absorb small trades by small producers but not yet deep enough to allow large producers to hedge as much as they would like. Anecdotal evidence suggests that large oil companies tend to hedge only the gap between current production levels and the minimum amounts needed to keep facilities going.

A Role For International Financial Institutions

There are likely to be substantial benefits associated with developing markets that allow oil-producing governments to hedge their oil price exposure. As a result, there is a role for international financial institutions to address the status quo and coordination problems just outlined.

With respect to the status quo problem, the key to developing deeper markets for managing oil price risk is to educate government officials on the use of market-based instruments to deal with oil price risk. International financial institutions can also encourage the involvement of investment banks to develop the counter-party side of the financial market for oil price exposure.

Ultimately, if oil producers hedge a significant amount of their oil price risk, the largest fraction of the risk will need to be held by international institutional investors, because relatively little oil price exposure will be taken by petroleum users. However, this will require the development of securities that are tailored to the needs of international institutional investors, in addition to the removal of various regulatory and legal impediments.

This points to a potential third role for international financial institutions in facilitating securitization of oil resources. One possibility would be to develop an international agency, resembling Fannie Mae or Freddie Mac, to act as a conduit for banks that initiate loans with interest payments tied to oil prices. These loans could then be packaged, securitized, and divided into tranches to appeal to various investor clienteles. For example, a senior AAA tranche could be sold to pension funds and insurance...
companies. A BBB tranche could be sold to more adventurous institutions, like university endowments, that would like to have some oil exposure but are prohibited in their charters from holding derivatives. Hedge funds and other more specialized investors could hold the riskier junior tranches.

International financial institutions could play a role by helping originate the transaction and working to ensure that the developing country gets the best possible terms for financing. In addition, international financial institutions can provide guarantees and insurance against the risk of default, analogous to the role that Freddie Mac played in developing a market for mortgage-backed securities. In the initial development of the residential mortgage-backed securities market, Freddie Mac acted as a conduit between the savings and loan institutions that initiated the mortgages and the investment banks that offered the securities. As a conduit, an international financial institution might need a few billion dollars to initially hold some of the loans while they are being securitized and might want to set aside funds to insure the loans.\(^3\)

Thus, international financial institutions may be able to facilitate the development of a market for securitized oil-linked loans by dealing with legal and regulatory issues related to the introduction of new securities, addressing concerns about political risk, and marketing these securities to countries that import a substantial amount of oil to help them hedge oil consumption risk.

There are likely to be favorable externalities for the global economy as a result of the greater use of financial markets for managing oil price risk. When prices decline, a constrained producer may want to increase its production (rather than reduce it, as microeconomic theory would predict) to generate sufficient revenues to meet its investment needs. As a result, what would otherwise be a temporary price decline can be both amplified and prolonged because of the tendency of some producers to increase production and contribute to an oil glut.

Although oil consumers benefit from the decline in oil prices in these situations, the excess volatility resulting from this perverse incentive to increase production as prices fall creates substantial inefficiencies in consuming countries as well as in producers. For example, it is likely to be more difficult to maintain fuel efficiency efforts and to increase conservation efforts when oil prices are very volatile. In addition, oil price volatility is likely to contribute to political instability in oil-producing countries.

When a country hedges a significant portion of its oil revenues, its tendency to increase production as prices decline can be eliminated. Indeed, the country will be able to benefit by timing its oil production to coincide with higher oil prices. By doing this, the oil producers will generate higher revenues on average, and oil prices will be less volatile.

Finally, international financial institutions can work more closely with major international oil companies in sharing oil price risk with developing country governments. Major oil companies are already exposed to substantial amounts of oil price
risk. The largest international oil companies—British Petroleum, Exxon Mobil, Royal Dutch Shell, and Total Fina Elf—produce about eight to nine million barrels of oil a day. This amount exceeds the total production of non-OPEC South American and African producers—in other words, the production from developing economies that would benefit the most from hedging. What this means is that the major oil companies can take a significant amount of the oil price exposure of the smaller producers without substantially increasing their own exposure. The oil companies can do this either by buying the securities described or by structuring their contracts with the exporting countries in ways that shift the price risk to them.

Conclusion

To deal with oil price volatility and exhaustibility, countries have set up stabilization and savings funds. Although the funds have helped manage windfalls and turn depletable wealth into productive assets, their performance has been much weaker in reducing the effect of volatile oil prices on government revenues and spending. Especially important have been a sound fiscal framework and accountability and transparency in fund structure.

Two limiting influences on the effectiveness of oil funds are oil price processes, which are almost entirely beyond the control of policymakers, and political economy factors. Both tend to put downward pressure on the optimal size of funds and raise the efficiency gains from the use of market-based financial instruments.

In theory, the first-best strategy to deal with commodity price volatility is the use of market-based risk instruments. In practice, however, policymakers in developing areas are deterred by the status quo problem and by coordination failures in the development of international financial markets. On the status quo problem, Texas's experience suggests that government hedging of oil price risk is challenging but not impossible when the program has the clear objective of insurance and is introduced gradually, preferably as a pilot program by extending existing responsibilities for treasury operations rather than by creating a new, stand-alone program. Clear operational guidelines for the program, with appropriate layers of authorization for trades, automatic stop-loss provisions, separation of responsibilities, and internal reporting are critical to avoid speculation.

There is also a potential role for international financial institutions in promoting the use of market-based instruments to manage oil price risk, given coordination failures in international financial markets. Education, publicity, and technical assistance in working with oil-producing and -importing governments to manage oil price risk are possibilities. International financial institutions can also work to develop instruments to facilitate securitization of oil proceeds and can encourage oil companies to work with developing economies in sharing oil price risk. There are
likely to be significant externalities for the global economy associated with this role—namely, reduced volatility in oil prices.

Notes

Julie Devlin is senior private sector development specialist at the World Bank; her e-mail address is jdevlin@worldbank.org. Sheridan Titman is professor of finance, McAllister Centennial Chair in Financial Services, University of Texas at the University of Texas, Austin; his e-mail address is titman@mail.utexas.edu.

1. In Chile and Norway, commodity revenues as a percentage of overall government revenues are significantly lower than in most oil-exporting countries, implying that the impact of a volatile revenue stream is significantly lower. Copper prices are also less volatile than oil prices, lessening the degree of potential fiscal volatility. These effects were taken into account in Crain and Devlin (2002).

2. This discussion is based on Patterson (2001).

3. However, the proposed hedging strategy need not put very large amounts of much money at risk. First, it is assumed that the oil-producing country would put significant amounts of money on deposit at the banks to serve as partial collateral for the volumetric loans. Second, the loans would also be collateralized by oil, which presumably can be seized if the defaulting country tries to export the oil. It should be emphasized that the oil exporter will have an incentive to renege on the loan agreement only when oil prices are substantially higher than the contracted price. However, if the oil-importing countries agree to recognize these contracts, the exporting countries will not be able to renege and sell their oil at the higher market price. Of course, there is still political risk associated with political events that result in a disruption in the country’s ability to produce oil.

References


Does your library subscribe to this journal?

If not, why not use the online Library Recommendation Form for OUP Journals. Go to www.oupjournals.org/library, fill in your details, press 'submit', and the form will automatically be forwarded to your librarian.

LIBRARY RECOMMENDATION FORM

Dear Library's Acquisition Committee,

I would like to recommend a subscription to

[Leave a brief note here]

Name: .........................................................
Position: ....................................................
Department: ................................................
Signature: ...................................................

If you prefer, contact us:

OXFORD UNIVERSITY PRESS

Telephone: +44 1865 353000
Fax: +44 1865 353485
Email: journalsmarketing@oup.com

OXFORD, OXFORDSHIRE, OX2 6DP, UK