30230

Targeting of Transfers in Developing Countries

Review of Lessons and Experience

DAVID COADY

MARGARET GROSH

JOHN HODDINOTT

Targeting of Transfers in Developing Countries

Targeting of Transfers in Developing Countries: Review of Lessons and Experience

David Coady
Margaret Grosh
John Hoddinott

THE WORLD BANK
Washington, D.C.

IFPRI®

Contents

Contents of CD-ROM
   Book text in English
   Book text in Russian
   Book text in Spanish
   *Annotated bibliography with information on the 122 programs in the
   database*

Foreword

The World Bank and the International Food Policy Research Institute (IFPRI) work closely with partners in developing countries in designing, monitoring, and evaluating interventions to reduce poverty and help households to manage risks. A major issue associated with such interventions is the implementation of mechanisms to ensure that benefits flow to the poorest and most vulnerable individuals. While much has been written on targeting antipoverty interventions, this writing has often focused on case studies of individual programs that, while interesting, are often too specific to be of much value to developing country governments.

This book, the result of joint work by staff at the World Bank and IFPRI, seeks to fill this information gap. It assesses more than 100 case studies undertaken by the Bank, IFPRI, and many others involved in development work on the design, implementation, and effectiveness of methods designed to focus benefits on the poor. While the central message is clear—there are no magic bullets—the book goes beyond this finding and provides clear advice as to which approaches are most likely to succeed in different circumstances.

IFPRI and the World Bank share a common goal: a world free of hunger and poverty. We hope this book will assist others who share our goal by showing how careful design and implementation can maximize the benefits of interventions for the poor and vulnerable.

Robert Holzmann, Director of Social Protection, World Bank

Joachim von Braun, Director General,
International Food Policy Research Institute

Acknowledgments

Acronyms and Abbreviations

GDP	gross domestic product
GAPVU	Gabinete de Apoio à População Vulneravel
INAS	National Institute for Social Action
INEGI	National Institute of Statistics, Geography and Information
NGOs	nongovernmental organizations
OECD	Organisation for Economic Co-operation and Development
PIOJ	Planning Institute of Jamaica
PPP	purchasing power parity
PRAF	Family Allowance Program
PROGRESA	Programa de Educación, Salud y Alimentación
SISBEN	Sistema de Identificación y Clasificación de Potenciales Beneficiarios para los Programas Sociales
STATIN	Statistical Institute of Jamaica
VAT	value-added tax

1

Introduction

A consensus has emerged in the past two decades that economic growth is a necessary but insufficient condition for the alleviation of poverty.[1] Additional elements are required. First, the asset base of poor households needs to be built up so that they can participate in the growth process. Second, growth needs to be more intensive in the assets held by the poor and the sectors in which they predominate. Third, because it takes time for the benefits from such a strategy to accrue, short-term public transfers are required to protect and raise the consumption of the poorest households.

Implementation of this agenda for reducing poverty requires methods for reaching the poor. In part, this can be accomplished by spending on items such as universal primary education (van de Walle 1998) that reach a wide swath of society, including the poor. It also can be accomplished by providing resources directly to the poor. However, scarce government resources have encouraged efforts to concentrate resources on "target groups" of poor households or individuals. This will achieve the maximum impact from a given poverty-alleviation budget or achieve a given impact at the least budgetary cost. The attraction holds for many kinds of poverty reduction programs and expenditures, but perhaps most strongly for the transfer programs that constitute safety nets because these transfers confer a benefit that is largely a private good for the recipient household.

Although targeting has obvious benefits, numerous methods exist for directing resources to a particular group. The salient literature, while extensive, is largely dominated by descriptions of individual, sometimes idiosyncratic, programs. Even comparative analyses tend to cover a single region (e.g., Grosh 1994 for Latin America and the Caribbean; Braithwaite, Grootaert, and Milanovic 2000 for Eastern Europe and Central Asia) or 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.

Accordingly, this book provides a general review of experiences and lessons learned with methods used to target interventions in developing countries. Our primary audience is composed of policy makers and program managers in developing countries, in donor agencies, and in nongovernmental organizations (NGOs) who have the responsibility for designing interventions that reach the poor. Our objective is to convey available targeting options, anticipated results, and relevant information to assist in optimizing the implementation of the chosen option.

We stress that targeting is a means toward the end of poverty reduction. Assessing the effectiveness of targeting is an exercise in assessing one component of antipoverty interventions. It should not be confused with an assessment of *all* impacts of targeted interventions on welfare, a review that is beyond the scope of this book.[2] Programs may have other objectives than transferring money to the poorest households, and these objectives may involve a tradeoff with targeting performance. For this reason, policy makers who are thinking about intervention choices must consider the whole set of strengths and weaknesses of programs as they make their decisions.

Mindful of these caveats, we seek to convey five core messages about targeting effectiveness.

- *Targeting can work* . . . Across all programs for which we could obtain information on targeting performance, we find that the median program provides approximately 25 percent more resources to the poor than would random allocations. The best programs were able to concentrate a high level of resources on poor individuals and households. Argentina's *Trabajar* public works program, the best program in this regard, was able to transfer 80 percent of program benefits to the poorest quintile. The best 10 performers deliver to the poor two to four times the share of benefits that they would get with random allocations. Progressive allocations were possible in all country settings, in countries at markedly different income levels, and in most types of programs.
- *. . . but it doesn't always*. The state of the art as practiced around the world is highly variable. While median performance was good, in approximately 25 percent of cases targeting was regressive so that a random allocation of resources would have provided a greater share of benefits to the poor. For every method considered, excepting targeting based on a work requirement, there was at least one example of a regressive program.
- *There is no clearly preferred method for all types of programs or all country contexts*. In our sample of programs, 80 percent of the variability in targeting performance was due to differences *within* targeting methods and only 20 percent was due to differences *across* methods.
- *A weak ranking of outcomes achieved by different mechanisms was possible*. 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 relative

to targeting that uses self-selection based on consumption. Proxy means testing, community-based selection of individuals, and demographic targeting to children show good results on average, but with considerable variation. Demographic targeting to the elderly, community bidding, and self-selection based on consumption show limited potential for good targeting. This ranking cannot be taken as a blanket preference for one method over another. It does not consider cost and feasibility constraints. Furthermore, our regression results should be considered as showing correlations rather than causal relations because targeting methods are themselves choices.

- *Implementation matters tremendously to outcomes.* Some of the variability was explainable by country context. Targeting performance improved with country income levels (the proxy for implementation capacity), the extent to which governments are held accountable for their actions, and the degree of inequality. Generally, using more targeting methods produced better targeting. Unobserved factors, however, explained much of the differences in targeting success. Significant potential remains for improvements in the design and implementation of targeting methods. If programs with poor targeting success were brought up to median, the mean performance indicator would rise from 1.38 to 1.55.

Chapter two provides a brief review of targeting, discussing the benefits and costs of targeting, methods for assessing targeting performance, and a taxonomy of targeting methods. Readers familiar with these issues can go directly to chapter three, where we analyze quantitative evidence on targeting outcomes derived from an extensive review of existing studies. Chapter four complements this quantitative analysis with a qualitative treatment of common targeting methods. For each, we review international experience, how the method works, what determines how well it works, what its costs are likely to be, and appropriate circumstances for its use. Chapter five offers summaries and conclusions.

Notes

1. See *World Development Reports* for 1990, 1997, and 2000 for more detailed discussion on this and related issues.

2. We note that assessment of impact requires careful attention to the counterfactual, that is, what beneficiaries would have done in the absence of these interventions. Few studies of welfare programs in developing countries do so with any care, with some exceptions: Datt and Ravallion (1994), Ravallion and Datt (1995), Jalan and Ravallion (1999), and Skoufias (2001).

Targeting: An Overview

Managers have many methods available to target an antipoverty intervention. In developing an understanding of what methods are appropriate under what circumstances, it is helpful to begin by enumerating the benefits and costs of targeting.[1] Decisions about whether to target, how precise to be, and what method to use will depend on the relative size of these costs and benefits, which will vary by setting. An assessment of these benefits and costs requires the measurement of targeting performance, which is the third topic taken up here. Lastly, we outline a structure for classifying targeting methods.

The Benefits of Targeting

Targeting is a means of increasing program efficiency by increasing the benefit that the poor can get within a fixed program budget. The case for targeting is tantalizingly simple. Imagine an economy with 100 million people, 30 million of whom are poor. The budget for a transfer program is $300 million. With no targeting, the program could give everyone in the population $3. If the program could be targeted only to the poor, it could give each poor person $10 and spend the full budget, or it could continue to give each poor person $3 for a budget of only $90 million.

More generally, the motivation for targeting arises from the following three features of the policy environment:

- *Objective:* the desire to maximize the reduction in poverty or, more generally, the increase in social welfare
- *Budget constraint:* a limited poverty alleviation budget
- *Opportunity cost:* the tradeoff between the number of beneficiaries covered by the intervention and the level of transfers.

These three features imply that targeting transfers at poor households has a potential return, namely, that the amount of the transfer budget going to those households deemed to be most in need of transfers can be increased.

Figure 2.1.   Targeting Poverty Alleviation Transfer

Final Income

Original Income
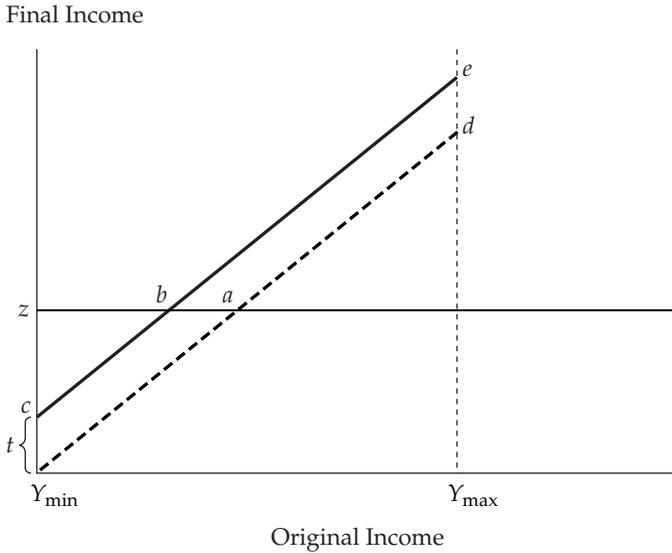
This concept can be expressed graphically (figure 2.1). As a policy maker, suppose we have a fixed transfer budget just sufficient to eliminate consumption poverty.[2] We have representative household survey data and, using this, we graph consumption levels of individual households before any transfers to them, ordering them from worst to best off. This ordering is represented on the *x*-axis as "original income," while a household's income after the transfer is given on the *y*-axis as "final income." The maximum and minimum household incomes in the survey are y_{max} and y_{min}, respectively, and z is the poverty line. The line dy_{min} shows that, by definition, before the transfer program is in place households' final incomes are equal to their original incomes. The optimal transfer scheme is one that gives a transfer to all poor households only (i.e., those with income less than z), with transfer levels equal to their individual "poverty gaps," that is, the distance between their original income and the poverty line, za. This transfer program brings all poor households up to the poverty line; all nonpoor households have equal final and original incomes. The poverty budget is represented by the area zay_{min} and is the minimum budget required to eliminate poverty.

Consider the case of a uniform transfer program, which gives the same transfer equal to $t\ (= c - y_{min})$ to all households, both poor and nonpoor. Because of the leakage of transfers to nonpoor households, the transfers to poor households are no longer sufficient to eliminate their poverty. Two forms of "inefficiency" are associated with the uniform transfer:

- Nonpoor households receive a transfer.
- Some poor households (those in the line interval ba) receive transfers greater than their poverty gaps.

As a result of these inefficiencies, the poverty impact of the uniform transfer scheme is less than that of the optimal transfer scheme, less by the area *zcb*. The total leakage of the budget (reflecting the two sources of inefficiency identified above) is given by the area *bade*, which for a fixed budget must also equal the area *zcb*, which equals the level of poverty after the uniform transfer program. Therefore, imperfect targeting results in a lower poverty impact for a given budget. Improved targeting involves screening some of the nonpoor households out of the program.

The Costs of Targeting

The scenario outlined above illustrating the benefits of targeting assumed that it was possible to distinguish who is poor and who is not. In fact, there are costs to acquiring information about who is needy and, even then, such information is rarely perfect. These costs can be classified as follows:

- *Administrative Costs:* These costs include the costs of collecting information, for example, conducting means testing of households or conducting a survey on which to base a poverty map. These costs mean that less of the budget is available to be distributed to beneficiaries. In general we expect that the costs of gathering information to target will increase with the precision of the targeting.

  It is possible that if finer targeting means that the total number of beneficiaries declines, the total administrative costs will decline, either absolutely or a share of total costs. This would result from two forces. First, a targeted program may serve a smaller number of people, so the overall scope of machinery to deliver benefits could be smaller. Second, if the tighter targeting allows a larger benefit per client, the share of administrative costs will be lower. Imagine a program that costs $1 per household to gather information about targeting and $5 per household for the administrative costs of delivering the benefit worth $100. If the program serves 1 million client households, then the total administrative cost would be $6 million, the total cost $106 million, and the share of administrative costs about 6 percent. Next imagine moving to much finer targeting, for example, from demographic targeting to a means test. The cost of gathering information for targeting might rise to $5 per household. The cost of getting the benefit into the client' s hands remains $5. However, now the program serves only 250,000 families, so administrative costs are $2.5 million. If the benefit is kept at $100 per family, then the total budget will be $27.5 million and the share of administrative costs about 10 percent. If some of the resources freed through the finer targeting are used to raise the benefit to $200 per family, then the total cost would be $52.5 million and the share of administrative costs would be about 5 percent, lower in both absolute terms and as a share of the total program budget.

It is important to note, however, that from the perspective of targeting the relationship between the level of costs incurred because of the decision to target transfers to the poor and the improved targeting performance resulting from these extra costs is of particular interest. While from this perspective it is always desirable to reduce the level of nontargeting-related program administrative costs, higher targeting costs are acceptable if they lead to sufficiently better targeting of transfers. When interpreting the relative size of administrative costs across programs, it is also important to recognize that some costs are fixed (i.e., independent of the number of households included in the program and/or of the transfer levels given to households) so that relative the cost-effectiveness of programs is sensitive to the size of the program. Focusing on fixed targeting-related costs, this means that expensive targeting methods are only likely to be warranted for large programs, that is, programs with large transfer levels and/or a large number of beneficiaries).[3]

- *Private Costs:* Households also incur private costs involved in taking up transfers. For example, workfare programs involve households incurring an opportunity cost in terms of forgone income opportunities. Queuing involves similar, though usually much smaller, opportunity costs. Households may face cash costs for obtaining certifications required for the program, such as a national identity card or proof of residency or of disability, and for transportation to and from program offices. Private costs, which are often overlooked when evaluating programs, may be quite important, especially when self-selection methods are used or when access to the program is conditioned on actions (e.g., keeping children in school) by the household. Indeed, Duclos (1995, p. 410), estimates that even for Great Britain's Supplemental Benefit—a means-tested cash transfer not particularly reliant on self-targeting—"approximately one-fifth of the total income support budget is lost to recipients in the form of various take-up inconveniences."

- *Incentive Costs*: These are often referred to as *indirect costs*. They exist because the presence of eligibility criteria may induce households to change their behavior in an attempt to become beneficiaries. For example, a program open only to those below a minimum income may cause some households to reduce their labor supply and thus their earned incomes. This is one of the reasons why transfers that guarantee a minimum income irrespective of earnings are not considered desirable. Other examples of such "negative incentive effects" are higher consumption of subsidized commodities, crowding out of private transfers (Cox and Jimenez 1995; Jensen 1998), relocation/migration, or devoting resources to misreporting. Indirect effects may also be positive, for example, when transfers are conditioned on household behaviors such as the enrollment of children in school or attendance at health clinics.[4]

Though labor disincentive effects are an important concern in the development of many OECD countries' welfare programs (Moffitt 1992, 2003), they may be less important in developing country safety net programs for several reasons:

- Direct means tests are not the most common targeting method and are especially rare in low-income countries.
- Transfers are rarely graduated. Thus, only those around the cut-off point have an incentive to change their behavior so as to be deemed eligible for transfers. The smaller the transfer is, the lower is the number of people likely to be affected.
- Benefit levels are usually low, implying that recipients will maintain a strong incentive to choose additional earnings over additional leisure when they have a choice.

Nonetheless, in principle, such labor-disincentive effects cannot be ignored or assumed not to exist.[5]

One way of minimizing disincentive effects would be to keep the population relatively uninformed about the detailed eligibility criteria being used, for example, letting the population know that it is based on some concept of poverty but not providing the details of how this is actually measured. Such lack of transparency may in itself be seen as an undesirable characteristic of program design. Basing eligibility on information or characteristics collected prior to the program is another way to eliminate the problem, assuming that households were not answering strategically in anticipation of a program. However, the need for periodic recertification will require the eventual use of updated information on characteristics so that the incentive problem will arise.

- *Social Costs*: These costs may arise when the targeting of poor households involves publicly identifying households as poor, which may carry a social stigma. If the poorest households do not take-up the transfer as a result, then this decreases the effectiveness of the program at getting transfers into the hands of the poorest. Such issues obviously take on additional importance when one appeals to concepts of poverty such as Sen's "capabilities" (Sen 1988).
- *Political Costs*: Excluding the middle classes may remove broad-based support for such programs and make them unsustainable if voter support determines the budget and is in turn determined by whether the voter benefits directly from the program.[6] On the other hand, efficient targeting to ensure that only those in need receive benefits may actually increase political support from those who support it based on its indirect benefits to them of reducing poverty (such as a feeling of social justice, being hassled by fewer beggars, lower likelihood of property theft, increased political stability, or lower taxes). Of course, political support may come from interest groups who are suppliers to the program or advocates for its beneficiaries—farmers' and teachers' unions may support school lunch programs on these grounds.

The relative importance of the above costs will differ across targeting methods and also across different sociopolitical environments. For example, it is likely that administrative costs are more important when individual or household assessment is used. Incentive costs are likely to be less important when categorical targeting is used. Private costs are likely to be more important when self-selection is used. While the nature and importance of social costs may differ widely with the form of self-selection inherent in the program design, all of these costs need to be considered when evaluating the targeting effectiveness of programs.

Measuring Targeting Performance

In practice program officials do not have perfect information about who is poor because this information is difficult, time consuming, and costly to collect. Thus, when basing program eligibility on imperfect information, they may mistakenly commit errors of inclusion—identifying nonpoor persons as poor and therefore admitting them to the program, or errors of exclusion—identifying poor persons as not poor and thus denying them access to the program. In a world of unlimited resources, such errors could be greatly minimized by collecting additional information. However, in a world of limited resources, policy makers and program managers need to know whether such costs are justified in terms of improved targeting. Further, governments will wish to determine how effective a given targeted intervention is. Both exercises require a measure of targeting performance.

A common approach to evaluate the targeting performance of alternative transfer instruments is to compare undercoverage and leakage rates. *Undercoverage* is the proportion of poor households that are not included in the program (errors of exclusion).[7] *Leakage* is the proportion of those who are reached by the program who are classified as nonpoor (errors of inclusion).[8]

Analysis is often presented in terms of a two-by-two matrix (table 2.1).[9] Consider a case where there are a 100 households and a poverty line that implies that 40 of these are classified as poor. Next consider a program that

Table 2.1. Calculating Leakage and Undercoverage Rates

	Welfare Status of Households		
	Poor	*Non-Poor*	*Total*
Households excluded from program	10 (U = 25%) (Exclusion error)	50 (Successful targeting)	50
Households included in program	30 (Successful targeting)	10 (L = 25%) (Inclusion error)	50
Total	**40**	**60**	**100**

Note: U and L denote under-coverage and leakage rates, respectively.

gives benefits to 40 households selected according to imperfect targeting criteria. Of these, 30 are poor (i.e., they have incomes below the poverty line) and 10 are nonpoor (i.e., they have incomes above the poverty line). Both the 30 poor households included in the program and the 50 nonpoor households excluded are treated as successful targeting. The 10 poor households excluded are treated as "errors of exclusion, while the 10 non-poor households are seen as "errors of inclusion." Here, $N_{p,o} = 10$, $N_p = 40$ and so the undercoverage rate is 25 percent. Further, $N_{np,i} = 10$, $N_i = 40$ and so the leakage rate is also 25 percent.[10]

In general actions taken to reduce one kind of error may cause the other to increase. Introducing more stringent rules in order to identify need so as to screen out the nonpoor will, for example, also make it more difficult for the poor to provide the necessary information. Thus, while meant to reduce errors of inclusion, it will also raise errors of exclusion. Similarly, raising the cut-off point in an (imperfect) proxy-means score in order to reduce undercoverage will also tend to increase leakage.

In practice, the inevitability of targeting errors affects the decision about whether to target, how precisely to target, and the method used for targeting. First, it reduces the potential benefit; the illustration in section 2.1 assumed perfect targeting and thus exaggerated the benefit from targeting. Second, the fact that both types of targeting errors will occur and are generally inversely linked means that policymakers must decide how well they can tolerate each. An error of inclusion wastes program resources (e.g., by leaving less for "poor" households or by increasing the budget required to have the same poverty impact) and thus makes the program inefficient. An error of exclusion leaves that person without help and makes the program ineffective at reducing poverty. Both are undesirable, and different policy makers may have different views about which is worse.

This approach has several limitations (Coady and Skoufias 2001):

- First, it discards much distributional information. Surely it is better to give a transfer to someone just over the poverty line than to someone at the very top of the distribution, but both count equally as errors of inclusion. Similarly, benefits to the very poorest as opposed to those just below the poverty line count equally as success cases, although the former is presumably more desirable.[11]
- Second, it focuses only on who gets the transfers and not on how much households get (i.e., the size of the transfer budget and the differentiation of transfer levels across households).
- Third, when comparing across programs it is often the case that those that do well on undercoverage simultaneously score badly on leakage. For example, universal programs would be expected to score relatively well on undercoverage but poorly on leakage, but the leakage/undercoverage approach does not address the issue of trade-off. The core problem is that a focus solely on leakage and undercoverage fails to make explicit how program managers, policy makers, or society itself weights the benefits of transferring resources to different groups, for example, the moderately versus extremely poor.

Three alternatives overcome these limitations:

- One approach is based on the *distributional characteristic* more commonly used in the literature on commodity taxation (Newbery and Stern 1987; Ahmad and Stern 1991; Coady and Skoufias 2001). This approach builds an index of society's welfare, summing across individuals and using explicit welfare weights for different kinds of individuals.[12] The attraction of this index is that welfare weights are made more transparent and that it generalizes from familiar simple cases. For example, if poor households are given a welfare weight of one and nonpoor households a weight of zero, and if we further assume that all beneficiary households receive the same level of transfer, then this index collapses to the proportion of households receiving transfers that are classified as poor (or 1 minus the rate of leakage). If, in addition, we know the level of benefits received by beneficiaries, then it collapses to the share of the program budget received by poor households. Where the "poor" are defined as households falling within the bottom deciles (e.g., 20 percent or 40 percent) of the national income distribution, similar indices can be calculated. Generally, all that is required to calculate the distributional characteristic is mean incomes by decile and decile shares in transfers. The administrative cost side of the program can be easily incorporated by including this cost in the denominator along with total transfers.
- An alternative to specifying welfare weights either implicitly or explicitly is to calculate the share of the program budget going to, for example, the various deciles or quantiles of the national income distribution. The numbers can relate to either proportions of beneficiaries or proportion of total transfers. One can focus on whatever part of the distribution that one wishes, although one should be clear that this implicitly involves specifying welfare weights. For example, focusing on the share of the transfer budget accruing to the bottom 20 percent of the distribution is equivalent to attaching a welfare weight of unity to these households and zero to others. If, in addition to the shares of total transfers received by each decile, one also presents mean incomes, then one provides sufficient information for the calculation of the distributional characteristic
- A third approach reframes the issue. Rather than asking how effective the program is at identifying the poor, it asks how effective it is at reducing poverty. It proceeds by comparing the relative impacts of the alternative instruments on the extent of poverty subject to a fixed common budget or, equivalently, the minimum cost of achieving a given reduction in poverty across instruments (Ravallion and Chao 1989; Ravallion 1993). This explicitly incorporates into the previous approaches the size of transfers and the budget, in addition to how transfer levels are differentiated across households in different parts of the income distribution.

A final complication in evaluating targeting outcomes stems from the fact that the program analyst faces many of the difficulties in correctly measuring welfare that the program official faces. Not only is income difficult to measure for those with irregular incomes or entwined household and small business accounts; the household survey information that the analyst usually relies on may not use exactly the same concepts for income, time period, or unit of observation that the program does. Moreover, household welfare may have changed between the time the household sought entry to the program and when it was surveyed. Duclos (1995) expands this analysis and shows that analyst error can lead to substantial mis-estimates of takeup rates and targeting errors.

Classifying Targeting Methods

Targeting methods all have the same goal—to correctly and efficiently identify which households are poor or which are not. To understand the effectiveness of these approaches, it is useful to distinguish between *methods* and *actors*. *Methods* refer to the approaches taken to reach a target group. Below, we divide these into three groups: individual/household assessment, categorical targeting,[13] and self-selection. *Actors* refer to the identity of the individuals who perform two roles: the implementation of the targeting method and the subsequent implementation of the intervention.

Individual/Household Assessment is a method in which an official (usually a government employee) directly assesses, household by household or individual by individual, whether the applicant is eligible for the program. It is the most laborious of targeting methods. The gold standard of targeting is a *verified means test* that collects (nearly) complete information on a household's income and/or wealth and verifies the information collected against independent sources such as pay stubs or income and property tax records. This requires the existence of such verifiable records in the target population, as well as the administrative capacity to process this information and to continually update it in a timely fashion. For these reasons verified means tests are extremely rare in developing countries where the poorest households receive income from a myriad of diverse sources and formal record keeping is nonexistent. Other individual assessment mechanisms are used in the absence of the capacity for a verified means test. Three common ones are simple means tests, proxy means tests, and community-based targeting.[14]

- Simple means tests, with no independent verification of income, are not uncommon. A visit to the household by a program social worker may help to verify in a qualitative way that visible standards of living (which reflect income or wealth) are more or less consistent with the figures reported. Alternately, the social workers' assessment may be wholly qualitative, taking into account many factors about the household's needs and means but not having to quantify them. These types of simple means tests are used for both direct transfer programs and

for fee-waivering programs, with or without the visit to the household. Jamaica's food stamp program, implemented in the 1980s, is an example (Grosh 1992).

- *Proxy means tests,* while relatively rare, are being instituted in a growing number of countries. We use the term to denote a system that generates a score for applicant households based on fairly easy to observe characteristics of the household such as the location and quality of the dwelling, ownership of durable goods, demographic structure of the household, and the education and, possibly, occupations of adult members. The indicators used in calculating this score and their weights are derived from statistical analysis (usually regression analysis or principal components) of data from detailed household surveys of a sort too costly to be carried out for all applicants to large programs. The information provided by the applicant is usually partially verified by either collecting the information on a visit to the home by a program official, as in Chile's unified family subsidy (Sancho 1992) or by having the applicant bring written verification of part of the information to the program office, as done in Armenia (World Bank 1999).

- Community based-targeting uses a group of community members or a community leader whose principal functions in the community are not related to the transfer program to decide who in the community should benefit. School officials or the parent-teacher association may determine entry to a school-linked program. A group of village elders may determine who receives grain provided for drought relief, or special committees composed of common community members or a mix of community members and local officials may be specially formed to determine eligibility for a program. The idea is that local knowledge of families' living conditions may be more accurate than the results of a means test conducted by a government social worker or a proxy means test.

Categorical Targeting refers to a method in which all individuals in a specified category—for example, a particular age group or region—are eligible to receive benefits. This method is also referred to as *statistical targeting, tagging, or group targeting*. It involves defining eligibility in terms of individual or household characteristics that are fairly easy to observe, hard to falsely manipulate, and correlated with poverty. Age, gender, ethnicity, land ownership, demographic composition, or geographical location are common examples that are fairly easy to verify. Age is a commonly used category, with cash child allowances predominant in transition countries, supplemental feeding programs for children under five common in poor countries, and noncontributory pensions for the elderly common in many places. As we show in chapter three, geographic targeting is even more common, often used in combination with other methods. Unemployment or disability status is somewhat harder to verify, but cash assistance to these groups may be categorically targeted as well. In chapter three, we review results for

geographic, demographic, and other categorical methods. In chapter four, we treat geographic and demographic targeting by age in depth.

Under *self-selection*, the program has universal eligibility, but the design involves dimensions that are thought to encourage the poorest to use the program and the nonpoor not to do so.[15] This is accomplished by recognizing differences in the private participation costs between poor and non-poor households. For example, this may involve:

- 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 jobs
- restriction of transfers to take place at certain times with a requirement to queue
- transfer of in-kind benefits with "inferior" characteristics (e.g., low quality wheat or rice)
- location of points of service delivery (e.g., ration stores, participating clinics or schools) in areas where the poor are highly concentrated so that the nonpoor have higher (private and social) costs of travel.

Universal food subsidies can be viewed as a form of self-selection since they are universally available and households receive benefits by consuming the commodity. In practice, households can often determine not just whether to participate but also the intensity of their participation. Tunisia's reformed milk subsidy program, whereby milk subsidies are higher for reconstituted milk in inconvenient and small packages than for other grades and packaging of milk, is an example of a self-targeted intervention (Tuck and Lindert 1996), as is a public works program in Maharashtra State, India, called the *Employment Guarantee Scheme* (Datt and Ravallion 1994).

Whereas methods refer to "how" targeting is undertaken, actors refer to "who" targets and "who" implements these interventions. Actors can include central government officials; lower state, municipality, or district level officials; private sector contractors; and community members such as teachers, health clinic staff, and elders. The decision whether to decentralize both the identification of beneficiaries and the provision of the program will hinge on several factors: which actors can provide the most cost-effective source of information on individual, household or locality circumstances; which actors can deliver the intervention most cost-effectively; and whether different actors have the incentive to target and implement the intervention in the manner desired by those who fund the program.

In reviewing this menu of targeting options, policy makers should be mindful of two important considerations. First, individual targeting methods are not mutually exclusive and can be used in different combinations and sequences. A child allowance (categorical targeting) may be means tested (individual assessment). Subsidized coarse grain (self-targeting) may be available for sale only in food shops in poor neighborhoods (geographic targeting). In fact, the use of a single targeting method is not the norm; 60 percent of the interventions described in the next section used two or

more methods. Second, when assessing whether a particular intervention reaches its intended beneficiaries, it is important to be cognizant of four dimensions:

- type of interventions chosen—for example, a food-for-work program will, by design, exclude poor people who are physically unable to work
- targeting method chosen
- identity of the actor who undertakes this targeting
- identity of the actor who provides the intervention.

We describe each of these features in our more detailed descriptions of individual programs in the bibliography (Coady, Grosh, and Hoddinott 2003).

Notes

1. Discussions of the benefits and costs of targeting in developing countries are also found in Atkinson (1995), Besley and Kanbur (1993), Devereux (1999), Grosh (1994), Hoddinott (2001), and van de Walle (1998).

2. See Besley and Kanbur (1993) for a more detailed discussion of the issues discussed in this paragraph.

3. See Coady, Perez, and Vera-Llamas (2000) for more detailed discussion.

4. Skoufias (2001) provides a review of these positive effects in his synthesis of the impact of the anti-poverty program *PROGRESA* in Mexico.

5. Sahn and Alderman (1995) provide an instructive case study of the impact of the means-tested rice subsidy on labor supply in Sri Lanka. They find no impact on the likelihood of work, but that days worked (conditional on working) fell by around 10 percent. They caution, however, that increases in leisure are utility enhancing, particularly in environments where individuals engage in heavy manual labor.

6. Case studies of these political issues are found in Pinstrup-Andersen (1988), Besley and Kanbur (1993), Adams (1998, 2000) and Ahmed, Bouis, Gutner, and Lofgren (2001).

7. In mathematical notation:

$$U = \frac{N_{p,o}}{N_p}$$

where $N_{p,o}$ is the number of poor households who are left out of the program and N_p is the total number of poor households.

8. In mathematical notation:

$$L = \frac{N_{np,i}}{N_i}$$

where $N_{np,i}$ is the number of nonpoor households in the program and N_i is the total number of households in the program.

9. Cornia and Stewart (1995) provide further discussion and explanation.

10. The fact that undercoverage and leakage rates coincide reflects the fact that the number of persons in the program is equal to the number of poor persons so that the number of persons wrongly excluded will always equal the number of persons wrongly included. If the number of eligible persons is greater (less) than the number of poor persons, then this will inevitably tend to increase (decrease) leakage

and decrease (increase) undercoverage rates. It is important to note that this mistargeting is in a sense due to the wrong program size (i.e., the number of eligible persons being different from the number of poor persons) as opposed to imperfect targeting *per se*. For example, if the number of eligible persons were 30, then even if all those selected were poor undercoverage would still be 10 percent.

11. This shortcoming can be overcome somewhat by plotting the percentages of households in, say, each decile that are mistargeted against mean decile incomes. Mistargeting left of the poverty line represents undercoverage within each decile, while that to the right represents leakage by decile, Skoufias and Coady (2002).

12. The common form of the index is defined as:

$$\frac{\sum_{h} \beta^h T^h}{\sum_{h} T^h} = \sum_{h} \beta^h \theta^h$$

where β^h is the social valuation of income transferred to household h (or its "welfare weight"), T^h is the level of the transfer to the household, and θ^h is each household's share of the total program budget.

13. Collectively, these two methods are often referred to as forms of *administrative targeting*.

14. A fourth example is the assessment of *nutritional status,* such as low weight for age or growth faltering. They can be used to target increased medical care or transfers in cash or kind, such as food during disaster relief. Such programs often operate in conjunction with the health system. We do not review them here.

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

3

The International Evidence on Targeting Outcomes

While there is a fairly rich literature on targeted programs, much of it documents single programs. Even comparative pieces tend to cover only a single region, method, or intervention. This partial coverage frustrates efforts to make broader assessments about the effectiveness of different targeting methods or to draw policy-relevant lessons. The purpose of this chapter is to address this problem. We do so by first explaining how we constructed a database of targeted antipoverty interventions and then use these data to address three questions:

- What targeting outcomes are observed?
- Are there systematic differences in targeting performance by targeting methods and other factors?
- What are the implications of such systematic differences for the design and implementation of targeted interventions?

This quantitative analysis is complemented by a qualitative discussion of individual targeting methods in the chapter that follows.

Database Construction

The first step in our analysis was an extensive literature review and the construction of a database of targeted antipoverty interventions.[1] To our knowledge, this work represents the most extensive attempt to construct such a database.

Our criteria for inclusion in this database were the following:

- The intervention had to be situated in a low or middle-income country.
- A principal objective of the intervention is poverty reduction defined in terms of income or consumption.

- Documentation on the intervention contains information on the targeting method used, its implementation, and outcomes.
- The intervention is relatively recent (generally from 1985–2002).

Included in our data are cash transfers (including welfare and social assistance payments, child benefits, and noncontributory pensions), near-cash transfers (such as quantity rationed subsidized food rations 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, a number of interventions that are included have objectives that include, but are not limited to, direct poverty reduction. Social funds are a good example. While short-term poverty reduction can be an important component of these interventions, so too 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 requires caution when translating evidence on targeting performance into broader assessments of the relative welfare impacts of these program types.
- Second, focusing the review in this way necessarily means excluding a number of interventions that may be targeted and that may have some impact in terms of poverty reduction. Excluded are the following:
  - Occupationally based transfer schemes such as formal sector unemployment insurance or occupational old-age or disability pensions: In these, the principal mechanisms that determine eligibility and benefit levels are employment and contributions history rather than current poverty status.
  - Credit and microcredit schemes: although these are often targeted, they are largely motivated by credit market failures.
  - Supplementary feeding programs: our foray into the vast literature on this type of intervention did not yield studies that satisfied the criteria described above.
  - Most short-term emergency aid: although this has a clear poverty focus and is often targeted by need, the time scale on which it operates typically precludes an assessment of the distribution of the benefits.
  - Fee waivering in health care (see a companion piece by Bitran and Giedion 2003 commissioned in parallel with this book).

Most studies of targeting—especially those outside of Latin America and the Caribbean—do not appear in peer-reviewed journals. Consequently, we undertook searches of the "gray" literature using Web search engines found at the World Bank, Eldis, and IFPRI websites using the following key words: *safety nets*, *targeting*, *social funds*, *pensions*, *public works*, *subsidies*. Additional cases were found via canvassing colleagues about work that had not yet

been catalogued in these places. Searches were also undertaken in the following academic journals for the years 1990–2002: *Economic Development and Cultural Change*, the *Journal of Development Economics*, the *Journal of Development Studies*, the *Journal of Public Economics*, the *World Bank Economic Review*, the *World Bank Research Observer* and *World Development,* and *Economic and Political Weekly* for 1998–2002. Additional cases were found through reviews of existing compilations such as Grosh (1994) and Braithwaite, Grootaert, and Milanovic (2000).

Given the nature of such a search, it is important to remember that our sample of interventions is not necessarily reflective of the distribution of programs that exist in the world. Rather our sample represents those programs that have some measurement of targeting outcomes and that have been published in the catalogued English language literature that we have been able to access. 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.
- It is a program that by virtue of methods or setting is deemed attractive by analysts and editors. For example, we suspect that programs using community-based methods and agents are underrepresented. These are often only funded locally and the method is chosen when there are poor data and low administrative capacity, features that all reduce the likelihood of an evaluation being conducted and finding its way into the international literature. For similar reasons, it is likely that we underrepresent the literature on public works in Sub-Saharan Africa. In contrast, proxy means tests are well represented, with a large share of all such programs in the world showing up in this sample.

Lastly, we note that some programs described here have been reformed or in some cases eliminated since the studies were written. The lessons from how the old versions of programs worked are still valid guides on program design and implementation. However, the sketches of programs provided both here and in chapter four may no longer reflect actual practice in the program or country.

Programs Identified

Based on the criteria described above, we were able to collect information for 122 interventions drawn from 48 countries. Table 3.1 provides a description of the distribution of these interventions across regions, income groups, and intervention types. This sample of interventions provides a fairly broad

Table 3.1. Distribution of Interventions by Region and Country Income Levels

	Transfers			Subsidies		Public works for	
	Cash	Near cash	Food	Food	Nonfood	Job creation	Program output (e.g., social funds)
By Regions							
Latin America and Caribbean, 32	14	3	4	0	2	4	5
Eastern Europe and former Soviet Union, 26	24	1	0	0	0	0	1
Middle East and North Africa, 14	1	0	0	13	0	0	0
Sub-Saharan Africa, 13	4	0	1	4	1	2	1
South Asia, 21	1	13	3	0	0	4	0
East Asia, 16	5	1	4	1	2	2	1
By Income Level							
Poorest, 63	17	15	6	10	2	8	5
Less poor, 59	32	3	6	8	3	4	3
Total 122	49	18	12	18	5	12	8

Notes:
1. Numbers in italics are number of interventions by region and income level.
2. Poorest countries have per-capita GDP in PPP dollars below 1,200; less-poor countries have per-capita GDP above 1,200 and below 10,840.
3. Near-cash transfers include interventions such as food stamps or the right to purchase a limited quantity of food at a subsidized price.

regional coverage. Although cash transfer programs account for a large proportion (40 percent) of the interventions, the other intervention types are well represented. In some regions, a particular intervention type dominates: cash transfers in Eastern Europe and the Former Soviet Union and Central Asia (ECA), universal food subsidies in the Middle East and North Africa (MENA), and near-cash transfers in South Asia (SEA). By contrast, there is a wider mix of reported interventions in other regions. Most of the cash transfer programs occur in Latin America and the Caribbean (LAC) and ECA, most of the near-cash transfer programs occur in South Asia, most of the universal food subsidies occur in MENA, and most of the social funds occur in LAC. Dividing the sample by per capita Gross Domestic Product (GDP) levels, we find that cash transfer programs are more likely to be found in less poor countries and near-cash transfers in the poorest countries.

Table 3.2 provides information on the distribution of interventions and their targeting methods. The first thing to notice is that interventions use a combination of targeting methods; in all we have 253 occurrences of different targeting methods, so that the interventions in our sample use just over two different targeting methods on average. Just 48 interventions use a single targeting method, while 42 use two methods, 21 use three methods, and 11 use four methods.

There are some marked differences by region. Most of the interventions using means and proxy means testing are concentrated in ECA and LAC. A legacy of the central planning era in ECA has been an extensive administrative system that is suited to the individual assessment of individual circumstances 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 in this region is based either on some form of individual assessment or individual characteristic such as age. Reliance on food subsidies explains why self-targeting based on consumption patterns is the dominant targeting method in MENA. SEA is notable for its extensive use of geographic targeting as well as a relatively high reliance on self-selection based on work or consumption. LAC countries also use geographic targeting extensively, but this is more often accompanied by either direct individual assessment (i.e., means or proxy means testing) or by targeting children. The small number of documented programs for sub-Saharan Africa (SSA) and East Asia and the Pacific (EAP) show more mixed patterns.

There are also broad differences across income levels. Generally, poorer countries tend to rely more on self-selection methods and categorical targeting, whereas forms of individual assessment 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.

Although certain program types are synonymous with certain targeting methods, most use a combination of methods, presumably because there is synergy from the perspective of targeting efficiency. Public works programs typically use a combination of geographic targeting and self-selection based

Table 3.2. Distribution of Targeting Methods by Region, Country Income Level, and Program Type

	Individual Assessment			Categorical				Self-selection		
	Means tests	Proxy means tests	Community assessment	Geography	Age: elderly	Age: children	Other	Work	Consumption	Community bidding
By region										
Latin America and Caribbean, 68	8	5	3	20	4	14	4	4	0	6
Eastern Europe and former Soviet Union, 46	14	1	3	1	6	11	7	2	0	1
Middle East and North Africa, 23	4	0	0	2	1	1	2	0	12	1
Sub-Saharan Africa, 25	3	0	2	3	5	1	4	2	4	1
South Asia, 49	2	1	3	16	2	1	10	4	10	0
East Asia, 42	3	1	3	10	6	8	8	1	1	1
By income level										
Poorest, 147	12	3	10	37	10	14	28	8	19	6
Less poor, 106	22	5	4	15	14	22	7	5	8	4
By program type										
Cash transfer, 103	24	4	5	9	19	24	16	2	0	0
Near-cash transfer, 36	4	3	0	12	1	2	4	0	10	0
Food transfer, 35	0	1	5	9	3	9	7	0	0	1
Food subsidy, 23	3	0	0	2	0	0	0	0	0	1
Non-food subsidy, 9	3	0	0	2	1	0	2	0	17	0
Public works, job creation, 29	0	0	2	10	0	0	6	11	0	0
Public works, program output (e.g., social fund), 18	0	0	2	8	0	0	0	0	0	8
Total 253	34	8	14	52	24	36	35	13	27	10

Notes:
1. Many programs use more than one targeting method. Thus the total number of targets methods tallied is greater than the number of programs.
2. Poorest countries have per-capita GDP in PPP dollars below 1,200; less-poor countries have per-capita GDP above 1,200 and below 10,840.

on low wages and a work requirement. In practice, however, public works also often require additional rationing of employment using categorical targeting if demand exceeds supply at the wage paid. 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, which are most likely to have some form of individual assessment, are often conditioned on other characteristics, such as age in the case of pensions or child benefits.

Indicators of Targeting Performance

To compare the performance of the different targeting methods used in the range of programs considered in our analysis, we need a comparable performance indicator for each program. In such a meta-analysis, the definitions, methods, and presentations in the original studies vary in ways that make it difficult to assemble such 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/nonpoor classification that differs by country. Other studies report none of these measures but use other less common ones. Of course, the measure of welfare used is not always strictly comparable from study to study. Thus we are faced with how best to compare targeting performance outcomes using data that are not strictly comparable. (See box 5.1 for a plea for greater comparability in reporting.)

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

- the proportion of total transfers received by 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 or beneficiaries going to "poor" households, where the poor are defined in terms of some specified part of the welfare distribution (e.g., falling in the bottom 35 percent of the income distribution).

As discussed in chapter two, ideally we would like to know the proportion of total transfers received by households falling within different centiles (40th, 20th, 10th, and so on) of the national income distribution. This is a better measure than the proportion of beneficiaries by centile because in the case of the latter, we do not necessarily know anything about variations in the levels of transfers. These two measures—proportions of total transfers and proportions of beneficiaries—are only equivalent when transfer levels are uniform across beneficiaries.

Given that no single common measure of targeting performance is available, we have constructed a measure based on a comparison of actual

performance to a common reference outcome, namely, 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. One can think of neutral targeting as arising either from the random allocation of benefits across the population or a universal intervention in which all individuals received identical benefits. The indicator used in our analysis is constructed by dividing the actual outcome by the appropriate neutral outcome. For example, if the bottom 40 percent of the income distribution receive 60 percent of the benefits, then our indicator of performance is calculated as (60/40) = 1.5, thus a higher value is associated with better targeting performance. A value of 1.5 means that targeting has led to the target group (here those 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 less than one for regressive targeting, with unity denoting neutral targeting.

The performance indicator used in the analysis below is based on a lexicographic selection process among the available incidence indicators. Where it is available, we base performance on the proportion of benefits accruing to the bottom two quintiles. Where this is not available, we base it on the proportion of benefits accruing to the bottom quintile, then benefits to the bottom decile, and lastly, the share of program benefits received by individuals deemed to be below a poverty line. We can calculate such a performance indicator for 85 programs in 36 countries. In section 3.6 we discuss how sensitive our results are to the use of this method.

Descriptive Analysis

Table 3.3 lists all programs for which we can construct our performance indicator from best to worst. Targeting performance varies enormously, ranging from 4 for the Trabajar public works program in Argentina to 0.28 for value-added tax (VAT) exemptions on fresh milk in South Africa. The median value is 1.25, so that the "typical" program transfers 25 percent more to the target group than would be the case with a universal or random allocation. However, a staggering 21 of the 85 programs—25 percent—are regressive, with a performance index less than 1. In these cases, a random selection of beneficiaries would actually provide greater benefits to the poor. Median performance rises to 1.3 if interventions using self-selection based on consumption, which account for a large proportion of regressive programs, are withdrawn from the sample. Doing so also reduces the proportion of regressive interventions to 16 percent.

It is instructive to focus on the worst and best 10 programs. The worst 10 have a median score of only 0.60, ranging from 0.28–0.78, and are mainly from SSA and MENA, with three from South Africa's VAT exemption program. Seven out of the 10 are food subsidy programs, and two of the remaining three programs involve cash transfers. It is also noticeable that only

Table 3.3. Targeting Performance of Antipoverty Interventions for Sample Programs in 1985–2003

			Individual Assessment			Categorical				Self-selection			Share of transfers going to:		
Country	Program	Performance	Means test	Proxy means test	Community assessment	Geographic	Age: elderly	Age: children	Other	Work	Con-sumption	Community bidding	Poorest 10%	Poorest 20%	Poorest 40%
Argentina	Trabajar public works	4.00								✓			60.0	80.0	
Estonia	Cash social assistance	3.47	✓										34.7		
Dominica	Cash transfers	3.00	✓						✓					60.0	
Hungary	Cash social assistance	2.72	✓										27.2		
Albania	Ngihme Ekonomika cash	2.65			✓				✓				36.0	53.0	
Yemen	Social welfare fund cash	2.15	✓			✓	✓		✓				68.0	75.0	86.0
Poland	Social assistance cash	2.10	✓			✓	✓	✓	✓						
Romania	Minimum income guarantee	2.08	✓						✓				21.0		
Chile	SUF cash transfers	2.08	✓							✓				62.0	83.0
Nicaragua	RPS conditional cash transfer	2.02		✓		✓		✓	✓					57.0	83.0
Honduras	PRAF cash transfer	1.99				✓		✓	✓				32.6	55.0	80.9
Chile	School lunches	1.98		✓		✓		✓					22.1	42.5	79.5
														53.0	79.0
Slovenia	Otroski Dodatek child benefit	1.95			✓			✓					42.0	65.0	78.0
Bulgaria	Cash social assistance	1.95	✓			✓							22.0	60.5	78.1
Bolivia	ESF public works	1.93	✓									✓	25.0	31.0	77.0

Table 3.3. (*continued*)

Country	Program	Performance	Individual Assessment			Categorical				Self-selection			Share of transfers going to:		
			Means test	Proxy means test	Community assessment	Geographic	Age: elderly	Age: children	Other	Work	Con-sumption	Community bidding	Poorest 10%	Poorest 20%	Poorest 40%
Kyrgyz Republic	Unified monthly cash benefit	1.85												54.0	74.0
Chile	PASIS cash to poor elderly	1.83		✓			✓							50.0	73.0
Republic of Korea	LPS cash for home care	1.81	✓				✓						32.3	56.8	72.3
Peru	"Glass of milk" program	1.80				✓		✓				✓		42.0	72.0
Chile	PEM/POJH public works	1.78								✓				51.0	71.0
Costa Rica	Noncontributory pensions	1.73	✓				✓		✓					45.0	69.0
Indonesia	JPS-BK health subsidy	1.68				✓			✓					33.6	
India	Rural AP State PDS rice	1.63	✓			✓					✓		19.1	35.8	65.2
Costa Rica	School Lunches	1.63				✓		✓						39.0	65.0
Mexico	LICONSA milk subsidy	1.60	✓			✓		✓							64.0
India	PDS-subsidized jowar	1.58				✓					✓				63.0
Hungary	Child cash allowance	1.57						✓					23.5	39.5	62.6
Mexico	PROGRESA conditional cash	1.56		✓		✓		✓					22.0	39.5	
Colombia	Housing subsidy	1.50	✓											35.0	62.4
Colombia	ICBF day care / school food	1.50		✓				✓						28.0	60.0
Indonesia	JPS-PDK public works	1.48				✓			✓	✓				29.6	60.0

Country	Program											
Jamaica	Food stamps—MCH	1.45					✓	✓			29.0	58.0
Indonesia	JPS education subsidy	1.44			✓		✓	✓	✓		28.8	54.5
India	Rural Mh. State PDS	1.36		✓	✓		✓		✓	10.4	26.3	54.0
Zambia	Housing subsidy	1.35	✓		✓		✓				28.0	54.0
Uzbekistan	Child allowance	1.35	✓		✓		✓	✓	✓		28.4	53.0
Latvia	Family child allowance	1.33			✓		✓	✓		15.0	29.0	53.0
India	PDS-subsidized rice	1.33			✓			✓				
Indonesia	JPS-OPK rice subsidy	1.32		✓	✓		✓		✓		26.4	52.0
Bolivia	Social investment fund	1.30			✓					✓	24.5	52.0
Jamaica	Food stamps program	1.30	✓		✓	✓		✓		✓	31.0	50.0
Honduras	Social fund program	1.25		✓	✓		✓			✓	32.0	50.0
Chile	Viviendas Basicas mortgage	1.25		✓	✓		✓				28.0	50.0
India	PDS-subsidized oil	1.25	✓		✓		✓				28.0	50.0
Sri Lanka	Food stamps	1.25	✓		✓							
South Africa	Maize VAT exemption	1.23					✓	✓	✓			
Vietnam	SGF food transfers	1.22			✓		✓	✓			24.4	
India	PDS-subsidized kerosene	1.20			✓		✓	✓	✓			48.0
Bangladesh	Food for education	1.20			✓	✓	✓	✓				
Morocco	Food subsidies, flour	1.18			✓		✓	✓			23.0	47.0
India	PDS-subsidized sugar	1.13			✓		✓	✓				45.0

Table 3.3. (continued)

Country	Program	Performance	Individual Assessment			Categorical				Self-selection			Share of transfers going to:		
			Means test	Proxy means test	Community assessment	Geographic	Age: elderly	Age: children	Other	Work	Con-sumption	Community bidding	Poorest 10%	Poorest 20%	Poorest 40%
Armenia	Family cash benefit	1.13						✓					14.0	25.0	45.0
Peru	Social fund program	1.10				✓						✓		23.0	44.0
Bulgaria	Cash pensions	1.10					✓							22.0	
Nicaragua	Social fund program	1.10				✓						✓	9.0	23.0	44.0
India	Urban AP State program	1.09				✓					✓		10.4	21.5	43.6
India	PDS rice	1.08				✓					✓		15.0	25.0	43.0
Zambia	Social fund program	1.05	✓			✓							7.6	17.8	41.7
Mozambique	GAPVU cash transfers	1.04	✓			✓			✓				21.0	21.0	41.0
India	Urban Mh. State PDS	1.03				✓					✓		22.6	22.6	40.5
Tunisia	Universal food subsidies	1.01									✓		15.0	15.0	40.0
Uzbekistan	Child/low-income benefit	1.00	✓		✓				✓				11.0	20.0	40.0
Egypt	Universal flour subsidy	1.00									✓			20.0	
India	PDS-subsidized wheat	1.00				✓					✓			21.0	40.0
Latvia	Social assistance program	1.00	✓										21.0		40.0
Romania	Family child allowance	1.00						✓							40.0
Egypt	Universal bread subsidy	0.98									✓			21.0	39.0
Bulgaria	Child/family cash allowance	0.95						✓						19.0	

Country	Program										
Egypt	Universal sugar subsidy	0.95						✓		18.0	38.0
Egypt	Universal oil subsidy	0.95	✓					✓		18.0	38.0
Armenia	Social fund program	0.93	✓						8.0		
Tunisia	Food subsidies	0.93	✓							15.0	37.0
Poland	Family child allowance	0.90	✓					✓	9.0	17.0	37.0
Romania	Discretionary assistance	0.90			✓					18.0	
Morocco	Food subsidies, sugar	0.85	✓		✓					15.0	34.0
Colombia	SSB utility subsidies	0.82		✓						15.0	33.0
South Africa	Beans VAT exemption	0.79	✓								
Latvia	Cash pensions	0.70	✓		✓				5.0	12.0	28.0
Algeria	Universal food subsidies	0.70	✓							12.0	28.0
South Africa	Oil VAT exemption	0.68	✓								
Morocco	Food subsidies, oil	0.60	✓							11.0	24.0
Armenia	Humanitarian assistance	0.58	✓			✓			3.0	8.0	23.0
Colombia	Subsidio Familiar cash	0.50			✓	✓				7.0	20.0
Yemen	Universal food subsidies	0.45	✓							7.0	18.0
Vietnam	Transfers to old/infirm	0.40			✓		✓				
South Africa	Milk VAT exemption	0.28	✓							8.0	
Republic of Korea	Old age pension	0.26			✓						10.5

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

one of the poorly performing programs uses either means or proxy means targeting methods, none of them is geographically targeted, and they come from across the income spectrum. The top 10 programs have a median score of 2.15, range from 2.02–4.0, and are from either LAC or ECA. Eight out of the 10 involve cash transfers. Nine out of the 10 make use of means, proxy means, or geographic targeting, and seven out of the 10 are in less-poor countries.

The fact that cash-transfers feature in both the best and worst 10 programs highlights the possibility that variations in targeting performance may reflect poor implementation rather than poor potential for such programs *per se.* It is, however, noticeable that whereas public works are all in the top half of the performance table, social funds are nearly all in the bottom half. This is consistent with a trade-off between the objective of reducing 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 dominance of less-poor countries among the top half suggests that characteristics correlated with income such as administrative capacity are important determinants of targeting performance.

Table 3.4 provides summary statistics on targeting performance—sample size, median, interquartile range (*iqr*) and the *iqr* as a percentage of the

Table 3.4.   Targeting Performance by Targeting Method

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

median—by targeting method. First impressions suggest that table 3.4 yields a clear hierarchy in terms of targeting performance. Interventions using forms of individual assessment have better incidence than those relying on forms of categorical targeting that in turn out perform interventions that use self-selection, much as one would expect. A closer inspection, however, reveals that such impressions are too general to be very useful. First, there is much heterogeneity within these broad methods of targeting. Most notably, the category of self-selection includes interventions utilizing a work requirement that have the highest median performance and self-selection based on consumption, which has the lowest median. Second, three specific methods—categorical targeting to the elderly, self-selection 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 *iqr* as a percentage of the median. This suggests that, *ceteris paribus*, even the best examples of these targeting methods produce relatively small targeting gains. By contrast, while other methods report higher median values, they are also characterized by proportionately higher variations in targeting effectiveness. So while these methods offer potentially large gains, there is no guarantee that they will improve targeting performance.

One way of exploring the source of variation in targeting outcomes is by using a Theil inequality index. A desirable feature of the Theil index is that it allows the analyst to allocate total variation in a sample into that part due to differences between the means of subgroups and that part due to variation within each of those subgroups. Grouping by targeting method (according to whether they use geographic, means/proxy means, both, or other targeting methods) explains only 20 percent of the total variation. When programs are grouped by region, we find that variation in average performance across continents explains only about 28 percent of total variation. Grouping according to program type, we find that variation in average performance among programs explains 36 percent of the total variation.

One way of interpreting these large variations is in terms of implementation effectiveness. No matter how well one chooses among methods or programs, effectiveness of implementation is a key factor in determining targeting performance. This point is further illustrated 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. In chapter four, we return to this issue in much more detail.

Regression Analysis

Readers unfamiliar or uncomfortable with regression analysis might wish to skip ahead to section 3.6, the caveats, or to section 3.7, a summary of these findings.

Although factors other than choice of method or program may be relatively large, this does not mean that these choices are unimportant. To elucidate the importance of these choices, table 3.5 presents the results of a

Table 3.5. Multivariate Analysis of Targeting Performance

	Basic Results				Dependent variable is in levels (5)	Uses median regression (6)
	(1)	(2)	(3)	(4)		
Log GDP per capita	0.120		0.181	0.150	0.201	0.133
	(2.10)		(3.27)	(2.95)	(2.88)	(2.35)
Log Voice	0.161			0.176	0.217	0.159
	(2.06)			(2.88)	(2.52)	(1.65)
Log Gini coefficient	0.363			0.110	0.439	0.333
	(1.75)			(0.39)	(1.17)	(1.14)
Means testing		0.284	0.285	0.300	0.397	0.306
		(2.43)	(2.59)	(2.90)	(2.84)	(2.36)
Proxy means testing		0.252	0.108	0.078	0.021	−0.054
		(1.43)	(0.74)	(0.45)	(0.10)	(0.28)
Community assessment		0.198	0.119	−0.012	−0.017	−0.156
		(0.97)	(0.59)	(0.09)	(0.10)	(0.59)
Geographic		0.260	0.336	0.341	0.381	0.370
		(2.72)	(3.25)	(3.33)	(2.56)	(2.42)
Age: elderly		−0.055	−0.089	−0.117	−0.124	−0.145
		(0.37)	(0.68)	(0.091)	(0.75)	(0.59)

	(1)	(2)	(3)	(4)	(5)	(6)
Age: young		0.198	0.128	0.140	0.131	0.188
		(1.98)	(1.34)	(1.53)	(0.94)	(1.89)
Other categorical		−0.132	0.028	0.137	0.061	0.088
		(0.88)	(0.22)	(1.23)	(0.32)	(0.42)
Work		0.511	0.404	0.404	0.650	0.314
		(3.66)	(3.74)	(3.73)	(2.83)	(2.22)
Community bidding		−0.018	−0.012	−0.029	−0.117	−0.149
		(0.17)	(0.11)	(0.24)	(0.69)	(1.12)
F-statistic	3.46	3.56	5.41	7.20	5.89	
R^2	0.318	0.406	0.484	0.545	0.572	
Sample size	84	85	85	84	84	84

Notes:

1. All specifications contain controls, not reported, indicating whether performance measure is based on proportion of benefits going to the (a) bottom quintile, (b) poorest decile, (c) to the "poor" or (d) proportion of poor found in population.

2. Specifications (1)–(5) estimate standard errors using the methods proposed by Huber 1967 and White 1980. Specification (6) calculates standard errors using the bootstrap with 50 repetitions.

3. Specifications (1)–(4) and (6) express the dependent variable in logs; specification (5) uses levels.

4. Absolute value of t-statistics in parentheses.

series of regressions that identify how performance varies systematically across these choices as well as country characteristics. In doing so, we note that targeting methods are themselves choices; they are not "exogenous" or "pre-determined." Consequently, it is incorrect to treat these results as causal relations. Rather, they are measures of partial correlation or association.

Our first specification explores how country characteristics such as income, government accountability, and inequality are associated with (log) incidence. Income is measured as log GDP per capita (in purchasing power parity (PPP) dollars) as of 1995.[2] 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). Compiling subjective perceptions regarding the quality of governance in different countries using sources such as polls of experts, commercial risk-rating agencies, and cross-country surveys, they define voice, perhaps more accurately described as "voice and accountability," as a composite measure based on aspects of political processes, civil liberties, and political rights. In so doing, they capture the extent to which citizens participate in the selection of their governments as well as the extent to which citizens and the media can hold governments accountable for their actions. We use countries' percentile ranks (their ranking relative to each other) as these provide an easier way of interpreting the estimated coefficients. In our sample, Vietnam has the lowest percentile rank for "voice," 6, while Costa Rica has the highest percentile rank, 88.[3] Lastly, we include country-specific Gini coefficients on the grounds that it might be easier to identify potential beneficiaries when income or consumption differences across individuals are greater. We also include, but do not report, controls indicating whether the performance measure is based on the proportion of benefits going to the bottom quintile, the poorest decile, with the "poor" defined with reference to a poverty line or the proportion of poor found in population. Doing so takes into account confounding effects arising from the use of different measures of incidence in the studies on which this analysis is based. Standard errors are computed using the methods proposed by Huber (1967) and White (1980).[4]

The results shown in specification (1) show that as country income rises and as inequality rises, so does the targeting performance of antipoverty interventions. The former is consistent with higher-income countries having a greater capacity to design and implement finer targeting methods. The latter is consistent with countries having higher-income inequality attaching greater importance to targeting both due to the greater potential gains from targeting as well as possibly to a greater ability to differentiate among households at different parts of the income distribution. Targeting is also better in countries where government accountability is better; this is consistent with the higher level of accountability of these country governments for the effectiveness of their poverty programs.

Specification (2) looks solely at the impact of choice of targeting method. We include dummy variables for nine targeting methods described

above: three forms of individual assessment (means testing, proxy means testing, community selection of individual beneficiaries), four forms of categorical targeting (geographic, the elderly, the young, and others), and two types of selection (work requirement, community bidding for projects). The omitted category is self-selection based on consumption. We chose this as the base category for two reasons. It is often argued that this form of targeting should be seen as a transition tool while the capacity for more precise mechanisms—such as means testing—is developed.[5] Conversely, others have expressed skepticism over the ability of alternative targeting methods to reach the poor when compared to self-selection based on the consumption of food.[6] Hence, an attractive feature of this specification is that one should interpret the coefficients on these methods relative to self-selection based on consumption. Also, since the typical universal food subsidy program has neutral targeting performance (see table 3.4), the coefficients on the other targeting methods can be approximately interpreted as the proportional difference from neutral targeting. For example, a coefficient of 0.2 indicates that the method has a mean performance of about 1.2 so that, on average, poor households receive 20 percent more than their population share. 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 quintiles relative to self-selection based on consumption. Proxy means testing, community assessment, and targeting the young are also associated with improved incidence, though these are measured with larger standard errors. Targeting the elderly, other types of categorical targeting, and selection based on community bidding are not associated with better incidence relative to our base category, self-targeting based on consumption. Given the huge expansion of interest in proxy means testing in the last year or two and the limited number of programs for which results are available in our database, we performed some sensitivity analysis with respect to these programs, discussed below.

Countries with better capacity for program implementation may do better at directing benefits toward poorer members of the population either by choosing finer targeting methods or implementing their choices more effectively. As such, the associations in specification (2) may be misleading; they may merely reflect correlation between unobserved implementation capacity and observed targeting methods. We explore this possibility 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—among 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 country income is associated with a 1.8 percent increase in targeting performance. To give a sense of the magnitude of the voice effect, raising the voice rank from 37 (Pakistan's voice rank) to 67 (India's voice rank) would be

associated with about a 30 percent improvement in targeting performance.[7] However, it is possible that geographic targeting will also be more effective in countries with marked inequalities; indeed, when we drop geographic targeting from our specification (but not any other method), we obtain a parameter estimate for the log Gini coefficient almost identical to that reported in specification (1).

We performed a number of additional specific checks 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. Our basic results remain unchanged: means testing, geographic targeting, and targeting based on a work requirement raise targeting performance relative to the omitted category, self-selection based on consumption. There is no meaningful change in any of our other results.

Specification (6) takes a slightly different approach, estimating median regressions, which express differences in performance in terms of differences in medians.[8] 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

Table 3.6. Association between Targeting Performance and Number of Methods Used

	(1)	*(2)*
Number of methods used	0.165	
	(3.97)	
Used two methods		0.182
		(1.66)
Used three methods		0.300
		(2.88)
Used four or five methods		0.533
		(3.11)
Log GDP per capita	0.141	0.141
	(3.06)	(2.93)
Log voice	0.229	0.232
	(3.49)	(3.07)
Log Gini	0.280	0.278
	(1.35)	(1.31)
F-statistic	6.42	5.08
R^2	0.426	0.419
Sample size	84	84

Notes:

1. Specifications (1) and (2) contain controls, not reported, indicating whether performance measure is based on proportion of benefits going to the (a) bottom quintile, (b) poorest decile, (c) to the "poor" or (d) proportion of poor found in population. Specifications (1) and (2) estimate standard errors using the methods proposed by Huber 1967 and White 1980.

2. Absolute value of *t*-statistics in parentheses.

similar to specification (4)—which uses an identical set of regressors, sample and dependent variable—with the one exception being that the log Gini coefficient has a markedly larger coefficient.[9]

Our 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. We have not explored the association between combinations of targeting methods and targeting performance despite the fact that use of multiple methods is common. Table 3.6 addresses this issue. In addition to controls for income, voice, governance, inequality, and how the performance measure is constructed, we add to specification (1) the number of targeting methods used. The results show that use of more methods is associated with improved targeting; each additional method improves performance by 15 percent. In specification (2), we represent the number of targeting methods by a series of dummy variables. This produces a similar finding. While our sample size is too small to explore the association between specific groupings of methods and targeting performance, these results suggest that such an approach improves targeting.

Caveats and Limitations

A number of caveats and limitations should be made explicit with regard to interpreting our performance measure and the analysis based on it.

- First, our performance measure is a mishmash of various measures, although for the vast majority of the interventions (80 percent) we use the percentage of benefits accruing to either the bottom 40 percent or 20 percent of the national income distribution. This raises concerns regarding comparability. For example, one may believe that it is more difficult to target the poorest 20 percent compared to the poorest 40 percent so that programs for which we use the former may appear ineffective solely because of the performance indicator used.

   We addressed this issue in a number of ways. We calculated a second performance measure that gives, through its lexicographic ordering, priority to the proportion of resources flowing to the bottom decile, then bottom quintile, then bottom two quintiles. Doing so does not change in any meaningful way the results reported in tables 3.3 and 3.4. We also ran all regressions (reported below) using both measures of targeting performance and again found no meaningful change to our results. This is not completely surprising given that our performance measure and the alternative have correlation coefficients (in terms of levels and ranks respectively) between 0.94 and 0.97. As a further check, in the multivariate regression analysis we always include variables that control for the performance measure used.

- Second, by focusing on the percentage of benefits accruing to the bottom parts of the income distribution we are ignoring 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 those just above this income cutoff. This is partly why we give priority to the 40 percent measure of performance when constructing our performance index. It is also arguably the case that such a focus coincides more closely to the objectives of most targeted programs. In any case, the fact that our results are extremely insensitive to the ordering is at least suggestive that where between 20 percent and 40 percent one draws the cutoff point is somewhat inconsequential.
- Third, the data we have collected are only a sample of the hundreds of antipoverty interventions. Further, we could only calculate our performance indicator for two-thirds of this sample. These observations when taken together point to the possibility of "sample selection bias," that is to say, that there may be certain characteristics of these programs—for example, the fact that they were evaluated and documented—which are themselves associated with our measures of targeting performance. A good example of this possibility relates to community targeting. Our sample is only a fraction of the studies listed in Conning and Kevane (2002); it could well be the case that only successful interventions using community targeting have been well documented.
- Fourth, the number of assessments of targeting continues to increase. Our review is based on programs we had found up to approximately July 2003. As this literature grows and new findings become available, we expect that some of our quantitative findings will change as a result. Though the precise numbers change with each new program included, we expect the findings to be robust where we already have a large body of literature—it is unlikely that more observations geographic targeting's will reduce apparent usefulness or show unexpected potential for general food price subsidies. The conclusions may, however, change where the number of existing assessments is relatively small. For example, proxy means tests are in the early stages of design or implementation in many countries, but we have data from only nine programs in five countries in our database. If several of the new programs yield results as good as those shown for Chile or Mexico, then proxy means tests would join the list of methods generally showing good results.
- Fifth, some of the mis-targeting 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. Methods that rely on static indicators of living standards, such as proxy means tests, are

likely to perform less well than those that rely on self-selection when there is considerable movement of households in and out of poverty.

We remind the reader that we have been able to focus on only one narrow piece of the targeting and program choice decisions. Our performance index focuses solely on the benefit side of the equation and ignores cost; the latter may be an extremely important factor in choosing targeting methods or programs to transfer income to the poor. For example, it is often argued that well-designed public works programs can be very effective at concentrating benefits in the hands of the poor. However, the high nontransfer costs associated with such programs (including nonwage costs and forgone income) substantially reduce the cost-effectiveness of such transfer programs.

Our omission of quantitative consideration of the cost side largely reflects data restrictions. In conducting the literature review we collected the available evidence on administrative costs, hoping to comment on how these varied by method. Unfortunately, such data were scant. We have limited cost data for 32 programs, but both cost and our performance indicator for only 20 programs. 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. These numbers, which were based on budget or expenditure records for program administration, include only official costs. No attempt is made to determine how much of program benefits are siphoned off due to corruption or theft. In contrast much of the cost data on South Asian programs is constructed from knowing a total budget and having data from a survey sample on the value of benefit received by households. Through appropriate extrapolations based on the sampling weights, a figure for the total cost per dollar of benefit received is calculated. In most cases it appears that corruption and theft contribute more to total program expenses than legitimate administrative expenses, although little is said about these latter. 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, while always important, is often only one of the objectives of interventions. Therefore, the extent of tradeoffs between these other objectives and that of effective targeting needs to be taken into account when arriving at an overall evaluation of any program. However, it may be the case that these other objectives impinge as much, if not more so, on the program design, the targeting process, and the way in which the program is "sold" and delivered. Presumably most policy analysts would at least accept that monitoring the targeting performance of programs dedicated mainly to poverty alleviation is always desirable, especially in the context of developing countries where poverty is high, budgets are tight, and other policy instruments (e.g. a comprehensive income tax system) are less developed, less sophisticated, and less progressive.

Summary

Using a database of targeted antipoverty interventions found in low- and middle-income countries specially constructed for this comparative analysis, we have addressed three questions in this chapter:

- What targeting outcomes are observed?
- Are there systematic differences in targeting performance by targeting methods and other factors?
- What are the implications for such systematic differences for the design and implementation of targeted interventions?

With respect to targeting outcomes, we find that the median value of our measure of targeting performance is 1.25, so that the median program transfers 25 percent more to the target group than would be the case with a universal or random allocation. In this sense, "targeting works." However, a staggering 21 of the 85 programs for which we can build our performance measure—more than 25 percent—are regressive, with a performance index less than 1. In these cases, a random selection of beneficiaries would actually 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 and rarely exhibit good targeting performance. However, even when these are dropped from our sample, we still find that 16 percent of targeted antipoverty interventions are regressive.

Differences in targeting performance are partly explained by differences in country characteristics:

- Countries with better capacity for program implementation, as measured by GDP per capita do better at directing benefits toward poorer members of the population.
- Countries where governments are more likely to be held accountable for their behavior—where "voice" is stronger—appear to implement interventions with improved targeting performance.
- Countries where inequality is more pronounced and presumably differences in economic well-being are easier to identify demonstrate better targeting outcomes.

Differences in targeting performance also reflect choice of targeting method. 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 relative to self-selection based on consumption. Proxy means testing, community-based selection of individuals, and demographic targeting to children show good results on average but with considerable variation. Demographic targeting to the elderly, community bidding, and self-selection based on consumption show limited potential for good targeting.

However, we caution that the results on the impact of targeting method should not be interpreted as a lexicographic ranking of methods. There is

considerable variation in targeting performance when we examine experiences with specific program types and specific targeting methods. Thus, informed choices on targeting methods should be extended beyond the 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? These issues are explored in the next chapter.

Notes

1. This is available in the form of an annotated bibliography: Coady, Grosh, and Hoddinott (2003), also found in the CD-ROM at the back of this book. For each program we obtained 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.

2. The income and inequality data are taken from the World Bank's World Development Indicators database.

3. In preliminary work, we also 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.

4. We also explored whether the study had been published in a refereed journal, a book, or was unpublished work undertaken by World Bank staff, IFPRI staff, or by 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 that 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. The 85 estimates of targeting performance used here come from the following sources: refereed journals, 17; books, 17; unpublished work by World Bank staff, 37; unpublished work by IFPRI staff, 8; and unpublished work by researchers based at other institutions, 6. When we include dummy variables for type of publication as additional controls, we do not find that these are jointly statistically significant and hence do not report them here. However, we do note that 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.

5. See, for example, Pinstrup-Andersen (1988) and Alderman and Lindert (1998).

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

7. 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.

8. More precisely, we estimated a quantile regression 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.

9. As a further specification check, we re-estimated specification (4), but restricted the sample in three ways: (1) including only studies reporting the share of benefits accruing to the bottom two quintiles; (2) including only studies reporting the share of benefits accruing to the bottom two quintiles or the poorest quintile; and (3) including only studies that report the share of benefits accruing either to the bottom two quintiles, the poorest quintile or the poorest decile. When we do so, we continue to obtain comparable results to those reported in specification (4).

4

Implementing Targeting Methods

This chapter provides a qualitative assessment of specific targeting methods complementary to the quantitative results outlined in chapter 3. Our objective is to assist program managers and policy makers in making effective choices of method and implementation. Because several targeting options are available for several program types and few absolute rules exist, both the choice of targeting method and its thoughtful implementation require knowing a great deal about how each method works in general and variations on how it can work in particular circumstances or variants of implementation.

We consider means testing, proxy means testing, community-based methods, geographic targeting, demographic targeting, and self-targeting, moving from the most information-intensive to the less-information intensive methods. We do not cover several categorical methods—especially disability status, people living with HIV/AIDS, and orphans—which are of interest for a number of programs but for which we have less expertise. For each method we review international experience, how the method works, what determines how well it works, what its costs are likely to be, and what are the appropriate circumstances for its use. Table 4.1 provides a brief overview of the advantages, limitations, and suitable circumstances for each of these targeting methods. Some readers may find that this will help them focus on those methods suitable to their circumstances. Readers may find it useful to refer back to table 3.2 to see the pattern of use of the various targeting methods across regions and program types.

Means Testing

Means testing is a form of individual assessment that compares resources such as income under the command of an individual or household with some threshold or cutoff. In our sample, 34 programs used some form of means test. Although these are found in all parts of the developing world,

Table 4.1.   Comparing Targeting Methods: Overview

---

1. Individual/household assessment
a) Means testing

Brief description	An official (usually a government employee) directly assesses, household by household or individual by individual, whether the applicant is eligible for the program. It has three main variants: those with third-party verification of income, those in which the applicant provides documents to verify income or related welfare indicators, and those in which a simple interview is used to collect information.
Advantages	In the best of cases, very accurate
Limitations	• requires high levels of literacy and documentation of economic transactions, preferably of income • administratively demanding where there are meaningful attempts at verification • most likely to induce work disincentives
Appropriate circumstances	• where declared income is verifiable or some form of self-selection limits applications by nontarget groups • where administrative capacity is high • where benefits to recipients are large enough to justify costs of administering means test

b) Proxy means tests

Brief description	A "score" for each household is calculated based on a small number of easily observable characteristics and a weight (ideally obtained from factor or regression analysis of household data). Eligibility is determined by comparing the score against a predetermined cutoff.
Advantages	• is verifiable, may allay concerns over politicization or randomness of benefit assignment • uses readily observable household characteristics • is less likely than means test to affect work effort
Limitations	• may seem mysterious or arbitrary to some • requires large body of literate and probably computer-trained staff, moderate-to-high levels of information and technology • inherent inaccuracies at household level, although good on average • insensitive to quick changes in welfare, as in a crisis or in some transition countries

Table 4.1.   (*continued*)

Appropriate circumstances	• reasonably high administrative capacity • programs meant to address chronic poverty in stable situations • where applicable to a large program or to several programs so as to maximize return for fixed overhead

c) Community targeting

Brief description	A community leader or group of community members whose principal functions in the community are not related to the transfer program decides who in the community should receive benefits
Advantages	• takes advantage of local information on individual circumstances • allows for local definition of need and welfare • transfers costs of identifying beneficiaries from intervention to community (this can also be seen as a limitation)
Limitations	• local actors have other incentives besides good targeting of the program • may lower authority or cohesion of local actors • may continue or exacerbate patterns of social exclusion • if local definitions of welfare are used, evaluation is more difficult and ambiguous
Appropriate circumstances	• where local communities are clearly defined and cohesive • for programs that propose to include a small portion of the population • for temporary or low benefit programs that cannot support an administrative structure of their own

2. Categorical targeting

a) Geographical targeting

Brief description	Eligibility for benefits is determined, at least partly, by location of residence. This method uses existing information such as surveys of basic needs or poverty maps.
Advantages	• administratively simple • no labor disincentive • unlikely to create stigma effects • easy to combine with other methods

(table continues on following page)

Table 4.1.   (*continued*)

Limitations	• depends critically on the accuracy of information • performs poorly where poverty is not spatially concentrated • can be politically controversial
Appropriate circumstances	• where considerable variations exist in living standards across regions • where administrative capacity is sufficiently limited so as to preclude use of individual/household assessment • where delivery of intervention will use a fixed site such as a school, clinic, or ration shop

b) Demographic targeting

Brief description	Eligibility is determined by age, gender, or some other demographic characteristic
Advantages	• administratively simple • low stigma • often politically popular
Limitations	• inaccurate where demographic characteristics are poor correlates of poverty; current research suggests that observed correlations are sensitive to assumptions made about household scale economies and adult equivalences
Appropriate circumstances	• where registration of vital statistics or other demographic characteristics is extensive • where a low-cost targeting method is required

3. Self-targeting

Brief description	A program, good, or service that is open to all but designed in such a way that takeup for it will be much higher among the poor than the nonpoor
Advantages	• administrative costs of targeting likely to be low • unlikely to induce labor disincentives
Limitations	• imposes costs, sometimes a quite substantial cost on the recipient, which lowers the net value of the benefit • stigma may be considerable • may be difficult to find a means of delivering a large benefit
Appropriate circumstances	• countries with very low administrative capacity • settings where individuals are moving rapidly in and out of poverty • where a wage or consumption patterns separates poor from nonpoor (e.g., consumption of inferior goods)

they are relatively more common in Latin America and East and Central Europe. The majority of means-tested programs (24 of 34) involve cash transfers.

How Does Means Testing Work?

The application of means testing requires the collection of information on the total income of households or individuals. This requires either a visit by a designated officer to the home of the potential beneficiary (home visit) or a visit by the beneficiary to a program office (office visit). Verification can involve simply an interview with the information provided taken at face value, supporting documentation provided by the potential beneficiary (individual documentation), and/or supporting documentation provided by a third party (third-party verification). Consequently, within the rubric of means testing are large variations in terms of the complexity of the means test, the level, and the distribution (between beneficiaries and implementing agencies) of costs and the accuracy of targeting achieved.

The "gold standard" means tests involves verifying information provided by the applicant with information from third parties often income- or property-tax records from the public sector, sometimes wage information from employers or financial information from banks. This approach is appropriate only where the targeted population of low-income families is largely employed in the formal sector and/or participates in a well-functioning income tax system. Although these conditions may exist in many OECD and some transition economies, neither is commonly met in poor countries; they may not be met for the poor populations of middle-income countries.

Alternatively, applicants may be asked to supply verification of some aspects of their welfare. For example, instead of requiring employers to supply information to the welfare agency in a uniform format and time schedule, applicants bring copies of their pay stubs to a social welfare office. The range of information may extend beyond income to expenditures assumed to be correlated with income and able to be readily documented, such as telephone bills or electricity usage. This will work poorly for individuals whose economic lives are largely undocumented, such as workers in the informal sector who do not have access to utilities. This procedure shifts the costs of collecting the information to applicants in a way that can entail significant transactions costs for them.

Lastly, means tests may be based on an interview with at best informal verification. In some cases the interview may be based in the household on a home visit, which allows a qualitative glimpse of the household's living standard. In the (now-defunct) Jamaican food stamp program, the form on which the applicant reported income had to be signed by a community representative such as a minister of religion or justice of the peace (Grosh 1992). In other cases the interview may be a simple in-office affair where there is nothing more than the consistency of the interviewee's story to verify or rebut his claims.[1]

What Determines How Well Means Testing Works?

Six out of the top 10 interventions—social assistance programs in Estonia, Dominica, Hungary, Yemen, Poland, and Romania as ranked by our performance measure used means testing, as did several with performance indicators of about 1—child allowances in Uzbekistan and Poland and social assistance in Latvia. The effectiveness of targeting based on means testing depends crucially on the ability to collect reliable information on total income at a reasonable cost. This requires either access to formal employment or tax records to validate reported incomes, or the collection of first-hand information on household incomes, wealth, and/or consumption. Such characteristics are typically associated with higher-income countries; in our data set, evidence suggests that, conditional on having chosen some form of means testing, targeting performance rises with country-income levels.[2]

What Are the Costs Associated with Means Testing?

Means testing appears to carry higher costs associated with the collection and verification of information than other methods, although hard evidence is scanty.[3] These can be incurred as program administrative costs or private costs by beneficiaries. Costs associated with the collection of comprehensive information on household incomes or consumption via a home visit can cost several dollars per applicant household. On the other hand, office visits may involve applicants incurring substantial time and money costs associated with acquiring the relevant documentary evidence and (possibly multiple) trips to program offices. These costs can be reduced by ensuring that potential beneficiaries have easy access to offices, that they are well informed about eligibility rules and the documentation required, and that program offices have the capacity to deal efficiently with applications. Using means testing in combination with other targeting methods can also reduce costs, for example, by avoiding the costs associated with collection of information from households or individuals that have little chance of being eligible to receive program benefits.

The social stigma costs associated with means testing, either related to the requirement that beneficiaries publicly identify themselves as poor or undertake actions in order to qualify, also means that use of means testing can be a politically sensitive issue. Some argue that current income is not a good proxy for access to opportunity and that it is the latter that requires public action. Others argue that the information requirements for means testing represent unacceptable intrusion into the lives of citizens. However, where means testing is seen to prevent the "undeserving" from receiving benefits, such costs may appear attractive, especially by those who perceive themselves as financing these programs.

Where means testing is based on income, it may create a disincentive to earn own income and thereby to work. The disincentive arises from the fact

that if benefits are withdrawn as income rises as a result of working, the net gain in income to the household is reduced. If withdrawals are high—as was the case in the United Kingdom until the early 1990s—households find themselves in a "poverty trap" (Dilnot and Stark 1989). Sometimes there are also other requirements, often to demonstrate that the recipient is doing something "worthy" such as working, in training, actively searching for work, caring for young children or sick or disabled household members. Though the motive for the requirements is clear, the monitoring can be quite difficult. It requires both extensive and continuous information that is hard to verify or assess. What constitutes looking hard enough for a job? How young must children be to excuse a parent from working? Such finetuning of incentives systems is uncommon in low and middle-income countries, though recently several new programs have been introduced that tie benefits to the provision of schooling and health interventions for children.

Appropriate Circumstances

Means testing is appropriate for countries with high administrative capacity and well-documented economic transactions that allow third party or applicant supplied verification of income. Because it is demanding, means tests will be most applicable to programs that provide large benefits. Indeed, for a program with very large benefits, it is very unlikely that the simplest methods—demographic, geographic or self-targeting alone will be sufficient.

Means testing is less appropriate for countries with more limited administrative capacities or as part of a rapid response to a shock. Among countries with per capita incomes less than 1000 PPP USD, one of the poor performing interventions was the means-tested GAPVU program, a cash transfer program operating in urban Mozambique. Implementation was highly uneven across urban localities, with widespread corruption (via the inclusion of fictitious names on beneficiary lists) co-existing with such rigorous application of means testing that at one point only 0.6 percent of the target population in Maputo was receiving benefits (Government of Mozambique, 1998). The performance indicator for GAPVU was 1.05, implying that the benefits were only about as well distributed as they would be with random allocation. GAPVU eventually was closed down before being relaunched as the National Institute for Social Action (INAS).

Proxy Means Testing

Because of the difficulties associated with collecting and verifying detailed information on household income or consumption levels in many developing countries, governments and development practitioners have tried to identify alternatives that overcome these difficulties. In this section we examine the use and application of one such approach, namely, proxy means testing. In our sample we have nine programs in five countries that

use proxy means tests, three from LAC and one each in ECA and East Asia. Proxy means testing is a relatively new tool in the targeting toolbox. Chile was the first country to use this approach when it introduced its *Ficha CAS* program in 1980. Since then, the tool has been monitored and its implementation and use refined over the years (Larrañaga 2003; Wodon and Clert 2000; Sancho 1992; Racynzski 1991; Castaneda 1990). The approach has spread elsewhere in Latin America, with large proxy means testing systems having been set up and evaluated in Colombia (first for subsidized health insurance and then for several programs including conditional cash transfers, workfare, and scholarships for vocational training; Castañeda 2003) and Mexico (conditional cash transfers, Skoufias and others 2001; Coady 2001). A number of other Latin American nations (Argentina, Costa Rica, Ecuador, Jamaica, Honduras, and Nicaragua) are developing proxy means testing systems and results from these are not yet available. However, the experience is no longer exclusively Latin American. Armenia has used a proxy means test since 1994 for humanitarian assistance and cash transfers (World Bank 1999, 2003), and Indonesia has used one for targeting its subsidized rice rations (Sumarto and others 2000). Turkey introduced such a system in 2002 as part of a response to the financial crisis (Ayala 2003), and other countries have done some piloting without fully setting them up (e.g., Russia, Egypt, and Zimbabwe, and as of our press date, Sri Lanka).

How Does Proxy Means Testing Work?

Proxy means tests use a relatively small number of household characteristics to calculate a score that indicates the household's economic welfare. This score is used to determine eligibility for receipt of program benefits and possibly also the level of benefits.

    The first step in designing a proxy means test is to select a few variables that are well correlated with poverty and have three characteristics:

- Variables should be few enough that it is feasible to apply the proxy means test to the significant share of the population that may apply for the program, possibly as much as one third.
- Variables selected must be easy to measure or observe.
- Variables should be relatively difficult for the household to manipulate just to get into a program. These variables are typically drawn from the data sets of detailed household surveys, for example, a household budget survey or a multitopic survey that include detailed information on consumption, employment, education, health, housing, and family structure.

In most cases the variables selected include indicators of the location of the family's home, the quality of its dwelling, its ownership of durable goods, the demographic structure of the household, labor force status, occupation or sector of work for the adults, and sometimes partial measures of income. The number of variables used is often on the order of two dozen.

Once the variables have been chosen, statistical methods are used to associate a weight with each variable. One common approach is to use regression analysis and regress total income or consumption of the household on the selected variables. Often these regressions are run separately by region (e.g., by province or rural/urban) so that variable weights differ across regions. This procedure is often iterative in that the variables initially selected are chosen on the basis of a more comprehensive statistical analysis that evaluates their predictive power, that is, how closely they are correlated with household income. Sometimes the weights are rounded to facilitate calculation of scores in the field. Analogous procedures are used when principal components analysis is used rather than regression analysis.

In all countries, the basic design of the system (e.g., choice of variables, design of forms, operational manuals) is handled by a single agency in the national government, usually in the ministry of planning or welfare. The staff power to fill out registration forms and do data entry for the hundreds of thousands or millions of applications has been placed in different agencies in different countries. In Chile, the municipality administers the test (Larrañaga 2003). In Armenia, this function is performed by the staff of the social welfare ministry (World Bank 1999). In Mexico, a special temporary cadre of survey workers administered the first wave of registrations in each area, but eventually these functions were carried out by personnel employed and organized directly by program officials (Skoufias and others 2001). This intensive first wave of outreach at the inception of a proxy means test is a common and recommended technique to set up the system.

Well-implemented systems include periodic recertification of beneficiaries to ensure that those whose welfare has improved are removed from the system. There may also be other program specific limitations on the duration of benefits or exit rules related to age or other criteria. Recertification may also present the opportunity to update the scoring system. However, although recertification may improve the accuracy of targeting, it also involves additional costs. We do not know of any country that recertifies more often than once a year; some recertify on a three-year cycle. All too often, there is no regular cycle but long periods without any perhaps followed by ad hoc decisions to recertify everyone.

In some cases the scoring system is made public. In theory making it public embodies the ideal of transparency. Individual applicants or their advocates are empowered to double check whether their application has been handled fairly, and civil society is fully informed on the policy choice and can debate the appropriateness of the formula. In Armenia, the formula has been public, usually posted on the walls of the welfare offices. However, qualitative fieldwork (Gomart 1998a) showed that even in a society with universal literacy and exceptionally high levels of tertiary and technical training, the scoring system was not well understood, especially since the formula was presented in its mathematical form with specialized notation. In Chile, the formula was publicly known for some years, with

the points assigned to each answer printed on the application card so that the interviewer could calculate the applicant's score on the spot. In 1987, however, the government revised the formula based on a new data set and decided to keep it confidential, partly because its complexity made it difficult to communicate clearly and partly due to concerns that applicants would know too clearly how to bribe social workers if it were publicly known (Sancho 1992). We are not aware, however, of evidence as to whether this concern was based on real incidents or just the hypothetical possibility.

A key feature of proxy means tests, which some see as an advantage and others see as a disadvantage, is the formulaic nature of its calculation of need. It has the merit of making replicable judgments using consistent and visible criteria. Because the information used is fairly straightforward to collect and simple to interpret, a well-instituted proxy means test should guarantee "horizontal" equity. This equity means—that the same or similar households (at least in terms of the variables chosen) will receive the same treatment or decision even if evaluated by different staff members or by the same staff member on different days or in different moods. Moreover, the training can be fairly simple and need not include a great deal of social work content. Appealingly, the degree to which this holds can be monitored quantitatively. For a government fighting actual or perceived corruption or politicization of program entry, this can be important indeed. The flip side of the formula is that it is rigid—it does not take into account special circumstances pertinent to the household. Additionally, there is built-in error: the formula is designed to be right on average but will not correctly categorize every household.

What Determines How Well Proxy Means Testing Works?

How well proxy means testing systems work depends on a number of factors:

First, one needs to be able to identify variables that exist in the surveys that are highly correlated with household income, that can be easily observed, and that cannot be easily manipulated by households in an attempt to get into the program. In many cases the formulae at best only explains about half of the variation in consumption among households and often substantially less. For example, in the case of Egypt's food rationing system, the variables used explained only 43 percent of the total variation in consumption, compared to the most comprehensive model that explained 62 percent of the variation (Ahmed and Bouis 2001). Moreover, the prediction for each individual household, even if unbiased, has a large standard error, a fact not really taken into account in current practice with proxy means tests. In Armenia, various candidate formulae explained only about a quarter of the total variation (Grosh and Glinskaya 1998).

Basing the choice of variables on analysis of existing household surveys will also rule out many variables that astute field observers in the country would suggest. In Armenia, for example, social workers visiting households make note of the presence or absence of fast food wrappers in the rubbish bin, and whether the household's clothing is apparently from before or after independence and the move to a free market (Gomart, 1998a and b). These indicators are probably very highly correlated with household income, but such questions are not on the standard survey on which proxy means tests are built so it is not possible to test this hypothesis or to assign such questions weights in the proxy means test formula. Though some special questions might be added to the survey, the problem remains of making them relatively easy for the social worker to observe and hard for the household to falsely manipulate.

Second, one needs to have an information system that is capable of reliably gathering this information from many households, in particular from the poorest households. A body of literate personnel is required to help fill out the forms for each applicant. This is probably the requirement that most limits the countries in which proxy means tests are used. In addition, a small central team with sophisticated analytic skills is needed to set up the scoring system and supervise its implementation. The information technology used can include sophisticated networks of computers calculating scores during the client interview and sharing information with all levels of the program in real time, or scores can be calculated by hand and stand alone computers used only for ex-post record keeping.

As with means testing, administrative arrangements associated with collecting and verifying information are vital to ensuring low errors of exclusion. No matter how well or badly the statistical formula works, if the poor do not register for the program it will have high exclusion errors. In our experience, this is an area that often fails to get the attention it deserves. Some governments have devoted much more effort and ingenuity to the statistical issues of finetuning the formula and administrative attention to winnowing out unqualified participants than to lowering errors of exclusion through a good outreach program.

Information about the benefits and procedures must be made to reach the participants in appropriate languages through appropriate channels. Postings and pamphlets in markets, village shops, places of worship, beer halls, schools and clinics, and conducting briefings of traditional leaders, community organizers, and NGOs will usually be more effective than a mass media campaign in reaching the poor. Transactions costs to register should be low, ideally involving a single visit between the household and registering officer, and this possibly in the applicant's home or involving only one trip and queue in the appropriate office (such as the department of social welfare or public health clinic). Social workers should have the time and transport allowance to allow them to do extension work to help reach those who find it

unusually difficult to register—homebound or infirm persons and those who live in areas remote from registration offices. In the most thorough cases, this can involve a complete census in poor areas upon program inception followed by periodic outreach. Requirements that applicants present paper-work (such as national identity papers, marriage or birth certificates for household members, or tax papers) should be reviewed critically as each of these potentially raise errors of exclusion. Where they are required, a system to ensure that such paperwork can be obtained quickly and at low transaction costs is desirable. International experience includes cases where, in contrast, it may take multiple trips to various offices to determine what paper work is required and to get it in order. In addition to time and transport costs of this, there may be official fees to get some paper work complete, and under-the-table payments for others, all of which tend to raise errors of exclusion.

A home visit as part of the application procedure can be a means by which administrative arrangements reduce errors of inclusion (this is the main job of the formula itself). Public information campaigns that empha-size that the goal of the program is to serve the poor may create a stigma that is also helpful in reducing leakage. Most of the Latin American programs re-quire an interviewer or social worker to visit the household to fill out the proxy means testing form (see Coady, Grosh, and Hoddinott 2003). This provides a means of verifying that the information on the form is true. It may also induce a certain amount of self-selection as those who live clearly too well to qualify may be discouraged from applying. However, the home visit takes resources. In contrast, in Armenia home visits were not a routine part of the application procedure; rather, the applicants were required to bring documentation of their household status into the welfare office, in-cluding registry papers to prove family composition, proof of residency in-formation related to earnings or unemployment status and receipt of other government programs, and a medical certificate of disability if applicable (World Bank 1999). Not surprisingly, the Armenian program's early results were regressive, though the reformed program shows much better out-comes (World Bank 2003).[4]

There are certainly limitations to verification through presentation of documents. Obviously the range of variables for which documentation can be expected is limited. A first limitation has to do with the general levels of literacy and documentation in the country. In low-income countries, these may be quite low indeed. Even in middle-income countries, paper docu-mentation is often better for ruling out the upper end of the income distrib-ution than for certifying poverty.[5] It is the nonpoor who are likely to have formal pay stubs, utility bills, or automobile tax records. Thus a program can rule out those who present such things, but it cannot tell whether those who do not present them are poor or merely concealing information. In a limited number of cases it may be possible to link the proxy means testing agency's records with those of the government agencies that monitor these other forms of paper work. This may, however, encounter formidable barriers in terms of compatibility of records.

What Are the Costs Associated with Proxy Means Testing?

The cost issues related to proxy means targeting are very similar to those for means testing. The collection and verification of information involves costs, and whether these are incurred as private costs by applicants or are financed from the program budget, will depend on whether information collection involves office or home visits. An extra cost is that personnel need to be trained to process the information using the scoring system and the levels of sophistication that this involves can vary widely. There are also the costs associated with developing the scoring system. The additional costs involved are lower if household survey data already exist so that a special data collection round is not required solely for the program.

In absolute terms we have data for three programs. In Armenia the administrative cost of the proxy means testing was estimated at about $1 per applicant (World Bank 1999), in Chile at about $5 per applicant in 1992 (Sancho 1992) and $10 in 2003 (Larrañaga 2003), and in Mexico $12 (Coady, Perez, and Vera-Llamas 2000). The Armenian costs are low both because local salaries are quite low and because of features of the proxy means test's implementation—in-office interviews and very limited outreach to help register those who face barriers in coming into the offices. Chile and Mexico have higher staff salaries, but both also have interviews conducted in the home and much more extensive outreach. Although there is a change in the absolute nominal Chilean unit cost figure between 1992 and 2003, both sources estimate that the Ficha CAS's administrative costs amount to about 1.3 percent of the expenditures targeted through the system. It was calculated for a mature system, with development costs amortized and periodic outreach for only a portion of registrations. In Mexico the costs were calculated when the system was newly introduced and all registrations were done through extensive outreach in targeted areas. Moreover, the Chilean form is a short two pages collecting only the information used in the formula. The Mexican form is 20 pages long and resembles more a survey questionnaire than a program application form. In addition to collecting information used in calculating the formula itself, it also obtains further data for subsequent poverty and impact analysis.

Targeting several programs using the system spreads the administrative cost over a larger benefit base and renders it more cost effective. Colombia's SISBEN, for example, certifies families as eligible for subsidized health insurance, for the system of conditional cash transfers, for workfare, and for scholarships for vocational training. It is possible for different programs to use different cutoff scores.

Combining targeting methods can also reduce costs. For example, Mexico's PROGRESA program chooses the areas where the program will operate from a detailed poverty map. Then teams of field workers canvas households in only those districts collecting information to be used in the proxy means test (Skoufias and others 2001). It is also certainly possible to use the proxy means test in conjunction with demographic categories, for

example, for a child allowance program, as done in the Chilean unified family subsidy program (Sancho 1992).

The potential social and political costs can vary widely across countries and may depend on how the introduction of the program is managed. The issue of using a formula often elicits a strong emotional reaction in many people. It appeals to some as "scientific," "rational," or a safeguard against corruption or prejudice. To others it will seem a mysterious black box ("the computer decided"), cold-hearted, or a coverup for what the government or social worker wanted to do (Gomart 1998a; Clert and Wodon 2000). The inclusion of certain variables may also be politically sensitive. For example, in the Egyptian system, binary variables indicating location at the governorate level were removed from the scoring system because they were seen as being too politically sensitive, in spite of the fact that they were very powerful predictors of poverty.

Appropriate Circumstances

As described above proxy means testing is administratively demanding, especially in its requirement for a large body of literate (and in some variants computer literate) personnel to carry out the registrations. There is usually also at least a moderately sophisticated computer network associated with the system. Most of the countries using proxy means tests thus far are middle income.

Proxy means tests can be (and have been) used for various cash transfers, subsidized food rations, health insurance, workfare, scholarships for vocational training, and housing and utility subsidies. Chile briefly used them for school feeding and fee setting in heath care but eventually stopped. In general they are best used for programs where it is not a logistical problem to delay a benefit until an applicant is certified. A cash transfer program, for example, is not inconvenienced if a given applicant has not yet had the paperwork completed. In contrast, it can be a problem for a hospital trying to use a proxy means test to determine whether to grant access or what fee to charge for, say, emergency room visits. Inevitably some patients will turn up, claiming to be needy, but they will be without the proper certification in the means testing system. If proxy means tests are to be used for access to such services, then one or more of the following arrangements have to be allowed: the hospital has to be able to administer the proxy means test itself (even without a home visit), it has to be able to admit the patient and waive the fees without the certification, or it has to be able to turn away the patient.

The indicators used in proxy means tests are usually rather static. Proxy means tests are, therefore, best used to reach the chronically poor rather than to distinguish those suffering most acutely from a newly confronted crisis (be it economy wide or household specific). This is partly because household surveys typically do not collect many of the more quickly changing variables that indicate household welfare, and partly because the subset of variables selected from the survey for use in the proxy means test are meant

to be those that are fairly easy to observe and not much subject to manipulation. Thus even if a survey asks questions along the lines of "how many meals did you eat yesterday?"—which can change from season to season or when a family confronts a crisis—those sorts of questions are rarely included in the variables selected for the proxy means test. This is surely something of a frustration to the policy makers searching (sometimes desperately) for a tool to ameliorate crises such as those suffered from East Asia's financial crash and its knockon effects in Russia and Latin America. However, the flaw must be put in context. Even if the formula did incorporate more quickly changing indicators, the ideal sorts of beneficiary lists that evolve rapidly as the crises wax and wane would require recertifying households every few months. In most cases, this is too demanding an administrative burden to be practical.

Community-Based Targeting

Of the 122 programs in our sample, 14 use community-based methods. They are spread fairly evenly across continents and over a large range of program types. We suspect that we have under-represented the use of this method, especially in low-income countries and that those we have may perform unusually well. We know, for example, that several of the programs that use community targeting use other methods as well that may be responsible for a good deal of their power. Mexico's PROGRESA, for example, first uses a poverty map to select poor areas in which to work, then a proxy means test to screen out non-poor residents of those areas, then the community-based targeting to fine-tune the proposed beneficiary lists (Skoufias, Davis, and de la Vega 2001).

How Does Community-Based Targeting Work?

In this case a group of community members or a community leader whose principal functions in the community are not related to the transfer program will decide who in the community should benefit and who should not. School officials or the parent-teacher association may determine entry to a school-linked program. A group of village elders may determine who receives grain provided for drought relief. Special committees composed of community members or a mix of community members and local officials may be specially formed to determine eligibility for a program.

There is little documented evidence on community-based targeting as compared to other methods. Conning and Kevane (2002) provide the most comprehensive summary and conclude that we know little. The ideas contained in this section are therefore rather more speculative than in most parts of the paper.

What Determines How Well Community-Based Methods Work?

Perhaps the most persuasive rationale for community-based targeting is that local actors have more information available to them or at lower costs

than would officials from a welfare agency. Ministry social workers may know only what applicants put on the application form in the office or what can be observed on a brief visit to the applicant's home, which will be costly to the agency in terms of time and transport and during which the applicant may try to conceal information. Community members may know much more about the applicant's resources and needs without even having to gather data beyond what they see in the course of daily transactions. Complementary to the argument based on information is one based on social capital. Where community members are enmeshed in multiple relationships, they may have less incentive or face greater repercussions if they hide or misuse information.

A second reason to choose community-based targeting is to allow communities to define need as they find most appropriate. National formulae, as discussed in the section on proxy means tests, may not work very well in all places. Access to networked utilities may be a key marker of welfare in urban areas, but it may not be significant in rural areas where even the wealthy are not served by utility networks. Additionally, the differing definitions can be routed in deeper cultural norms of what constitutes well-being.[6]

However, there are circumstances in which community-based targeting works less well. By definition, it gives decision-making power to a person or group whose principal responsibility is not related to the transfer program. That agent may have incentives other than, or at least additional to, providing the best targeting.[7] The community may wish to avoid dissent, the nurse or teacher may want to build a relationship with the parents of the children she serves, the municipal worker may want to get as many resources for the municipality from the national budget as possible, the mayor may wish to be re-elected, the local religious leader may wish to reinforce the social norms of his religion. The multiple interests of the community actors imply that even if they have excellent information on who is needy, they may not use that information in the way that the central welfare agency funding the program might prefer.

The duty of serving as targeting agent may also impinge upon the actors' abilities to perform their original and primary functions in the community. If the teacher says that the family should not qualify for free textbooks, will the parents listen as attentively when she says that their child behaves badly in class? If a neighbor says that a community member should not receive a benefit, will the denied applicant be as likely to help him out when he is on hard times? Such problems indicate that there may be costs to community targeting that are not easily identified or quantified.

Community-based targeting may also perpetuate local power structures. This may mean that local leaders will direct benefits to their families or client networks. It may mean that patterns of exclusion of certain groups are reinforced: minority ethnicities, those with HIV/AIDS, disabled persons, or unwed mothers. Conning and Kevane (2002) report that in Slovak villages Roma were excluded, and in Mexican *ejidos* Chola-speakers were excluded by non-Chola-speakers.

Community-based targeting presumes that the community agrees with the center that some of its members should receive transfers and others not. Sometimes this is not the case. In Indonesia, for example, field work showed that in some villages the elders, rather than specify which households should receive the subsidized rice ration, divided the village's quota among all members.

Finally, the very notion of community is problematic. The idea of community-based targeting is easiest to apply in small communities where it is clear who is a member, where everyone belongs to one and only one community, and where members know each other. This model may hold reasonably well for rural areas where households are clustered in villages or for urban areas with well-defined and tightly-knit neighborhoods. However, a great deal of the world lives in areas where the concept of a community is more amorphous. Some rural areas have dispersed settlement patterns. In many urban areas, close neighbors may not know each other well and boundaries between "communities" may be very blurry; indeed pertinent communities may not be geographic but defined by social criteria such as place of origin, affiliations with workplace or school, or affiliations with religious or occupational communities.

What Are the Costs Associated with Community-Based Targeting?

Community-based targeting usually has, or at least appears to have, low administrative costs. Often community actors involved in community-based targeting are not paid for their time or expertise. They may not receive assistance in covering the travel costs of doing their jobs thoroughly or for stationery, communications, and the like. These costs exist, but they may be passed to the community or agency for which the community leader works rather than appear on the welfare program's budget.

An important administrative task and cost of community-based methods (one that is not always given enough emphasis) is that involved in working with many thousands of individuals who are not dedicated to the program. Few or none will start out knowledgeable about the program and its workings; some may not be sympathetic. They may not have the same incentives to learn about or follow program rules that a dedicated staff member would and will have many constraints on their time that a dedicated staff member would not. Illiteracy and lack of fluency in the national language will be greater problems than when dealing with a cadre of civil servants. There will be no preestablished channel of communications. Thus the job of training and motivating actors will be even greater than if they were staff members and, if community-based targeting is to work, this must receive due attention.

Appropriate Circumstances

Community-based targeting often seems to be chosen where other options just do not seem feasible. Means testing or proxy means testing may be too

demanding administratively. Geographic and demographic targeting may result in too large a group of beneficiaries to be served by the budget or groups that are not very poor on average. Self-targeting may not be very applicable to the type of program being designed.

In what circumstances will community-based targeting probably work best? We have already indicated that one premise of the method is a well-defined community with good social consensus. We suspect that the method will work best when the community is asked to choose only a few members for program receipt, say five or ten percent, rather than when closer to half the community might benefit. For one thing, the method is most often used in very poor countries where the difference in incomes between the moderately poor and poorer may be in terms of nickels a day and thus hard for even those in the community to observe and agree on. When benefits can go to only a very few, most community members will stand little chance of getting them and may thus act in a more disinterested or altruistic way in helping to make or allowing to stand the decisions about who should benefit.

Community-based targeting will probably work best where there is a hard budget constraint given to the community—so many rations or scholarships or fee waivers to grant. Certainly this will help with overall program planning. Otherwise, the community will have no incentive to ration and may just declare everyone, or at least a large share of the population, poor. Such a hard budget constraint will usually be accompanied by some geographic targeting to allocate the budget to the community.

Community-based targeting may also work relatively well in conjunction with demographic targeting. It may, for example, be a suitable tool for choosing among widows if widows as a whole are too numerous for a program to help, or if widowhood is not highly correlated with poverty on average because the distribution of widow's welfare is bimodal. The community may be very well able to distinguish between types of widows—the matriarch of a family with several adult income earners and who is coresident or supported by one or more of them, versus the very old, frail woman living alone who receives no support from family members.

We also suspect that the problems with undermining the principal roles of those involved in community-based targeting decisions may be minimized if the program is temporary or gives relatively small benefits.

Geographical Targeting

Geographic targeting is popular; indeed in our sample of programs, it was the most common method, used in 52 out of 122 programs. It was the only method used by all program types, but it was relatively more important for public works and social funds and least important for universal food subsidies. It was used in all regions and countries in all income groups, but especially common LAC, SEA, and East Asia. In 51 of the 52 applications, it was used with at least one other method; only in the targeting of Colombian utility subsidies was it the sole targeting device. In this section, we briefly discuss issues related to the implementation of geographic targeting. For a

more detailed and comprehensive discussion of geographic targeting, see Bigman and Fofack (2000).

How Does Geographic Targeting Work?

Geographic targeting involves allocating resources to geographic areas using information that is thought to be a good indicator of the extent of poverty in these areas. For this reason, this approach is now commonly referred to as "poverty mapping." The areas used may be political subdivisions of the country (states or counties), or they may be the catchments of specific service providers such as clinics or schools. There are a number of approaches to poverty mapping; these differ essentially according to the amount of information used and how it is combined to evaluate the extent of poverty in each area. The main approaches are illustrated below moving from the least to most data intensive.

There are also issues of conceptual notions of poverty. Economists traditionally have focused on income or consumption as a measure of welfare, the latter typically being interpreted as a better proxy for "permanent" or lifetime income. In contrast, much of the history of poverty mapping has used a "basic needs" approach with poverty defined in terms of access to basic services. The indicators used are often interpreted using one of these approaches.

The simplest form of geographic targeting involves the use of a single variable such as nutritional status, which is often used when the program being targeted has a strong nutritional objective. In the Honduran cash transfer program PRAF, municipalities were chosen based on nutritional levels, using as the criterion the measure of height-for-age z-scores available from the 1997 height census for first-grade children (Morris 2001). In an early stage of the Venezuelan *Beca Alimentaria*, the "mapping" was less formal, based principally on the judgment of program officials about which schools served poor areas. That judgment was based, of course, on having been exposed to various sorts of poverty maps and to field conditions, but it was not formally quantified nor necessarily reliant on the same factors in different areas of the country. In such cases the underlying concept of poverty is obviously very subjective and not especially explicit or transparent.

A more sophisticated version of poverty mapping uses principal component analysis to calculate the summary poverty indicator. It is often interpreted as a reflection of basic needs status or capabilities, and the choice of variables is largely guided by a combination of philosophy and data availability. For example, in its initial geographic targeting stage for the PROGRESA program, Mexico used seven variables from three data sources—the XI General Population and Housing Census of 1990, the Population and Housing Count 1995, and the Geographic Integration Catalog, all constructed and carried out by the National Institute of Statistics, Geography and Information (INEGI). Seven variables were used:

- over-15 illiteracy rate
- percentage of households with access to running water
- percentage of households without access to drainage

- percentage of households without access to electricity
- average occupants per room
- percentage of houses with earth floors
- percentage of labor force in agriculture.

Principle component analysis was used to calculate a Basic Index of Marginalization for each locality in the census based on these seven variables, and this indicator was usually interpreted as a reflection of basic needs (Skoufas and others 2001).

More recent applications of poverty mapping have used consumption as the welfare or poverty indicator. The idea is to combine the information on consumption from nationally representative household-level income and expenditure surveys with the information available in a national census for all households in the country (Hentschel and others 2000). Using variables common to both data sets, a model of household consumption is estimated using the expenditure surveys. The estimated coefficients are then used to predict consumption for every household in the national census. Poverty rates are calculated for each geographic unit based on some poverty line and poverty index (e.g., the poverty headcount, the poverty gap, or the severity of poverty). This approach is essentially a more sophisticated form of proxy means tests applied to geographic units.

A special characteristic of this new approach is that greater recognition is given to the problems associated with the precision of predictions (Elbers, Lanjouw, and Lanjouw, forthcoming). As the degree of disaggregation increases, so do standard errors so that precision declines. For example, in the Ecuadorian example used in Hentschel and others (2000), the administrative level below the canton (of which there are about 400) is the parish (of which there are about 1,000). The standard errors become high enough to compromise comparisons for small parishes (below about 500 households). This Ecuador example served as pioneer and pilot for this approach. Since then similar applications are finished, underway, or about to start in Brazil, Guatemala, Mexico, Nicaragua, Panama, Kenya, Madagascar, Malawi, Mozambique, South Africa, Tanzania, Uganda, China (one province), Indonesia (one province), Thailand, and Vietnam (Lanjouw 2002).

Because construction of such poverty maps requires having access to census data at the household level, this implies either that the work be done within the census office, or that full unit record data for the census be released to external agents, which may break precedents of data access policies. The construction of consumption-based poverty maps requires that the household survey have a large subset of the variables that are also in the census and that they are coded in the same way, which is not always the case.

What Determines How Well Geographic Targeting Works?

Various factors influence the effectiveness and appropriateness of geographic targeting: its accuracy in the setting, the type of programs, and the alternatives available.

The potential for this approach depends heavily on being able to identify and measure accurately variables that are highly correlated with the welfare concept underlying the analysis, whether it is basic needs, capabilities, or consumption. A poverty map based on poor data cannot be expected to be accurate any more than a means test based on wrong information would be. However, only for the case of consumption-based poverty mapping can one really evaluate this underlying association, and this is undoubtedly one of its attractions. In our discussion below we will primarily be concerned with consumption-based poverty mapping, but much of what we say will also apply to maps based on other concepts of welfare.

The correlation between space and poverty is an obvious determinant of the accuracy of geographic targeting. Often there is a strong correlation, with some areas much wealthier than others due, for example, to differences in climate or natural resource base, different histories of policy and administration, or stock of infrastructure. These may occur more frequently in large countries, those with diverse geographic features, and/or those with decentralized administrations. In Brazil, a large, federal country, poverty rates are only 5.1 percent in the southeastern state of Sao Paulo, but 10 times that at 51.4 percent in the northeastern state of Maranhao (World Bank 2001). In contrast, in Jamaica, a small nation with a unitary government, although the wealthiest parish has mean consumption twice that of the poorest (STATIN and PIOJ 1994), only 11 percent of total inequality in consumption is due to differences between parish averages (STATIN 1989); the rest is due to differences within parishes. In the former Soviet Union, one of the legacies of central planning is that the cities tend to have more homogeneity of wealth across areas than found in urban areas in market economies.

In most countries the census is the only source of information available for all small geographic units, but it has only a few variables, usually about the demographics of the household and the standards of its housing. In some countries there may also be a few other sources available—perhaps an annual census of the height and weight of first-grade children in school or maps of soil quality and rainfall patterns. These may be a little more closely correlated with welfare, but they are still likely to be imperfect substitutes, especially if many children are not enrolled in school or if agriculture does not form a large part of the economic base. In the past this limited either the choice of variables or the level of disaggregation possible and thus was a major constraint on the accuracy of poverty mapping. The introduction of census-survey prediction methods begins to address this constraint.

For a given spatial distribution of income the proportion of total inequality explained by variations in mean incomes between areas will increase with the level of disaggregation. The advantage of using the national census is that it usually contains information at the lowest geographic unit available among all data sources, and this can greatly enhance the precision of targeting. However, two factors often work against using the map at the lowest level of disaggregation possible. First, it may be that even at this level the spatial variation may be low in some countries so that targeting only

certain areas will lead to large undercoverage rates. One can instead allocate resources to all areas, say, using a formula based on the index. Second, since the precision of predicted consumption decreases at lower levels of disaggregation, one cannot be very confident that targeting will in practice be effective. For both reasons, the best approach may be to use geographic targeting for allocating resources across areas and then use another targeting method to target within areas.

There are also some administrative issues about how the poverty map is used that relate to the level of service delivery. Even when poverty maps are accurate for very small units, it may be administratively awkward to take full advantage of that information. It may be difficult for a school district to offer free textbooks in some schools in its district but not in others, or for a health center to charge fees to those in some neighborhoods but not in others. It might not, on the other hand, be administratively difficult to have some school districts have free books or for some types of fees to be waived in all rural areas.

Some types of programs or services can use geographic targeting better than others. Poverty maps are very often used to guide investments in infrastructure, for example, where to build new roads or schools. In the more limited context of safety net programs, they are used in various ways.

For public works programs and social funds, citing infrastructure is important and poverty maps are used both to ensure that the area served by the infrastructure is poor and unserved and that the surrounding population from which the workers will likely be drawn is poor.

Poverty maps can be used to select areas where a subsidized service will operate or where transfers will be available. How well this works will depend on how often the services or purchases are needed, especially in urban areas. People will not travel as far to use a service that they frequently need as they will for one they need only occasionally. Thus we would expect that the targeting that results from citing in poor area shops that sell subsidized tortilla, which must be purchased fresh daily, would be better than from shops that sell a month's ration of subsidized flour or rice. If stigma, danger, or inconveniences are issues we might also expect geographic targeting to work better for programs that one must use in person. A rich household can, for example, send its servant to a poor neighborhood to buy a month's supply of subsidized milk powder, but it probably would not send its child daily to a drink a glass of milk at a feeding center in a poor neighborhood.

Community-based targeting methods, by definition, work only to allocate a budget given to the community for a specific program and thus must have some other allocation mechanism among communities if overall targeting is to be effective. Geographic targeting is the most sensible adjunct. For example, in the Indonesian scholarships program launched after the financial crisis of 1997/98, geographic targeting was used to allocate the budget to districts, and within the districts to the individual schools. Within schools, community targeting was used with the neediest children chosen

by a committee made up of community members and local officials (Government of Indonesia 1998).

A poverty map alone is rarely used to target a large benefit, especially one in cash, because of the inherent limits to its accuracy.

Finally, the feasibility of other options will influence whether to use geographic targeting. Where administrative capacity is high enough to run a means-tested cash transfer program, geographic targeting may not be a useful substitute or adjunct. Where means testing is out of the question, one alternative might be to use a combination of geographic and community targeting; another might be to use a combination of geographic and self-targeting.

What Are the Costs Associated with Geographic Targeting?

Geographic targeting is popular in large measure because it requires so few administrative resources. A small team of analysts can build a map (and indeed this can even be contracted out) using pretty much whatever data are available, though clearly the accuracy of the map will be greater if good data are gathered at disaggregated levels every few years. The map can be used by a host of agencies that need only an intuitive understanding of how it was constructed.

How a poverty map is used is often determined by political and social as well as economic factors. A poverty map ranks areas by poverty and may give some sort of absolute scale of poverty. However, the map itself does not say anything about how many resources to give to which areas. The degree of gradation in resource allocation according to the poverty map can vary considerably. In some cases the gradation is slight, so that on a per capita basis the poorest district may receive only, say, 10 percent more per capita than the richest. In other cases (the difference is not so likely to be over time as from place to place or program to program) the gradation is quite sharp, with the poorest areas receiving several times as much as the richest. Indeed, a program may rule out wealthier areas altogether and give graduated levels of resources among areas that participate.

Where the operation of the program involves high fixed setup costs it may be more cost-effective to concentrate resources in only a subset of regions. However, excluding some regions may be deemed politically infeasible. For example, in building their system of geographic targeting, early planners in the Bolivian Social Investment Fund once thought to rank *cantons* (counties) nationally and serve only the poorest grouping. Because some departments were so much better off than others, they had few or no high-priority *cantons* within their boundaries; indeed, most of the priority *cantons* fell in just three departments. The representatives of the other departments protested at funds going to only those *cantons*, especially as the broader issues of decentralization and the sharing of revenues between the national and departmental governments were receiving much attention at the time. The priority ranking system had to be modified so that it was used

to establish priorities with departments but not among them. Within each department, *cantons* were ranked according to literacy, school enrollment, and infant mortality rates. The ranked lists of *cantons* were divided into four priority groups containing approximately equal shares of the population. The cutoff point of the poverty index that delineated the difference between the first and second tiers in the priority system thus varied by department. Priority 1 *cantons* in each department were given equal treatment in the project promotion and appraisal (see Grosh 1994). Similar forces influenced the geographic allocation of funds for the Egyptian food subsidy system as well as for the food-for-education program in Bangladesh. In the Egyptian case, the allocation was not very highly correlated with the underlying poverty map so that other forces were obviously important (Ahmed and others 2001). In the Bangladesh case, for political reasons all *thanas* were eligible for the program, but within these geographic units the poorest *unions* were chosen (Ahmed and del Ninno 2001).

Political factors may also have an influence on the targeting index used. For example, if the poorest regions are less densely populated, then it may be that targeting based on a poverty headcount measure allocates more funds to wealthier regions, whereas targeting based on a severity of poverty measure results in a greater concentration of resources in fewer but poorer regions. The outcome may thus depend in part on whether the poorest districts are concentrated in a few provinces or scattered among them and depending on the political features of the country (e.g., the role of regionally based legislators in budget allocations, the general nature of decentralization or federalism and resource use, or perhaps confounding factors of regional ethnic divides with different welfare levels). Also, where the program starts out being very concentrated in a few regions there is usually increasing pressure to expand the program to other areas, especially when the program involves health and education services for which equality of opportunity (and not just alleviation of current poverty) is seen as the driving force. It may be that some areas are more income poor but that all are equally education or health poor.

How funds are allocated will also raise important social issues. For example, concentrating resources in the poorest areas, even when cost effective, will tend to exclude some who are equally poor but just happen to live in a less poor region. This goes against the principle of horizontal equity, which requires that those who are alike in all relevant characteristics should be treated similarly. Adherence to the principle of horizontal equity requires either that program resources go to all regions in proportion to the extent of poverty or that the allocation of resources for other programs (possibly using different targeting methods within regions) rectify the situation.

In our experience, a great deal of effort has gone into the creation of competing poverty maps based on different *ad hoc* composite indicators of poverty and then the debates about which map is best overall or for a specific purpose. The art of geographic targeting could probably be advanced were that energy rechanneled into formulating maps that combine census

and survey information and, perhaps most importantly, into addressing how the poverty map should be used. We have, for example, seen countries spend years arguing over minor changes in data source or *ad hoc* indices and then only very slightly differentiate the resources assigned to different areas according to the poverty ranking.

Appropriate Circumstances

Poverty maps are applicable for countries of all levels of income and administrative capacity. They can be and are used for all sorts of programs. Geographic methods can be combined with most other targeting methods and, in fact, in our sample no program used only geographic targeting. Some methods, for example, community-based targeting effectively requires a prior level of geographic targeting. Thus at some level they are very broadly applicable. Yet by themselves they are not terribly accurate. Thus the choice of whether to use them alone or in combination with other methods will depend on several factors. All else being equal, they will be best used where poverty is highly spatially correlated, where the benefit is linked to a service used daily and in person, where the benefit is small, or where other methods are not feasible.

Demographic Targeting

Demographic targeting—by age and/or sex—is a common form of categorical targeting. Our database contains 24 examples of programs targeted to the elderly and 36 for young children; 19 of the 24 programs targeted to the elderly are cash transfers spread fairly evenly between LAC, Eastern Europe, Sub-Saharan African, and East Asia. For the programs targeted to children, 24 of the 36 are for cash transfers. LAC dominates with 14 programs; Eastern Europe and East Asia follow with 11 and 8 respectively; and Sub-Saharan Africa, MENA, and South Asia have only one program each.

The median demographically targeted intervention by age or sex in our database performed no better or worse than non-demographically targeted interventions.[8] However, there were significant variations within these demographically targeted programs. Programs targeted toward the young on average did better at transferring resources to the poor than those targeted to the elderly. Interventions that combined demographic targeting with some other method performed considerably better than those that used demographic targeting alone. The best demographically targeted intervention in our database that used no other targeting method is the payment of family allowances in Hungary with a performance score of 1.57. Yet, 12 of the 23 demographically targeted interventions that also used means testing or geographic targeting transferred a greater proportion of program resources to poor households.

In our database, the programs with the most progressive incidence targeted to the elderly were found in Chile (the CAS-PASIS) and Costa Rica

(*Pensiones no Contributivas*) while the poorest were transfer programs in Vietnam and Latvia. In Chile, eligibility was determined jointly by age and a proxy means test based on a home visit. The test was overseen centrally, but the determination of eligibility was contracted out to local authorities and the private sector. In Costa Rica, individuals had to apply for benefits at local offices; social workers reviewing applications had considerable latitude in determining eligibility (Grosh 1992). Among programs that used only demographic targeting to the elderly, that with the most progressive incidence, Bulgaria's pensions, had a performance score of 1.10.

The basic notion of demographic targeting is very simple: to select groups defined by easily observed characteristics (e.g., the old, the young, or female-headed households,) that are poorer than average and to make them eligible for some sort of benefit. In some cases, such as the universal child allowance, no other criteria are used. However, demographic targeting can be combined with other methods; about three quarters of programs targeted toward children also use means testing or geographic targeting.

While the criteria for demographic targeting appear remarkably simple, implementation may still require some effort. For programs limited to young children and the elderly, proof of age is often required and a national identity number if the country has such a system. If these are readily available, the targeting mechanism does not add to administrative requirements. In the absence of such documentation, program administrators risk errors of inclusion if proof of age is not required or errors of exclusion where documentation is required but is difficult to obtain.

Although in principle, obtaining these documents should be straightforward, in many countries, and the more so in poorer countries, poor people may have problems getting them. The poor may not be aware of the importance of the document or understand the procedures to get one. They may face barriers of language or literacy. In addition to having to pay official fees and possibly bribes, they face transactions costs—in time and transport costs—associated with visiting registry offices. If the offices providing these documents have poor service—erratic hours, long lines, and unreliable service standards—the transactions costs for clients increase where multiple visits are necessary. If proof of own or parents' marital status has to be supplied, the problem can be compounded, especially where many unions are not officially registered. Problems are usually more marked for programs for the elderly than for the young. Registry systems are better now in many countries than they were 60 years ago. It is easier to find witnesses needed for informal testimony or formal affidavits for recent events than for distant ones if these can be used in lieu of or to obtain a birth certificate. Of course, it is easier for program workers to guess whether a child is above or below five years of age than an adult above or below 65 (or whatever the relative cut offs are) and to proceed without documentation. Even the former may be difficult where malnutrition is a significant problem.

Programs targeted to pregnant and lactating women usually require a medical certificate verifying pregnancy and stating the expected date of

birth. This may be a source of additional transactions costs, though perhaps not for women who are getting regular prenatal care. For women who would not seek such care in the absence of the intervention, the exam has a cost to them but should confer some medical benefit in return. Sometimes women must get recertified into a "lactating" category once their babies are born.[9] Experience shows, however, that such a requirement should usually be avoided. It adds a bureaucratic process costly to both the program particuipant and the government that can be easily eliminated by calculating the exit date of eligibility the appropriate number of months after the expected birth date of the child.

Two other administrative aspects of demographic targeting should be noted.

- Dispensing individualized benefits can require millions of transactions—interventions as disparate as Peru's "Glass of Milk" program and old-age pensions in Vietnam serve more than one million beneficiaries (Coady, Grosh, and Hoddinott 2003). The broad targeting criteria result in a large client base and thus in a large operational administrative requirement to ensure that transfers are provided reliably, that new beneficiaries can be added as they become eligible, and that individuals who become ineligible are dropped.
- Both the very young and very old are more physically immobile than the rest of the population and this has important implications for the take-up of benefits, especially in rural areas of poor countries. An instructive comparison is that of South Africa and Namibia. In South Africa, the physical delivery of pensions is undertaken by a government agency. Mobile automatic teller machines are driven through the countryside, stopping at meeting points such as markets or shops. Take-up rates in rural areas are about 80 percent (Case and Deaton 1998). By contrast, in Namibia, the delivery of pensions was contracted out to a private firm that provided few pay points. In more remote northern areas, some pensioners had to travel as much as 100 km to receive their pensions. This feature, combined with difficulties in obtaining documentation and registering beneficiaries, resulted in take up averaging 48 percent across the country, with coverage being as low as 30 or 34 percent in some areas (Subbarao 1998).

What Determines How Well Demographic Targeting Works?

As a stand-alone method, the effectiveness of demographic targeting depends on whether the group identified—the elderly, children, or other categorization—is poorer on average than groups that are excluded. There are enormous variations in the poverty levels of these categorizations. For example, the incidence of poverty among Polish preschool children is twice as high as it is for individuals aged 65 or older. By contrast, the

incidence of poverty among the elderly in nearby Estonia is 50 percent higher than it is among preschoolers (Braithwaite, Grootaert, and Milanovic 1999). However, a number of important issues is associated with measuring these poverty levels.

- The first, which is common to all poverty assessments, is the importance of distinguishing between the incidence and severity of poverty. In the Polish and Estonian examples above, although incidence differed dramatically across age groups, the poverty gap, a measure of the depth of poverty, was approximately the same for both the old and young in both countries.
- Second, it is important to understand the nature of living arrangements when targeting by age. In South Africa, it is argued that because young children often reside with their grandparents, the old-age pension is an effective means of reaching young children; Duflo (2000) argues that pension income received by women has had a significant impact on improving the nutritional status of children, particularly girls.
- A third issue relates to the construction of the measure of household welfare. If children need less than adults, per-capita measures will, all else being equal, overstate poverty in households with many children. Adjustments for this are referred to as *equivalence scales*. Certain expenses, such as heating, lighting, and to a certain extent housing, are household rather than individual expenses. For such items, a number of people living together can do so more cheaply, in per-capita terms, than living separately. Adjustments for this come under the heading of *scale economies*.

Lanjouw, Milanovic, and Paternostro (1998) demonstrate the importance of carefully considering both equivalence scales and scale economies in their seven-country study of Eastern Europe and the Former Soviet Union. With no equivalence scales, in all seven countries the elderly are less than averagely poor. Households with three or more children are poorer than average, sometimes markedly so. However, even a modest adjustment to equivalence scales (assuming children have consumption needs that are 70–90 percent those of adults) causes this ranking to be reversed. This matters powerfully for policy—should money go to pensions or to child allowances and services for children?

Deaton and Zaidi (2000) and Lanjouw, Milanovic, and Paternostro (1998) provide excellent source materials on the construction of equivalence scales and scale economies. Their general approach is via the following equation: Adult Equivalents $= (A + \alpha K)^{\beta}$ where α adjusts for age equivalences and β for economies of scale. A per-capita measure of household welfare assumes that there are no economies of scale ($\beta = 1$) and that children and adults have the same requirements ($\alpha = 1$). If household consumption is largely food, as in the case of the ultra poor in very poor countries, there are few economies of scale; thus β is close to one. Since children eat less than adults,

equivalence scales would be important and much different than 1 for young children, since infants need few calories relative to adults, thus $\alpha < 1$. As households and nations grow wealthier, consumption patterns change. The share of resources spent on food declines and the share of household "public" goods such as housing and durable goods rises, so the scale economies increase, implying that $\beta < 1$. At the same time, children consume more nonfood goods such as clothing and toys, all of which add to the costs of supporting them and reduce the importance of food-based equivalence scales, causing α to rise closer to 1.

What Are the Costs Associated with Demographic Targeting?

Programs using demographic targeting appear to have lower administrative costs than other targeting methods, although this statement should be treated cautiously given that cost data are available only for a small number of interventions. Private costs are largely related to the transactions costs associated with enrolling and obtaining benefits; as the South African and Namibian examples show, these can vary greatly. An additional appeal of programs for the very young and old is that they are directed at members of society for whom work, or at least heavy work, is not generally expected, thus granting transfers to them does not raise questions of whether it fosters low work effort.

Demographically targeted programs often have high political acceptability. Programs for children are politically popular almost everywhere, partly because they appeal to the arguments of building human capital for the next generation and providing equality of opportunity. Programs for the elderly are usually very popular as well; many societies accord the elderly great respect and no one likes the specter of destitute and disabled elderly, either as a reflection on society's caring or as a possible fate for oneself.

Appropriate Circumstances

Demographic targeting alone is suitable when the correlations between poverty and age are particularly strong, when the political appeal of a program universal to the age group is important, or when there are no other sensible options. It appears to work better in combination with means tests, proxy means tests, or community-based methods as a way of first reducing the target population to a smaller subset in the relevant age cohort, as long as such additional methods are feasible.

Self-Targeting

Our sample includes 13 programs self-targeted with a work requirement, 27 self-targeted through the purchase of subsidized goods, and 10 in which the community must put forward a proposal to benefit, as under a social fund.

Thus taken all together, self-targeting is a commonly used technique. It is closely associated with program types—the work requirement and public works go hand in hand; food price subsidies and self-targeting through the quality of the good do likewise, and community bidding and social funds. The correlations with regions are a bit less marked. Social funds are present throughout the world, though our sample contains only a small number and most of these are found in Latin America. Public works are found mostly in Latin America, Africa, and South Asia. Our sample of food subsidies is heavily based in MENA and is probably reflective of current world experience (though 20 or 30 years ago subsidies were found much more widely).

The results from self-targeting through a work requirement or through community bidding are rather less variable than for other methods, with the work requirement showing better incidence than the community bidding. The work requirement has high and low performance indicators ranging from 1.48 for Indonesia's JPS-PDK to 4.0 for Argentina's *Trabajar* program, and from .93 for the Armenian Social Fund to 1.30 for the Bolivian Social Fund. For self-targeting through purchases of subsidized goods, the results are less good, from .28 for the VAT exemption on fresh milk in South Africa to 1.63 for the subsidized rice ration in urban Andhra Pradesh, India.

How Does Self-Targeting Work?

Self-targeting programs are open to all, but they are designed in such a way that they are used mainly by the poor. The nonpoor choose, of their own accord, not to use them. The factors that contribute to this choice are private or transactions costs of participation, stigma associated with the use of the service or program, and preferences about quality.

Transactions costs can be small or large, implicit or explicit, in the value of time, or in cash. It is always important to understand the order of magnitude of transactions costs. Even if they are not planned to be a targeting tool, they may well affect program participation rates. Let us consider some examples.

The time cost of workfare is a classic form of self-targeting. To receive a payment in cash or food, individuals must perform significant labor. Usually the jobs involve unskilled, heavy manual labor. Usually the jobs are organized offering full-day or nearly full-day employment on days worked. Some programs offer a job for several weeks' or months' duration; others allow individuals to work only occasional days or a week here or there. Such full-time labor means that the workers must reduce the hours spent on other activities. Most of the workers would, in the absence of their public works job, be seeking and getting at least some employment, often as casual day labor or working on their own land or in their own micro-enterprise. Thus they would be generating some earnings in the absence of their workfare jobs. The transaction costs to them of holding the workfare job are the earnings foregone. The extent of these will be influenced by program design. Programs in rural areas in slack agricultural seasons; those that offer day

labor or those that have short work days better allow participants to carry out other labor activities and will have lower foregone earnings than do those that are full time for concentrated periods of several months.

A few programs require some "volunteer" labor or time in helping to organize the program or to render token service to the community in exchange for the benefit. Peru's famous and extensive community soup kitchens require users to take their turn preparing food (Yamada 1994). An old version of the Honduran food stamps program requested a day per month of time spent collecting trash, though it is not clear how rigorously this was enforced (Grosh 1994). Usually the order of magnitude of such volunteer labor is much smaller than for the workfare and does not preclude full employment, though it may slightly reduce hours of casual labor or domestic chores.

A third variant on this theme is the effort required of communities who solicit projects under Social Funds. The rules of course vary somewhat by country, but in all cases a certain organizational effort is required to convene the community or a subgroup of it to put together and shepherd the application. In many cases some sort of formal contribution to the project is also required, sometimes on the order of 5-15 percent of its costs. These contributions may be in cash, materials, or unpaid labor time, depending on the rules of the social fund.

Most other kinds of programs involve transactions costs, although of a smaller magnitude and often not explicitly stated or thought out. Virtually all programs require some action to sign up and/or to collect benefits. This in turn requires, for example, going to a program office, queuing, sometimes paying fees or even bribes to complete paperwork. Depending on the program design and the quality of service offered, multiple trips and long lines can be involved, and recipients may be required to get relatively small benefits. The time and cash costs can be a significant deterrent to program participation even in cases where they are not an explicit part of the targeting mechanism. Examples abound; we give only one as illustration. In Ecuador's *Bono Solidario* cash transfer program, recipients had to go monthly to participating banks to collect their money. This was hardest for those in rural areas, where the portion of eligible households who participated in the program was alleged to be lower than in urban areas. A small survey conducted among those who did collect benefits (Leon 1999) showed that while for most recipients collecting the monthly benefit was manageable, for some transactions costs were high. A third had to wait in the line for more than three hours. Three quarters had to pay some transport costs, and 8 percent had transport costs that were above 10 or 20 percent of the value of their benefit (depending on which category they were in). Four percent had to spend a night away from home. Faced with the prospect of a similar problem, the designers of Colombia's *Familias en Accion* program decided to exclude from its program municipalities that did not have a branch of the commercial bank, thereby excluding about 12 percent of the poorest areas otherwise chosen from its poverty map.

Stigma may be a factor in people's decisions about whether to participate in a program. The degree of stigma associated with program receipt seems to be rather variable depending on such things as the country's general belief structures about the causes of poverty, equality of opportunity, or the role of the state in providing minimum living standards,[10] as well as on program-specific features.

At the program level, how the public outreach, application, and benefit processes are designed can make participation more or less prone to stigma. In public outreach, governments can encourage or discourage stigma. In publicizing the new cash transfer system that grew out of humanitarian assistance, the Armenia government emphasized that the new program was meant only for the poor, essentially trying to encourage stigma as a self-targeting device (Lund 2002). In contrast, in Jamaica when the food stamps program was organized in 1984, the publicity for the maternal-child part of the program included television ads with the pregnant wife of a cabinet official enrolling for her stamps. There the emphasis was on boosting enrollment and acceptability of the program (Grosh 1992). In an increasing number of targeted programs, partly as a feature of good governance and partly to use stigma to deter leakage to the non-poor, a listing of all program beneficiaries is posted in a public place such as the local welfare office or municipality. This, of course, also leaves the poor open to stigma. Sometimes the means of publicly identifying beneficiaries has to do with the form of benefit. Food stamps, ration cards, or vouchers are visibly different from cash and thus clerks and fellow shoppers may notice when participants make purchases with them. Cash, in contrast, does not carry this visible label of provenance.

Is stigma good or bad? Clearly it is a cost to the program recipient who feels it. Whether it is mildly unpleasant or soul-destroying depends a great deal on the context, as well as on the sensitivity of the individual. As with other sorts of transactions costs, stigma is a tool that can help discourage leakage. However, it is quite a blunt tool insofar as it can discourage participation among the poor and work against the promotion of dignity and self-worth as an outcome of development. Thus stigma should be analyzed carefully and used very judiciously.

Features of program quality will also affect peoples' decision about whether to use a subsidized product or service. The second classic example of self-targeting after public workfare is the subsidization of a basic food stuff more consumed by the poor than the non-poor. The idea is to find different staples or variations on them that are nutritionally equivalent or closely so but differ in terms of prestige—sorghum vs corn, broken rice vs whole, coarse flour vs fine, and yellow vs white corn, are examples where the former is usually less prestigious but nutritionally equivalent to the latter. If the price of the less desired commodity is subsidized enough, the poor who are still trying to meet their calorie needs will buy it, while the non-poor purchase the more prestigious one. Of course, the sorting will be inexact and dependent on the relative strengths of preferences and differences in prices.

What Influences How Well Self-Targeting Works?

The size of transactions costs, the degree of stigma, and the differences in sensitivity to them between the poor and the nonpoor are the direct influences on how well self-targeting works.

In self-targeting through low wages, the critical factors are the wage paid relative to the market wage for such labor, and the distribution of wages in the economy. In Argentina's *Trabajar* program, the maximum wage paid was initially set at the minimum wage and subsequently lowered (and was subsequently about the equivalent to the earnings of the lowest decile of the population) and a few districts elect to pay somewhat less (Jalan and Ravallion 1999). Its performance indicator, at 4, is the best observed for any program in our sample 4.0. The Bolivian Emergency Social Fund, in contrast, paid the prevailing wage in the construction industry (Newman, Jorgensen, and Pradhan 1991). Targeting was somewhat less progressive, with a performance indicator of 1.93 because the reference wage was not undercut and because construction workers were not among the very poorest. If there are a lot of people earning near the public works wage, targeting will not be as good as it will be when the wage gradient is steeper. There is an inherent contradiction between fine targeting and the level of benefit. A low wage will ensure good targeting but will limit the benefit level. In extremely poor settings where the market wage is already very low, it may be important to verify that the net wage (after taking into account the caloric expenditures required to do the job) is high enough to meet welfare objectives.

Even in cases where the wage is set low enough to ensure that applicants for jobs are poor, if the program is not large enough relative to demand, then some kind of other rationing system will be needed. Sometimes this is informal—who knows the foremen or the party politicians. Other times is it formal—use of a lottery (as was considered in Argentina's *Trabajar* program; Jones 2002) or proxy means test (as in Colombia's *Manos a la Obra)* or some community group decision (as in the South Africa's Western Cape; Adato and Haddad 2002).

It can be difficult to identify vehicles suitable to transfer substantial resources. Workfare can do so, but it is suited only to those who can work and conventionally to only those who can do hard physical construction labor. This certainly excludes children, elderly and physically handicapped persons, and the very ill (including AIDS sufferers), and, depending on social conventions, may exclude others. In Indonesia's *padat karya* workfare programs, women participated to significant degree only in (Hindu) Bali; in other (Muslim) provinces their participation was negligible.[11] Such jobs may not be taken up by those with secondary education or who had held white-collar jobs.

Aside from workfare, it can be hard to find a self-targeting tool that allows both good targeting and provides a substantial benefit. For a cash payment of any size, the queuing costs alone would not be sufficient to guarantee good targeting. For food price subsidies, there may not be a

commodity that is consumed more by the poor than by the non-poor (especially if this is judged in absolute terms rather than in relative ones). Even if there is one, it needs to have a production and trade chain that makes it easy to attach the subsidy. For example, grain grown by smallholders and sold in a thriving private market to dispersed outlets will be harder to subsidize than a product that is largely imported by a monopoly state trading agency. Consumption patterns are important. Sorghum or millet, for example, may be consumed not only by poor humans but also used as animal feed. Thus subsidies on these grains may result in a costly indirect subsidy to the livestock industry.

What Are the Costs Associated with Self-Targeting?

First, self-targeting relies on social and transactions costs to the participants. Social costs are difficult to quantify. Transactions costs in the form of foregone earnings for workfare can be on the order of one quarter to over one half of the wage paid. Ravallion and Datt (1995) report them as one quarter for the Maharashtra Income Guarantee Scheme. Jalan and Ravallion (1999) report them for Argentina's *Trabajar* to be about half of gross earnings. Newman, Jorgensen, and Pradhan (1992) calculate them as being 60 percent for workers on Bolivia's ESF.

Nonwage costs in workfare are also significant. Skilled labor, materials, equipment, and administration are important inputs and incur considerable costs. The share of unskilled labor in total costs tends to vary from a low of around 20 percent to about 60 percent at the higher end of the spectrum. Occasionally programs may have higher wages shares, but these programs tend to operate only a narrow portfolio of work types (which can limit their size or impact) and/or put less emphasis on the long-run return to the works done (Subbarao and others 1997).

It must also be acknowledged that though the targeting part of workfare programs can be simple, administering the program as a whole is complex. Most programs run a plethora of small works sites, widely scattered over the whole country or the poorer areas of it. Centrally run programs may have to do the whole gamut of tasks, for example, from identifying works, designing the engineering specifications, procuring materials, hiring workers, or supervising construction. In Social Funds, some of these roles are delegated to the soliciting agencies with the Social Fund limiting its role to that of financier and supervisor. Even so, it is a complex job. Where the public works program has features that will make it an especially useful anti-poverty tool (e.g., offering jobs during the slack agricultural season, offering short term or daily work, or guaranteeing employment within a certain distance of the worker's home) the administrative complexities can increase enormously. How well the program is administered in general will certainly influence its value and probably its size both in the short and long run.

In other sorts of programs that have transactions costs high enough to play an important role in targeting outcomes, these are usually as much the

result of low administrative budgets and capacity as of intentional design and thus do not raise the costs found on the program budget, but rather those borne by the applicants.

Administrative costs for food price subsidy programs are usually rather low, though often unquantified. There may, however, be significant incentive costs in the economy depending on how the subsidy is applied and financed. Implicit taxation of the agricultural sector or direct purchase by marketing boards at prices below those on the world market are fairly common and produce a chain of distortions in the sector. Rarer these days, though once more commonly associated with general food price subsidies, use of disequilibrium exchanges rates produces even more widespread distortions in the economy (Pinstrup-Andersen 1988; Alderman 1992).

Political costs are usually low for self-targeting programs. Indeed, such programs can be very popular, though the reasoning will vary slightly by the type of self-selection used. Food subsidies may reach a very large share of the population and thus have many direct beneficiaries and supporters. Subsidies to basic public health or education services, or to add-on benefits attached to them, usually benefit from the halo of the desire to provide/human right to have access to these services. Again, a fair share of the population may use such providers for at least some services or some part of their life cycle. Benefits from workfare are usually much more finely targeted, but since workers on public works schemes are required to work, usually at jobs that are hard, uncomfortable, and dirty, they are perceived to deserve their benefit.

Appropriate Circumstances

Self-targeting is especially desirable when other methods are less feasible than usual, notably where administrative capacity is particularly low, in crisis settings, and where incomes are irregular. The first case is perhaps obvious and more need not be said. In crises settings, there can be a cascade of households into (deepened) poverty as jobs are lost, public services or transfers are cut, and household's coping abilities are exhausted. It is hoped that the crisis will quickly abate and that households would start to move out of poverty quickly as the economy picks up and their employment or wages recover. In such a setting, to be very accurate, a means tested or proxy means tested program would have to be certifying and recertifying many households very rapidly, a task which is daunting and expensive at best and often completely infeasible. Proxy means tests may not even be accurate in a crisis, as they are usually based on largely static indicators. Similar issues can apply in settings where household incomes are quite uneven, and especially where they are unpredictable as well as uneven. Farmers can at least plan for the annual cycle of harvest sales. In Armenia, in contrast, qualitative poverty research (Gomart 1998b) showed that many households had insufficient regular sources of income and moved in and out of poverty during the year as they managed to find a temporary job, sell

a durable good, or receive a transfer from abroad, all of which were hard to predict.

Notes

1. It is our impression that these are fairly common for in-facility based systems of fee waivers in the health sector. In essence, a member of the medical staff decides who gets a fee waiver. If the staff member has no information other than his/her own knowledge of the applicant and automony to make the decision, then this becomes an example of community-based targeting.

2. When we estimate a regression for the subsample of programs that use means testing, we find that targeting performance rises with country income, although for the reasons discussed earlier, such results should be treated with care.

3. As noted earlier, the absence of cost data prevents making such statements less tentative. Our data suggest that the median cost of transferring $1 to a beneficiary is $1.38 in the case of means tested programs, as opposed to $1.16 in the case of non-means tested interventions, though we are unable to make such a comparison controlling for program type. Targeting and administrative costs for means tested interventions reported in Grosh (1994, pp. 36–38) are generally higher than for other methods, on the order of 10 percent of total program costs. Lastly, Atkinson (1995) shows that the administrative costs associated with welfare payments to families in the U.K.—called "Family Credit"—are, at 5.3 percent of expenditures on benefits, more than twice as high as the costs associated with the universal child benefit program, implying that the means testing costs about 2–3 percent of the program budget.

4. The next worst performer in the category is the Viviendas program, targeted using the Chilean proxy means test, which does require a home visit. It has a performance 25 percent better than neutral targeting. Its poorer performance compared to the other Chilean programs targeted with the CAS is perhaps due to the requirement in the housing program of a substantial downpayment and mortgage service requirements, which tends to exclude the poor (see Sancho, 1992).

5. In Costa Rica, candidates for university student loans are asked to present the same kind of documentation discussed here. Because the base of applicants—university students—is well off compared to the population in general and because literacy and documentation are extensive in Costa Rica, this does not present the same problem it would for a program targeted to a poorer client base, much less in a country with less extensive literacy and written documentation of economic transactions (see Trejos, 1992).

6. Allowing for different definitions or interpretations of what constitutes need makes the evaluation of targeting outcomes difficult. If outcomes in different places are not the same when measured by a given standard (e.g., annual household consumption per capita as collected on a detailed household survey questionnaire), is one outcome better and the other worse? Or does each represent an optimal targeting where the definitions of need are different? If the outcomes are not strongly correlated with annual household consumption per capita, are they inaccurate? Or highly accurate according to the local definition?

7. The problem is generic to all actors in a program. For example, national politicians responsible for the program's design and major budget allocations are presumably interested in getting re-elected. But community-based methods at

minimum have an added layer of workers and moreover workers whose main motivations have nothing to do with the program's success and so may be affected more by this issue.

8. Provided we exclude interventions where the targeting method is self-selection based on consumption (e.g., self-targeted food subsidies).

9. As a criterion, lactation is not usually meant literally, but rather refers to a period of eligibility of six months or a year following the birth of a child.

10. For example, see Rainwater (1982) or Graham (2002).

11. There are design features that can help encourage women's participation in workfare, at least in some contexts, e.g., payment in food rather than cash, payment for piece rates, provision of adequate latrines, provision for onsite child care (see Subbarao, 2002).

5

Summary and Discussion

The primary audience for this book is policy makers and program managers in developing countries, in donor agencies, and in nongovernmental organizations who have responsibility for designing transfer programs to reach the poor. Our objective is to convey what targeting options are available and what results can be expected from them, as well as information that will assist in choosing among options and in implementing those chosen to good advantage. In this final chapter we summarize the results of the book and draw some implications from the findings.

What Can We Say about the Effectiveness of Targeting?

While our efforts to construct a database of targeted interventions were exhaustive (or at least exhausting), there are good reasons to believe that the database is neither a random sample nor a census of all targeted programs and their impacts. Further, as new assessments become available and as governments and donors continue to learn from their own experiences and the experiences of others, some of the quantitative findings will undoubtedly change. This should be borne in mind when considering these findings.

Targeting is a means toward the end, which is poverty reduction. Assessing the effectiveness of targeting is an exercise in assessing one component of antipoverty interventions. It should not be confused with an assessment of *all* impacts of targeted interventions on welfare, which is beyond the scope of this book. Programs may have other objectives than transferring money to the poorest households, and these objectives might involve a tradeoff with targeting performance. For example, social funds may be primarily concerned with creating community infrastructure and with building local capacity and social capital. That they show somewhat less progressive targeting outcomes than some of the purer transfer programs studied is not to say that they are bad policy. It does suggest that policy makers who

are thinking about intervention choices must consider the whole set of strengths and weaknesses of programs in making their decisions.

Mindful of these caveats, we seek to convey five core messages about targeting effectiveness:

- *Targeting can work* . . . Across all programs for which we could obtain information on targeting performance, we find that the median program provides—approximately 25 percent—more resources to the poor than would random allocations. The best programs we found were able to concentrate a high level of resources on poor individuals and households. Argentina's *Trabajar* public works program, the best program in this regard, was able to transfer 80 percent of program benefits to the poorest quintile. The best 10 performers deliver to the poor two to four times the share of benefits that they would get with random allocations. Progressive allocations were possible in all country settings, in countries at markedly different income levels and in most types of programs.
- *. . . but it doesn't always*. The state of the art as practiced around the world is highly variable. While median performance was good, in approximately 25 percent of cases, targeting was regressive so that a random allocation of resources would have provided a greater share of benefits to the poor. With the exception of targeting based on a work requirement, every method contained at least one example of a regressive program.
- *There is no clearly preferred method for all types of programs or all country contexts*. In our sample of programs, 80 percent of the variability in targeting performance was due to differences *within* targeting methods and only 20 percent due to differences *across* methods.
- *A weak ranking of outcomes achieved by different mechanisms was possible*. 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 relative to targeting that uses self-selection based on consumption. Proxy means testing, community-based selection of individuals, and demographic targeting to children show good results on average but with considerable variation. Demographic targeting to the elderly, community bidding, and self-selection based on consumption show limited potential for good targeting. This ranking cannot be taken as a blanket preference for one method over another. It does not consider cost and feasibility constraints. Furthermore, our regression results should be considered as showing correlations rather than causal relations because targeting methods are themselves choices.
- *Implementation matters tremendously to outcomes.* Some, but by no means all, of the variability was explainable by country context. Targeting performance improved with country income levels (the proxy for implementation capacity), the extent to which governments are

held accountable for their actions, and the degree of inequality. Generally, using more targeting methods produced better targeting. Unobserved factors, however, explained much of the differences in targeting success. There remains great potential for improvements in the design and implementation of targeting methods. If programs with poor targeting success were brought up to median, the mean performance indicator would rise from 1.38 to 1.53.

What Can We Say about the Implementation of Targeting Methods?

A recurring theme in this book is that the quality of implementation matters tremendously to the targeting outcome. While table 4.1 provides a brief summary of the advantages, limitations, and appropriate circumstances for the most important methods, it is important to reiterate that there is no clear recipe for how to target. For this reason, it is important for policy makers and program staff to understand a great deal about the details of the different methods and variations in their application. Chapter 4 provides a more detailed discussion of the methods, including information on variations in how the method can be implemented and in how it has performed. This is complemented by the CD-ROM found at the back of this book that provides an annotated bibliography describing the 122 programs in our database and supplying full references to further resource materials on them.

It is not possible here to review all of the variations and nuances of each method; for that the reader is referred to chapter 4. Some common cross-cutting themes that emerge from that treatment include:

- increased creativity, diligence, and/or administrative budget should be able to reduce errors of exclusion in all of the programs with which the authors are well familiar
- improved administration—streamlined procedures, better manuals, more training, more attention to quality control, adequate staff and equipment—often appear to be justified. This is a general judgment or hypothesis based on our qualitative reading of the cases, since cost data are so scarce. In a significant number of cases, there appear to be unexploited economies of scale because the single program is small and/or because structures could be but are not shared over several programs.

To make these generic points concrete, we recast them in the form of six questions. Program managers can use them to review programs for opportunities to improve implementation and targeting performance.

1. *Are there simple administrative changes that would improve targeting performance? Would they be cost-effective?* For example, would a better public communications scheme be worthwhile? More staff power or transport budget for visits to poor villages or neighborhoods?

Translation of materials or employment of staff fluent in nonofficial languages? Stricter enforcement of rules? A change in eligibility thresholds? Simplifying required paperwork?

2. *What administrative change would lower private costs? Social costs? Would such a change be cost-effective?* For example, is there a way to reduce the number of visits applicants must make to apply for benefits, or the waiting times or transport costs for all transactions? Can who is receiving benefits be kept confidential?

3. *How could the program's administration be improved, either by lowering costs or raising quality? Would such a change be cost-effective?* For example, would the program improve with better operational manuals, streamlined forms, more staff training, more equipment, a better computer system, redeployment or release of some staff, consolidation of overheads, or specific support functions with other programs?

4. *Could the targeting mechanism be used by other programs that are not doing so? Why aren't they? Would the program be better served by a different mechanism?* For example, could various programs define eligibility based on a single means test or proxy means test, thereby spreading the administrative cost over a wider base? Chile and Colombia have applied this principal to good effect in their proxy means testing systems, but most pure means tests are done on a program-by-program basis. Similarly, once a country has developed a poverty map, it can be used for targeting several programs. Are some programs using the "wrong" tool?

5. *Is the technical basis used consistent with "good international practice"?* For example, are the databases and statistical analyses underlying the proxy means test and poverty map sound? Are the measures of welfare used in the means test reasonable?

6. *How good is the monitoring and evaluation system?* Is there a regular management information system to track enrollment, delivery of benefits, and all the components of costs? When was the last assessment of the distributional performance undertaken? Have private transactions costs been assessed? When was the last beneficiary assessment? Has there been a full impact evaluation? How complete were these evaluations? Was the program or its implementation adjusted after the evaluations were done?

Final Remarks

Ten years ago, the platitudes of the day tended to either proclaim the panacea-like virtues of targeting or dismiss it as absolutely impossible or undesirable. Our review generalizes the finding of Grosh (1994) as well as those of individual case studies that targeting can work but that it does not always work everywhere. Targeting is neither panacea nor impossible; rather it is a widely useful but always limited tool. The more extensive database used in this review allows it to go further than earlier work in

establishing a weak ranking among methods by share of benefits concentrated on the poor. Still, this continues to fall short of a clear lexicographic ordering of methods.

What are some of the broad implications of these findings? First, the findings help to set standards or expectations by which to judge the performance of programs or countries. We know that moderately progressive targeting is possible in many places, so that should be the expectation. Countries should not be derided for performance that allows for some leakage—in our sample a performance indicator of 2 was quite good (which implies leakage of 20 percent of the benefit away from the lowest two quintiles of population). Nor, given the range of options available, should countries be excused from regressive outcomes for transfer programs. We can, however, have more limited expectations of targeting performance for programs that rely only on self-targeting of food commodities, targeting to the elderly, or community bidding, and for those in countries with low administrative capacity as proxied by measures of GDP, governance, and voice.

Second, the diversity in practice and outcomes within targeting methods implies that we need further work that deals with issues of implementation and cost effectiveness. Program managers need to be able to know more about the details of what was done elsewhere, why the choices were made, how they worked out, and what circumstances affected the outcomes. Our inability to say anything much about administrative costs points to a very important and specific gap in our knowledge base that should be addressed.

The findings on diversity of outcomes for a given targeting method also points to the importance of creativity and experimentation in devising and implementing targeting methods. This has, in fact, been happening as evidenced by the range of programs reviewed in this book, and we expect the trend to continue as the need to make the best use of limited resources for poverty reduction continues unabated. Moreover, the trend toward decentralization in governance generally implies that there will be increasing variation within countries and thus more from which to learn.

Third, a critical component of learning from such creativity and experimentation is a "culture of public evaluation." Our sense is that such a culture is now developing in much of Latin America, in Eastern Europe, and in parts of South and East Asia. However, we sense that such a culture is less prevalent in much of Sub-Saharan Africa and parts of the Middle East and North Africa. We were repeatedly struck by the difficulty we faced in obtaining evaluations of targeting performance in these regions. While it is certainly possible that we have just missed these evaluations, we probably spent as much time on fruitless searches for these evaluations on African countries as we did on amassing a considerable amount of information on programs in Latin America.

Apart from cases where evaluations were carefully carried out and documented, it was challenging to learn from the diversity of the experiences recorded here. Much of the detail on how methods are conducted and more of the story of why choices are made is never written down. To the extent

that it is, it is written in program manuals and directives that are usually not publicly available, and even if not confidential, not naturally disseminated beyond the agency employees that are meant to use them. Box 5.1 contains a plea for improved—and more consistent—reporting.

To gain a better understanding of what produces the outcomes observed will require a renewed emphasis on program monitoring and

---

Box 5.1. Minimal Desired Reporting for International Comparability

A frustration in reviewing the literature on targeting was the difficulty in making international comparisons. A significant problem was the incomparable nature of different measures of targeting performance. Many papers gave results for one of the bottom 10th, 20th, or 40th percentile, or even for the whole distribution. Others only reported the share of benefits going to the poor, but without specifying what percentage of the population was poor. Others used measures invented by the authors for use in a single country, but without the old standard of incidence and participation across the distribution. Since it is always at least a secondary benefit to any study to contribute to the international knowledge in the field (the first benefit being country-specific policy conclusions), we recommend that any study of targeting include, to the extent that information is available, the following basic information and measures even when the authors also use others.

- *Program Description:* A description of the objectives of the program, the program coverage, the type of transfer, and the structure and levels of transfers.
- *Targeting Methods:* A detailed description of the targeting methods officially used in the program (e.g., geographic, means, proxy means, or categorical) as well as any information regarding how well these operated in practice.
- *Targeting Agents:* Information on which bodies (e.g., central or provincial governments) is responsible for setting eligibility criteria, collecting and verifying this information, and implementing the program.
- *Targeting Performance:* Information on the incidence of beneficiaries and benefits across deciles (or quantiles) defined in terms of some standard welfare measure as well as a discussion of the construction of this welfare measure and analogous information on participation rates. This information can be presented based on individual- or household-level observations. It can also be presented in tables showing actual or cumulative shares.
- *Program Budget:* Information on the breakdown of the total program budget between transfers to households, illegal leakages, and administrative costs. Where possible, program costs associated with the targeting of the program should be identified separately. The overall targeting performance of the program can then be seen as a combination of the incidence of the actual transfers and the share of the program budget absorbed by program costs.

evaluation, but in an expanded sense. We need to know not just that errors of inclusion or exclusion were high or low, but what features in the program's design and implementation lead to those outcomes. Such learning requires administrative data and process indicators, as well as qualitative data from workers in the programs and their clients. Documentation should include a good deal of description, especially of administrative and institutional issues. This learning requires researchers who know programs intimately on the ground, and not just econometricians who have a handy household survey data set from a place they barely know. Such learning would greatly benefit from and contribute to south-south learning among program managers.

Because these types of evaluations become public goods, perhaps it is not surprising that they are underprovided. When evaluations are not made widely available, they become, at best, a form of folk wisdom, passed around a limited pool of individuals. So while governments are encouraged to commission and act on the evaluations we describe here, there is also a role for national governments and international agencies to facilitate the dissemination of descriptive information, operational manuals and forms, or other outputs not usually available in the more traditional publications such as academically driven books and journals. The same trend to decentralization that may foster more experimentation in practice implies that federal governments may now have incentives to ensure that learning takes place among the subnational units. Of course, once such lessons are available nationally, it is not hard to make them available internationally as well. This may prove to be an important supplement to the support to cross-country learning that various international agencies have provided via *ad hoc* efforts like this book.

Fourth, the fact that targeting can be successful in many different circumstances means that the question of the balance between transfers and other types of poverty reduction programs is important. If targeting were too hard, then transfer programs would not meet even the most minimal criteria for success and could be dismissed as a viable social policy option. Since they pass that first hurdle, we must understand better when and how to use them in fostering poverty reduction and development (see Ravallion 2003; Dercon 2002; World Bank 2000).

Lastly, in understanding the role of transfers in the broader policy context it is of interest to understand more about their impacts on recipients. Too often, it seems to us, the success of a targeted antipoverty intervention is cast in terms of its ability to exclude individuals or households who were not intended beneficiaries. While this is not unimportant, we see other issues as equally important or even more so:

- What happened to households as a result of this intervention?
- Were they able to avoid undesirable outcomes such as destitution, malnutrition for their children, untreated illnesses for all members, or loss of productive assets?

- Did it help them to attain desirable outcomes such as school completion for their children, more secure or remunerative employment, or greater earnings in self-employment?
- What aspects of the transfer programs were more conducive to positive outcomes?
- Were there any tradeoffs between good targeting and these other objectives?

Recognition of other issues in turn implies more sophisticated research techniques and data. For the analysis of targeting alone, a single cross-sectional data set is sufficient so long as it contains a good welfare measure and records participation or receipt of program benefits. However, for the analysis of fairly complex household behaviors and the longer-run impacts of program receipt, panel data that follow a household over time will be required.

References

Adams, R. 1998. "The Political Economy of the Food Subsidy System in Bangladesh." *Journal of Development Studies* 35(1): 66–88.

Adams, R. 2000. "Self-Targeted Subsidies: The Political and Distributional Impact of the Egyptian Food Subsidy System." *Economic Development and Cultural Change* 49(1): 115–36.

Adato, M., and L. Haddad. 2002. "Targeting Poverty through Community-Based Public Works Programmes: Experience from South Africa." *Journal of Development Studies* 38(3): 1–36.

Adato, M., and L. Haddad. 1999. "Targeting Poverty Programmes: Community-Based Public Works in South Africa." Food Consumption and Nutrition Division, International Food Policy Research Institute, Washington D.C. Processed.

Ahluwalia, D. 1993. "Public Distribution of Food in India: Coverage, Targeting and Leakage." *Food Policy* 18(1): 33–54.

Ahmad, E., and N. Stern. 1991. *The Theory and Practice of Tax Reform in Developing Countries*. Cambridge: Cambridge University Press.

Ahmed, A. 1992. "Operational Performance of the Rural Rationing Program in Bangladesh." Working Paper on Food Policy in Bangladesh 5. International Food Policy Research Institute, Washington D.C.

Ahmed, A. 2000. "Targeted Distribution." In R. Ahmed, S. Haggblade, eds., *Out of the Shadow of Famine: Evolving Food Markets and Food Policy in Bangladesh*. Baltimore: Johns Hopkins University Press.

Ahmed, A., and K. Billa. 1994. "Food for Education Program in Bangladesh: An Early Assessment." Manuscript 62, International Food Policy Research Institute, Dhaka.

Ahmed, A., and H. Bouis. 1999. "Identifying the Needy: Proxy Means Tests For Food Subsidies in Egypt." International Food Policy Research Institute, Washington D.C. Processed.

Ahmed, A., and C. del Ninno. 2001. "Food for Education Program in Bangladesh: An Evaluation of its Impact on Educational Attainment and Food Security." International Food Policy Research Institute, Washington D.C. Processed.

Ahmed, A., H. Bouis, and S. Ali. 1999. "Performance of the Egyptian Food Subsidy System: Distribution, Use, Leakage, Targeting and Cost-Effectiveness." International Food Policy Research Institute, Washington D.C. Processed.

Ahmed, A., H. Bouis, T. Gutner, and H. Lofgren. 2001. "The Egyptrian Food Subsidy System: Structure, Performance and Options for Reform." Research Report 119, International Food Policy Research Institute, Washington D.C.

Ahmed, A., H. Bouis, A. Hegazy, and A. El-Ganainy. 1999. "The Application of Proxy Means Testing to Determine Eligibility for Ration Card Subsidies in Egypt." International Food Policy Research Institute, Washington D.C. Processed.

Alam, M. 1993. "Case Study from Bangladesh: Socio-Economic Impact of Food for Education Program." Draft Report, Bangladesh Institute of Development Studies. Dhaka.

Alderman, H. 1988. "The Twilight of Flour Rationing in Pakistan." *Food Policy* 13: 245–56**.**

Alderman, H. 1992. "Nutritional Considerations in Bank Lending for Economic Adjustment," Paper presented at the 12th Agricultural Symposium. World Bank, Washington, D.C.

Alderman, H. 1998. "Social Assistance in Albania." World Bank Living Standards Measurement Study Working Paper 134. Washington, D.C.

Alderman, H. 2001. "Multi-tier Targeting of Social Assistance: The Role of Intergovernmental Transfers." *World Bank Economic Review* 15(1): 33–54.

Alderman, H. 2001. "Subsidies as a Social Safety Net: Effectiveness and Challenges." World Bank Social Protection Discussion Paper. Washington, D.C.

Alderman, H. 2002. "Do Local Officials Know Something We Don't? Decentralization of Targeted Transfers in Albania." *Journal of Public Economics* 83(3): 375–404.

Alderman, H., M.G. Chaudhry, and M. Garcia. 1988. "Household Food Security in Pakistan: The Ration Shop System." Working Paper on Food Subsidies 4, International Food Policy Research Institute, Washington D.C.

Alderman, H., and K. Lindert. 1998. "The Potential and Limitations of Self-Targeted Food Subsidies." *World Bank Research Observer* 13(2): 213–29.

Alwang, J. 1991. "A Literature Review of Public Food Distribution in Bangladesh." Working Papers on Food Policy in Bangladesh 1, International Food Policy Research Institute, Washington D.C.

Anh, V.T. 2000. "The Poor at Risk: Surviving the Economic Crisis in Southeast Asia—Vietnam." Centre for Southeast Asian Research/Institute of Asian Research, University of British Columbia, Vancouver. Processed.

Atkinson, A. 1995. "On Targeting Social Security: Theory and Western Experience with Family Benefits." In D. van de Walle and K. Nead, eds., *Public Spending and the Poor.* Baltimore: Johns Hopkins University Press.

Ayala Consulting. 2002. "Final Report: Workshop on Conditional Cash Transfers Programs' Operational Experience." Paper presented at workshop in Puebla, Mexico, April 29–May 1.

Bangladesh Institute of Development Studies and International Food Policy Research Institute. 1985. "Development Impact of the Food-for-Work Program in Bangladesh." Technical Paper, submitted to the World Food Programme.

Besley, T., and R. Kanbur. 1993. "Principles of Targeting." In M. Lipton and J. van der Gaag, eds., *Including the Poor*. Washington D.C.: World Bank.

Bigman, D., and H. FoFack. 2000. "Combining Census and Survey Data to Study Spatial Dimensions of Poverty: A Case Study of Ecuador." In *Geographical Targeting for Poverty Alleviation*. Washington, D.C.: World Bank.

Bitran, R., and U. Giedion. 2003. "Waivers and Exemptions for Health Services in Developing Countries." World Bank Social Protection Discussion Paper 0308. Washington, D.C.

Blank, L., T. Castañeda, J. Goldberg, W. Jack, and C. Beckford. 2000. "Jamaica Social Safety Net Assessment." World Bank, Washington D.C. Processed.

Bourguignon, F., F. Ferreira, and P. Leite. 2002. Ex-ante Evaluation of Conditional Cash Transfer Programs: The Case of Bolsa Escola. Processed.

Braithwaite, J. C. Grootaert, and B. Milanovic. 2000. *Poverty and Social Assistance in Transition Countries*. New York: St. Martin's Press.

Braithwaite, J., and others. 1999. "Russia: Targeting and the Longer-Term Poor." World Bank Report 19377-RU. Washington D.C.

Cameron, L. 1999. "The Role of Social Safety Net Scholarships in Preventing Mass School Drop-Out During the Indonesian Economic Crisis." Department of Economics, University of Melbourne. Processed.

Case, A., and A. Deaton. 1998. "Large Cash Transfers to the Elderly in South Africa." *Economic Journal* 108(4): 1330–61.

Castañeda, T. 1990. *Para Combatir la Pobreza*, Centro de Estudios Publicos, Santiago, Chile.

Castañeda, T. 2003. "Targeting Social Spending to the Poor with Proxy-Means Testing: Colombia's SISBEN System." Social Protection Sector, Latin America and Caribbean Region, World Bank, Washington, D.C. Processed.

Castañeda, T., J. Cercone, and L. Fernandez. 2002. "Costa Rica: Social Spending and the Poor." World Bank Report 24300-CR. Washington D.C.

Clay, D., D. Molla, and D. Habtewold. 1999. "Food Aid Targeting in Ethiopia: A Study of Who Needs It and Who Gets It." *Food Policy* 24: 391–409.

Clert, C., and Q. Wodon. 2001. "The Targeting of Government Programs in Chile." In Gacitúa-Marió, E., and Q. Woton, eds*., Measurement and Meaning: Combining Quantitative and Qualitative Methods for the Analysis of Poverty and Social Exclusion in Latin America*. World Bank Technical Paper 518, Washington, D.C.

Coady, D. 1997. "Agricultural Pricing Policies in Developing Countries: An Application to Pakistan." *International Tax and Public Finance* 4: 39–57.

Coady, D. 2001. "Evaluating the Distributional Power of PROGRESA's Cash transfers in Mexico." Food Consumption and Nutrition Division Discussion Paper 117, International Food Policy Research Institute, Washington, D.C.

Coady, D., M. Grosh, and J. Hoddinott. 2003. "Targeted Anti-poverty Interventions: A Selected Annotated Bibliography." International Food Policy Research Institute, Washington D. C. Processed.

Coady, D., R. Perez, and H. Vera-Llamas. 2000. "Cost-Effectiveness Analysis of the Education Component of PROGRESA." In D. Coady, ed., *The Application of Social Cost-Benefit Analysis to the Evaluation of PROGRESA*. Final report submitted to PROGRESA. International Food Policy Research Institute, Washington D.C.

Coady, D., and E. Skoufias. 2001. "On the Targeting and Redistributive Efficiencies of Alternative Transfer Programs." Food Consumption and Nutrition Division Discussion Paper 100. International Food Policy Research Institute, Washington, D.C.

Conning, J., and M. Kevane. 2002. "Community-based Targeting Mechanisms for Social Safety Nets: A Critical Review." *World Development* 30(3): 375–94.

Cornia, G., and F. Stewart. 1995. "Two Errors of Targeting." In D. van de Walle and K. Nead, eds., *Public Spending and the Poor.* Baltimore: Johns Hopkins University Press*.*

Cox, D., and E. Jimenez. 1995. "Private Transfers and the Effectiveness of Public Income Redistribution in the Philippines." In van de Walle and K. Nead, eds., *Public Spending and the Poor*. Baltimore: Johns Hopkins University Press.

Coudouel, A., S. Marnie, and J. Micklewright. 1998. Targeting Social Assistance in a Transition Economy: The Mahallas in Uzbekistan. Innocenti Occasional Papers, Economic and Social Policy Series 63. Florence.

Datt, G., and M. Ravallion. 1993. "Regional Disparities, Targeting, and Poverty in India." In Michael Lipton and Jacques van der Gaag, eds., *Including the Poor*. Washington, D.C.: World Bank.

Datt, G., and M. Ravallion. 1994. "Transfer Benefits From Public-Works Employment: Evidence for Rural India." *Economic Journal* 104 (November): 1346–69.

Datt, G., E. Payongayong, J. Garrett, and M. Ruel. 1997. "The GAPVU Cash Transfer Program in Mozambique: An Assessment." Food Consumption and Nutrition Division, International Food Policy Research Institute, Washington D.C. Processed.

Deaton, A., and S. Zaidi. 2002. "Guidelines for Constructing Consumption Aggregates For Welfare Analysis." World Bank Living Standards Measurement Study Working Paper 135. Washington, D.C.

del Ninno, C. 1998. "Efficiency of Targeted Food Programs." Food Management and Research Support Project, International Food Policy Research Institute, Washington D.C.

Dercon, S. 2002. "Income Risk, Coping Strategies, and Safety Nets." *World Bank Research Observer* 17(2): 141–66.

Devereux, S. 1999. "Targeting Transfers." *IDS Bulletin* 30(2): 61–74.

Dilnot, A., and G. Stark. 1989. "The Poverty Trap, Tax Cuts and the Reform of Social Security." In A. Dilnot and I. Walker, eds., *The Economics of Social Security.* Oxford: Oxford University Press.

Dorosh, P., and S. Haggblade. 1995. "Filling the Gaps: Consolidating Evidence on the Design of Alternative Targeted Food Programs in Bangladesh." *The Bangladesh Development Studies* XXIII(3–4): 47–80.

Duclos, J. 1995. "Modelling the Take-up of State Support." *Journal of Public Economics* 58: 391–415.

Duflo, E. 2000. "Child Health and Household Resources in South Africa: Evidence from the Old Age Pension Program." *American Economic Review*, Papers and Proceedings, 90(2): 393–98.

Dutta, B., and B. Ramaswami. 2001. "Targeting and Efficiency in the Public Distribution System: The Case of Andhra Pradesh and Maharashtra." *Economic and Political Weekly* XXXVI(18): 1524–32.

Edirisinghe, N. 1987. "The Food Stamp Scheme in Sri Lanka: Costs, Benefits and Options for Modification." Research Report 58. International Food Policy Research Institute, Washington D.C.

Edmonds, E. 2001. "Targeting Child Benefits in a Transition Economy." Department of Economics, Dartmouth College, New Hampshire. Processed.

Elbers, C., P. Lanjouw, and J. Lanjouw. (forthcoming). "Micro-level Estimation of Poverty and Inequality." *Econometrica* 71(1): 355–64.

Ezemenari, K., and K. Subbarao. 1999. "Jamaica's Food Stamp Program: Impacts on Poverty and Welfare. World Bank, Washington D.C. Processed.

Fajth, G. 1994. "Family Support Policies in Transitional Economies: Challenges and Constraints." Innocenti Occasional Papers Economic Policy Series, Number 43. Florence.

Foster, J., J. Greer, and E. Thorbecke. 1984. "A Class of Decomposable Poverty Measures," *Economertrica* 52(3): 761–5.

Gaiha, R., K. Imai, and P. Kaushik. 2001. "On the Targeting and Cost-Effectiveness of Anti-Poverty Programmes in Rural India." *Development and Change* 32: 309–42.

Galasso, E., and M. Ravallion. 1999. "Distributional Outcomes of a Decentralized Welfare Program." World Bank Policy Research Working Paper 2316. Washington D.C.

Gomart, E. 1998a. "Poor Beneficiary Assessment." World Bank Human Development Unit, Country Department III. Europe and Central Asia, Washington, D.C.

Gomart, E. 1998b. "The Poorest of the Poor in America." World Bank Human Development Unit, Country Department III. Europe and Central Asia, Washington, D.C.

Government of Indonesia. 1998. "Program Implementation Plan (PIP): Scholarship and Block Grant Program for Primary and Secondary Schools." Department of Education and Culture, Department of Religious Affairs, Department of Home Affairs, Jakarta.

Government of Mozambique. 1998. "Understanding Poverty and Well-Being in Mozambique: The First National Assessment (1996–97)." Ministry of Planning and Finance, Maputo.

Government of the United States. 1999. *Social Security Programs Throughout the World*. Social Security Administration, Washington D.C.

Graham, C. 2002. "Public Attitudes Matter: A Conceptual Frame for Accounting for Political Economy in Safety Nets and Social Assistance Policies." World Bank Social Protection Discussion Paper 0233. Washington, D.C.

Grootaert, C. 1997. "Poverty and Social Transfers in Hungary." World Bank Policy Research Working Paper 1770. Washington D.C.

Grosh, M. 1992. "The Jamaican Food Stamp Program: A Case Study in Targeting." *Food Policy* 17(1): 17–40.

Grosh, M. 1994. *Administering Targeted Social Programs in Latin America: From Platitudes to Practice*, Washington D.C.: World Bank.

Grosh, M., and E. Glinskaya. 1998. "Proxy Means Testing and Social Assistance in Armenia. World Bank Development Unit, Country Department III., Europe and Central Asia. Washington, D.C. Processed.

Haggblade, S., S. Rahman, and S. Rashid. 1993. "Statutory Rationing: Performance and Prospects." Draft Report, Dhaka: Bangladesh Food Policy Project, International Food Policy Research Institute.

Hassan, F., and R. Peters. 1995. "Social Safety Net and the Poor during the Transition: The Case of Bulgaria." World Bank Policy Research Working Paper 1450. Washington D.C.

Hentschel, J., J. Olson Lanjouw, P. Lanjouw, and J. Poggi. 2000. "Combining Census and Survey Data to Trace the Spatial Dimensions of Poverty: A Case Study of Ecuador." *The World Bank Economic Review* 14 (1): 147–65.

Hoddinott, J. 2001. "Targeting: Principles and Practice." In J. Hoddinott, ed., *Food Security Practice: Methods for Rural Development Projects*. Washington, D.C.: International Food Policy Research Institute.

Howes, S., and S. Jha. 1992. "Urban Bias in Indian Public Distribution System." *Economic and Political Weekly*, May: 1022–30.

Howes, S., and S. Jha. 1994. "Public Distribution of Food in India: A Comment." *Food Policy* 19(1): 65–8.

Huber, P. 1967. "The Behavior of Maximum Likelihood Estimates under Non-standard Conditions." In *Proceedings of the Fifth Berkeley Symposium in Mathematical Statistics and Probability*. University of California Press, Berkeley, California.

IFPRI (International Food Policy Research Institute). 1994. "Options for Targeting: Food Interventions in Bangladesh." Report by Working Group on Targeted Food Interventions. Washington, D.C.

Jacoby, H. 2002. "Is there an Intrahousehold 'Flypaper Effect'? Evidence from a School Feeding Program." *Economic Journal* 112(476): 196–221.

Jalan, J., and M. Ravallion. 1999. "Income Gains to the Poor from Workfare: Estimates for Argentina's Trabajar Program." World Bank Policy Research Working Paper 2149. Washington D.C.

Jamaica, Statistical Institute of Jamaica and Planning Institute of Jamaica. 1998. *Survey of Living Conditions, 1997*. Kingston, Jamaica.

Jarvis, S., and J. Micklewright. 1995. "The Targeting of Family Allowances in Hungary." In D. van de Walle and K. Nead, eds., *Public Spending and the Poor.* Baltimore: Johns Hopkins University Press*.*

Jayne, T., J. Strauss, T. Yamano, and D. Molla. 2001. "Giving to the Poor? Targeting of Food Aid in Rural Ethiopia." *World Development* 29(5): 887–910.

Jayne, T., J. Strauss, T. Yamano, and D. Molla. 2002. "Targeting of Food Aid in Rural Ethiopia: Chronic Need or Inertia?" *Journal of Development Economics* 68(2): 247–88.

Jensen, R. 1998. "Public Transfers, Private Transfers and the 'Crowding Out' Hypothesis: Evidence from South Africa." Faculty Research Working Paper R98-08. John F. Kennedy School of Government, Harvard University, Cambridge.

Jha, S. 1992. "Consumer Subsidies In India: Is Targeting Effective?" *Development and Change* 23(4): 101–28.

Kaufmann, D., A. Kraay, and P. Zoido-Lobaton. 1999. "Aggregating Governance Indicators," World Bank Policy Research Working Paper 2195. Washington, D.C.

Kim, Sang Kyun. 1999. "Cash Assistance in Korea," Seoul National University, Seoul. Processed.

Kim, S., and M.W. Lee. 2001. "Income Distribution and Public Transfers as Social Safety Nets in Korea." *Journal of Economic Development* 26(2): 61–75.

Kriesel, S., and S. Zaidi. 1999. "The Targeted Public Distribution System (TPDS) in Uttar Pradesh, India—An Evaluation." World Bank, Washington D.C. Processed.

Lanjouw, P., B. Milanovic, and S. Paternostro. 1998. "Poverty and the Economic Transition: How do Changes in Economies of Scale Affect Poverty Rates for Different Households?" World Bank Policy Research Working Paper 2009. Washington, D.C.

Larrañaga, O. 2003. "Focalización de Programas Sociales en Chile: El Sistema CAS." Social Protection Center, Latin America and Caribbean Region, World Bank, Washington, D.C. Processed.

Leach, J. 1998. "Cash Transfers A Case Study: Bolivia's Bonosol." World Bank, Washington, D.C. Processed.

Leon, M. 1999. "Informe Analitico de la Encuesta de Beneficiarios del Bono Solidario." Secretaria de Estado de Desarrollo Social, Quito, Ecuador. Processed.

Lund, L. 2002. "On Target in Armenia." *Spectrum*, Social Protection Unit, World Bank, Washington, D.C., pp. 27–31.

Lustig, N. 1986. "Food Subsidy Programs in Mexico." Working Paper on Food Subsidies 3. International Food Policy Research Institute, Washington D.C.

Mark, A. et al. 1999. "Improving Social Assistance in Armenia." World Bank Report 19385-AM, Human Development Unit, Country Department III, Europe and Central Asia, Washington, D.C.

Mason, A. et al. 2003. "Dominica Social Protection Review." World Bank Report 26254-DM. Washington, D.C.

Milanovic, B. 1992. "Distributional Impact of Cash and In-Kind Social Transfers in Eastern Europe and Russia," World Bank Working Paper: Transition and Macro-Adjustment WPS 1054. Washington, D.C.

Milanovic, B. 1998. "Explaining the Increase in Inequality during the Transition." World Bank Policy Research Working Paper 1965. Washington, D.C.

Milanovic, B. 2000. "Social Transfers and Social Assistance: An Empirical Analysis Using Latvian Household Survey Data." World Bank Policy Research Working Paper 2328. Washington, D.C.

Moffitt, R. 1992. "Incentive effects of the US Welfare System: A Review." *Journal of Economic Literature* 30(1): 1–61.

Moffitt, R. 2003. "The Negative Income Tax and the Evolution of U.S. Welfare Policy." *Journal of Economic Perspectives* 17(3): 119–40.

Morris, S. 2001. "Social Nutrition Programs with Explicit Targeting." International Food Policy Research Institute, Washington D.C. Processed.

Munro, L. 2003. "A Social Safety Net for the Chronically Poor? Zimbabwe's Public Assistance in the 1990s." Department of Policy, Planning and Development, Fordham University, New York. Processed.

National Research Council. 1995. *Measuring Poverty: A New Approach*, Washington, D.C.: National Academy of Sciences.

Newbery, D., and N. Stern, eds. 1987. *The Theory of Taxation for Developing Countries.* New York: Oxford University Press.

Newman, J., S. Jorgensen, and M. Pradhan. 1992. "How did Workers Benefit?" In S. Jorgensen, M. Grosh, and M. Schacter, eds. *Bolivia's Answer to Poverty, Economic Crisis and Adjustment: The Story of the Emergency Social Fund.* Washington, D.C.: World Bank.

Okrasa, W. 1999. "The Dynamics of Poverty and the Effectiveness of Poland's Safety Net (1993–96)." World Bank Policy Research Working Paper 2221.Washington D.C.

Osmani, S., and O. Chowdhury. 1983. "Short Run Impacts of Food for Work Programme in Bangladesh." *Bangladesh Development Studies* XI(1–2).

Palacios, R., and R. Rocha. 1998. "The Hungarian Pension System in Transition." In L. Bokros and J. J. Dethier, eds., *Public Finance Reform during the Transition: The Experience of Hungary.* World Bank, Washington D.C.

Park, A., S. Wang, and G. Wu. 2002. "Regional Poverty Targeting in China." *Journal of Public Economics* 86(1): 123–53.

Pinstrup-Andersen, P., ed. 1988. *Food Subsidies in Developing Countries: Costs, Benefits and Policy Options*. Baltimore: John Hopkins University Press.

Pinstrup-Anderson, P., and H. Alderman. 1998. "The Effectiveness of Consumer-Oriented Food Subsidies." In Per Pinstrup-Andersen, ed., *Food Subsidies in Developing Countries*. Baltimore: Johns Hopkins University Press.

Pradhan, M., L. Rawlings, and G. Ridder. 1998. "The Bolivian Social Investment Fund: An analysis of Baseline Data for Impact Evaluation." *World Bank Economic Review* 12(3): 457–82.

Prescott, N. 1997. "Poverty, Social Services, and Safety Nets in Vietnam." World Bank Discussion Paper No. 376. Washington D.C.

Pritchett, L., S. Sumarto, and A. Suryahadi. 2000. *Targeted Programs in an Economic Crisis: Empirical Findings from Indonesia's Crisis*. Mimeo. Kennedy School of Government, Harvard University, Cambridge.

Quinn, S. 1998. "Social Investment Funds. A Case Study: Brazil's Northeast Rural Poverty Alleviation Program." Case study presented at workshop on Transfers and Social Assistance for the Poor in the LAC Region, May. World Bank, Washington D.C.

Quinn, S. 1998. "Social Investment Funds. A Case Study: Nicaragua's *Fondo de Inversion Social*." Case study presented at workshop on Transfers and Social Assistance for the Poor in the LAC region. World Bank, Washington D.C.

Raczynski, D. 1991. "La Ficha CAS y La Focalizacion de los Programmas Sociales." Esudios Cierplan No. 141, La Corporación de Investigaciones Economica para Latinoamérica, Santiago, Chile.

Radhakrishna, R., K. Subbarao, with S. Indrakant, and C. Ravi. 1997. "India's Public Distribution System: A National and International Perspective." World Bank Discussion Paper 380. Washington D.C.

Rainwater, L. 1982. "Stigma in Income-Tested Programs." In I. Garfinkel, ed., *Income Tested Transfer Programs: The Case For and Against.* New York: New York Academic Press.

Rajan, S. I. 2001. "Social Assistance for the Poor Elderly: How Effective?" *Economic and Political Weekly* XXXVI(8): 613–7.

Ravallion, M. 1991. "Reaching the Poor through Public Employment." *World Bank Research Observer* 6(2): 153–75.

Ravallion, M. 1993. "Poverty Alleviation through Regional Targeting: A Case Study for Indonesia." In K. Hoff, A. Braverman, and J. E. Stiglitz, eds., *The Economics of Rural Organization: Theory Practice and Policy.* New York: Oxford University Press.

Ravallion, M. 1998. "Reaching Poor Areas in a Federal System." World Bank Policy Research Working Paper 1901. Washington D.C.

Ravallion, M. 1999. "Monitoring Targeting Performance when Decentralized Allocations to the Poor are Unobserved." World Bank Policy Research Working Paper 2080. Washington D.C.

Ravallion, M., and K. Chao. 1989. "Targeted Policies for Poverty Alleviation under Imperfect Information: Algorithms and Applications." *Journal of Policy Modeling* II(2): 213–24.

Ravallion, M., and G. Datt. 1995. "Is Targeting through a Work Requirement Efficient?" In D. van de Walle and K. Nead, eds., *Public Spending and the Poor.* Baltimore: Johns Hopkins University Press.

Ravallion, M., and Q. Wodon. 1999. "Does Child Labor Displace Schooling? Evidence from Behavioral Responses to an Enrollment Subsidy." World Bank Policy Research Working Paper 2116. Washington D.C.

Rawlings, L. 2002. "Colombia Social Safety Net Assessment." World Bank, Washington D.C. Processed.

Rawlings, L., L. Sherburne-Benz, and J. Van Domelen. 2003. *Evaluating Social Funds: A Cross-Country Analysis of Community Investments*. Washington, D.C.: World Bank.

Razmara, S. et al. 1999. "Consumer Food Subsidy Programs in the MENA Region." World Bank Report 19561-MNA. Human Development Group, Middle East and North Africa Region. Washington D.C.

Rignold, D., and E. Lasagabster. 2002. "Bulgaria: Public Expenditure Issues and Directions for Reform." World Bank Report 23929-BUL. Washington D.C.

Sahn, D., and H. Alderman. 1995. "Incentive Effects on Labor Supply of Sri Lanka's Rice Subsidy." In D. van de Walle and K. Nead, eds., *Public Spending and the Poor.* Baltimore: Johns Hopkins University Press.

Sancho, A. 1992. "La Ficha Cas Como Instrumento de Focalización del Gasto Social en el Nivel Local." In Margaret Grosh, ed., *Platitudes to Practice: Targeting Social Programs in Latin America*. Vol. II, Case Studies III, Latin America and the Caribbean Technical Department, World Bank Regional Study Program, Report 21. Washington, D.C.

Sedlacek, G., and S. Rocha. 1999. "The Bolsa-Escola Program." Paper presented at conference, Social Protection: Strategies for Policy Enhancement in LAC. World Bank, Washington D.C.

Sen, A. 1988. "The Concept of Development." In H. Chenery and T.N. Srinivasan, eds., *Handbook of Development Economics.* Vol. I. Amsterdam: North Holland.

Skoufias, E. 2001. "PROGRESA and its Impacts on the Human Capital and Welfare of Households in Rural Mexico: A Synthesis of the Results of an Evaluation by IFPRI. International Food Policy Research Institute, Washington D.C. Processed.

Skoufias, E., and D. Coady. 2002. "Are the Welfare Losses From Imperfect Targeting Important?" Food Consumption and Nutrition Division Discussion Paper 125, International Food Policy Research Institute, Washington, D.C.

Skoufias, E., B. Davis, and S. de la Vega. 2001. "Targeting and the Poor in Mexico: An Evaluation of the Selection of Households into PROGRESA." *World Development* 29(10): 1769–84.

SMERU (Social Monitoring and Early Response Unit). 1998. "Implementation of Special Market Operation (OPK): Results of a SMERU Rapid Field Appraisal Mission in Five Provinces." SMERU Special Report, Jakarta.

Social Fund Office. 2000. *Social Investment Fund Annual Progress Report for the Year 2000.* World Bank, Washington D.C.

Statistical Institute of Jamaica. 1989. Jamaica Survey of Living Conditions, November 1989. Kingston, Jamaica.

Statistical Institute of Jamaica and Planning Institute of Jamaica. 1994. "Jamaica Survey of Living Conditions," Report 1992. The Planning Institute of Jamaica and The Statistical Institute of Jamaica, July.

Subbarao, K. 1998. "Namibia's Social Safety Net: Issues and Options for Reform," World Bank Policy Research Working Paper 1996. Washington D.C.

Subbarao, K. 2001. "Systemic Shocks and Social Protection: Role and Effectiveness of Public Works Programs." World Bank Social Protection Discussion Paper 0302. Washington, D.C.

Subbarao, K., A. Bonnerjee, J. Braithwaite, S. Carvalho, K. Ezemenari, C. Graham, and A. Thompson. 1997. *Safety Net Programs and Poverty Reduction: Lessons from Cross-Country Experience*. Washington, D.C.: World Bank.

Sumarto, S., and A. Suryahadi. 1999. "Indonesia's Experience with Social Safety Nets: Lessons Learned and Future Prospects." Social Monitoring and Early Response Unit Working Paper.

Sumarto, S., A. Suryahadi, and L. Pritchett. 2000. "Safety Nets and Safety Ropes: Comparing the Dynamic Benefit Incidence of Two Indonesian 'JPS' Programs." Social Monitoring and Early Response Unit Working Paper, February.

Suryahadi, A., Y. Suharso, and S. Sumarto. 1999. "Coverage and Targeting in the Indonesian Social Safety Net Programs: Evidence from 100 Village Survey." Social Monitoring and Early Response Unit Working Paper, August.

Tesliuc, E. 2003. "Kyrgyz Republic: Enhancing Pro-Poor Growth." World Bank Report 24638-KG. Washington D.C.

Tesliuc, E. et al. 2003. "Romania: Poverty Assessment." World Bank Report No. 26189-RO. Washington D.C.

Trejos, J. 1992. "Becas Para Estudios Universitarios en Costa Rica. In M. Gras, ed., *From Platitudes to Practice: Targeting Social Programs in Latin America.* Volume II. Human Resources Division, Latin America and Caribbean Region, World Bank, Washington, D.C.

Tuck, L., and K. Lindert. 1996. "From Universal Food Subsidies to a Self-Targeted Program: A Case Study in Tunisian Reform." World Bank Discussion Paper 351.Washington, D.C.

United Nations. 2001. *Strengthening Policies and Programmes on Social Safety Nets: Issues, Recommendations and Selected Studies*. Social Development Division, Economic and Social Commission for Asia and the Pacific, Social Policy Paper No. 8, Thailand.

van de Walle, D. 1998. "Protecting the Poor in Vietnam's Emerging Market Economy." World Bank Policy Research Working Paper 1969. Washington D.C.

van de Walle, D. 2002a. "The Static and Dynamic Incidence of Vietnam's Public Safety Net." World Bank, Washington D.C. Processed.

van de Walle, D. 2002b. "Poverty and Transfers in Yemen." Middle East and North Africa World Bank Working Paper Series 30. Washington D.C.

Van Domelen, J., L. Rawlings, and L. Sherburne-Benz. 2001. *Evaluating Social Fund Performance: International Evidence from Community-Based Poverty Reduction Programs*. Washington D.C.: World Bank.

Vodopivec, M. 1998. "Transition from Cash Benefits to Work: The Case of Slovenia." *Empirical Economics*, 23: 173–202.

White, H. 1980. "A Heteroskedasticity-consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica*, 48: 817–38.

World Bank. 1990. *World Development Report: Poverty.* New York: Oxford University Press.

World Bank. 1997a. "Romania: Poverty and Social Policy." World Bank Report No. 16462-RO. Europe and Central Asia Region, Washington D.C.

World Bank. 1997b. *World Development Report: The Role of the State.* New York: Oxford University Press, New York.

World Bank. 1999. "Improving Social Assistance in Armenia." Human Development Unit, Country Department III, Europe and Central Asia Region, World Bank Report No. 19385-AM. Washington, D.C.

World Bank. 2000. *World Development Report: Attacking Poverty.* New York: Oxford University Press.

World Bank. 2001a. "Thailand Social Safety Nets: Policies and Programs." World Bank, Washington D.C. Processed.

World Bank. 2001b. "Rural Poverty Reduction in Brazil: Toward an Integrated Strategy." Vol. I & II, World Bank Report 21790-BB. Washington, D.C.

World Bank. 2001c. "Ukraine: Social Safety Nets and Poverty." World Bank Report 22677, Europe and Central Asia Region. Washington D.C.

World Bank 2003. Armenia: Public Expenditure Review. World Bank Report 24434-AM. Poverty Reduction and Economic management Unit, Europe and Central Asia Region. Washington, D.C.

Zaidi, S., and R. Murgai. 2002. *Bangladesh Public Expenditure Review*. World Bank Report No. 24370-BD. Washington D.C.

Index

Targeting of Transfers in Developing Countries: Review of Lessons and Experience reviews the lessons learned from 122 antipoverty interventions in 47 transition and developing countries to quantify targeting outcomes and their determinants and to inform the design and implementation of common targeting methods.

In addition to providing comparative quantitative analysis on targeting outcomes and their determinants, the authors provide a qualitative treatment of common targeting methods. In each case, they review international experiences: how the methods work, what determines how well they work, what costs are likely to be incurred, and what are appropriate circumstances for use. The authors also provide a brief review of targeting, discussing the benefits and costs of targeting, methods for assessing targeting performance, and a taxonomy of targeting methods.

The findings include the following:

· Targeting has the potential to increase the effectiveness of these interventions; the median program transfers 25 percent more to the poor than a universal or random allocation. However, targeting is not assured to work: about 25 percent of targeted programs actually delivered to the poor less than a universal or random allocation would have.

· No single preferred method has emerged for all types of programs or all country contexts. Indeed, of all variation observed in targeting outcomes, only 20 percent was accounted for by the choice of targeting method. Means testing, geographic targeting, and self-selection based on a work requirement are the most robustly progressive methods. Proxy means testing, community-based targeting, and demographic targeting to young children can be progressive but demonstrate large variability in outcomes. Demographic targeting to the elderly, subsidization of food commodities, and community bidding for projects show limited potential for good outcomes.

· Successful outcomes are critically dependent on the method of implementation.

Of particular interest to policymakers and program managers in developing countries, in donor agencies, and in nongovernmental organizations, this book offers important information to facilitate the effective design of interventions that reach the poor.

The CD-ROM provided contains the annotated bibliography with summary information for each of the 122 programs in the database compiled to support the research presented here.

Translations of the book in Russian and Spanish are included on the CD-ROM.
Traducciones al español y al ruso se incluyen en el disco compacto.
CD-ROM на задней обложке данной книги включает её переводы на русский и испанский языки.

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE
sustainable solutions for ending hunger and poverty
IFPRI®

THE WORLD BANK

15769

9 780821 357699