

Targeting Performance under Self-selection and Administrative Targeting Methods

DAVID P. COADY

International Monetary Fund

SUSAN W. PARKER

CIDE, Mexico

I. Introduction

There is a growing recognition in the development community that many existing safety net programs in developing countries are very badly targeted in that a large proportion of program benefits leak to higher-income households and a substantial proportion of poor households are excluded. A recent review of the targeting performance of such programs by Coady, Grosh, and Hoddinott (2004a) found a wide variation in targeting performance and also a strong positive correlation between the number of targeting methods used in a program and targeting performance. However, the data used in the analysis did not permit a deeper analysis of the relationship between each targeting method and targeting performance. This study provides new evidence on this issue by comparing the implications of two targeting methods that are commonly combined in practice, namely, self-selection by households and administrative selection of households by program officials.¹

In order to analyze the implications of different targeting methods for overall targeting performance, one needs to identify separately the different stages of the program participation process. In particular, one needs to identify which

We thank Iliana Yaschine, Bernardo Hernandez, and Citlalli Hernandez for very helpful clarifications on the program as well as untiring assistance with our data questions. We are also grateful to two anonymous reviewers and the editors for detailed comments that led to substantial improvements in the article. We alone are responsible for any remaining errors. The views expressed in this article should not be reported as representing those of the International Monetary Fund. Contact the corresponding author, David Coady, at dcoady@imf.org.

¹ Self-selection involves households or individuals making decisions on whether or not to apply for a program based on their perceived private costs and benefits of participation. Administrative targeting involves government officials deciding who among those applying should be deemed eligible based on household socioeconomic characteristics. For a more detailed discussion of alternative forms of targeting, see Coady, Grosh, and Hoddinott (2004b).

households knew about the program, whether they applied for the program, and, if so, whether they were accepted into the program. Knowledge of the existence of the program and how to apply can be seen as a combination of self-selection (i.e., a household allocating resources to finding out about the program) and the program strategy for informing households. Application for admission into the program will also reflect a self-selection process by households, and this **decision to apply will depend on the private cost of application and the expected level of program benefits if deemed eligible.** Under administrative selection of beneficiaries, acceptance into the program is determined by whether the household satisfies eligibility criteria, which may take the form of means testing or rely on categorical criteria (e.g., being landless, a female-headed household, elderly, or a household with young children).

Evaluating the targeting performance of a program also requires the specification of a performance indicator. In the context of poverty alleviation programs, it is conventional to take household consumption as the most reliable measure of household welfare and to evaluate targeting performance in terms of the effectiveness of the program at channeling a high proportion of benefits to lower welfare households (vertical targeting performance). This will essentially be the approach adopted in the present study. However, where relevant, we also highlight the adverse implications of improved vertical targeting for program coverage of the target poor population and discuss alternative design reforms that could reduce this trade-off.

The paucity of research on the relative contributions of different targeting methods to the overall targeting performance of a program reflects the absence of data that facilitate such an analysis. Typically, in household surveys one only has access to a single binary participation variable based on whether or not a household received benefits from a specific program and on the level of benefits received, so that one is unable to distinguish between the various stages identified above. In the context of poverty alleviation programs in developing or transition economies, we are aware of only one paper that distinguishes between the various stages.² Micklewright, Coudouel, and Marnie (2004) analyze participation outcomes for a social assistance program in Uzbekistan using nationally representative household survey data. However, their analysis of targeting performance was limited by the fact that the survey used does not contain a comprehensive measure of household consumption, and the

² For developed countries, see Heckman and Smith (2003), which combines data from a number of different sources to investigate the sources of inequality of participation among different groups of eligible individuals for the Job Training Partnership Act in the United States. However, data limitations resulted in both application and acceptance outcomes being combined into a single stage.

available data did not allow the authors to disentangle within-community targeting based on community selection from administrative geographic targeting.

The empirical analysis in this present paper is based on data collected for the evaluation of Mexico's *Oportunidades* program. The baseline survey for the evaluation of the program in urban areas was specifically designed to capture the relative contributions of two different targeting methods, namely, self-selection and administrative targeting. As well as containing information on household knowledge of the program, whether they applied, and if they were accepted, detailed information was also collected on household consumption. The specific tailoring of the questionnaire to the issue of targeting also means that many of the measurement problems encountered in earlier papers (e.g., in capturing eligibility criteria accurately) are likely to be substantially reduced. In addition, we are able to match the household data with program administrative data to obtain a more reliable indicator of program participation.

The format of the article is as follows. The next section briefly describes the program and the structure of the data used to evaluate targeting performance. Section III sets out the methodology used to evaluate targeting performance. In Section IV we present and discuss the results of the targeting evaluation and implications of alternative policy reforms for targeting performance, while Section V provides a summary of our findings.

II. Program and Data Description

A. Program Description

In 2002, the Mexican government began an expansion of an existing safety net program under its new name, *Oportunidades*. The existing program, previously referred to as PROGRESA, had focused on rural areas, and the expansion was aimed at including small and medium urban localities. The targeting approach used in rural PROGRESA was predominantly based on administrative selection. To identify eligible households, program officials visited every household in eligible localities and undertook a census, collecting information on a range of household socioeconomic characteristics. A subset of these characteristics was used to apply a previously developed proxy-means targeting approach, whereby numerical "weights" were assigned to these characteristics to calculate a household "score" and eligibility determined by comparing this score to a "threshold" score. In all, 80% of households in eligible localities were deemed program eligible.³

³ See Skoufias, Davis, and de la Vega (2001) and Coady (2006) for more detailed discussions of the design of PROGRESA and an evaluation of its targeting performance.

Because the percentage of eligible households in more urban areas was expected to be substantially smaller than in rural areas, the application of a purely administrative proxy-means approach was deemed too costly and so a strong element of self-selection was introduced. The introduction of a preliminary self-selection stage was intended to save on administrative resources allocated to screening out higher-income households as well as on resources allocated to visiting every household to do a census. Once participating communities were identified, an information campaign was initiated at the municipal and community levels to inform people of the existence and objectives of the program, the rules for program eligibility, and how to apply for the program. A range of media was used, including TV and radio advertisements; the distribution of flyers; the placement of posters in churches, schools, health clinics, and market places; and loudspeaker announcements. In principle, these were to be concentrated in the poorest blocks. The population was informed that a program office would be located in or near their locality during the months of June–August 2002, which they should visit to apply for the program. Decisions regarding the precise nature of the publicity campaign and its financing were decentralized to municipalities but with some oversight and assistance from program officials.

When households turned up at a program office they were asked to provide information on the specific socioeconomic characteristics that are used to calculate their score. Based on the administrative proxy-means model presented in table 1, the applicants were informed whether or not they were deemed eligible at this stage. Eligible households were subsequently visited to verify the information given. Program officials were expected to visit the potential beneficiaries in their home and to fill out a new questionnaire containing information on the same socioeconomic characteristics. This information was then processed back at the office, and the new eligibility status of the applicant was determined. Applicants were told to return to the module to confirm their eligibility status and be incorporated if selected. All program offices were closed at the end of August 2002, and households received their first transfers in November 2002 (see table 2 for details on the transfer schedule).⁴

B. Data Description

The data set used in this analysis is the baseline of the Urban Evaluation Survey of *Oportunidades* (2002), carried out between September and December

⁴ The program has essentially two objectives: (i) decreasing current income poverty through cash transfers and (ii) decreasing future poverty and breaking the intergenerational transmission of poverty by conditioning the receipt of transfers by eligible households on the accumulation of human capital. This study is concerned with the first of these objectives.

TABLE 1
VARIABLES AND WEIGHTS USED TO ESTIMATE THE DISCRIMINANT PROXY-MEANS SCORE

| Variables (x) | Definition | Coefficient |
|---------------|---|---|
| HACINA | Number of people/Number of rooms in the house | .139*HACINA |
| DEPEND | Total number of people in the household | .176*DEPEND |
| SEXO | The head of the household is a woman | -.02*SEXOJ |
| SS | Does not have access/right to medical service | .475 |
| NINOS | Total number of children <11 years | .255*NINOS |
| ESC* | Years of education of the household head (0 = never went to school or did not reach any level; 1 = primary education, first grade) | If (ESCJ1 = 1), mpESC = .380 If (ESCJ2 = 1), mpESC = .201 If (ESCJ1 = 0 & ESCJ2 = 0), mpESC = 0 |
| EDAD | Age of the head of the household | .005*EDADJ |
| BAO | BAO11 = do not have bath; BAO12 = have bath but without water | If (BAO11 = 1), mpBAO = .415 If (BAO12 = 1), mpESC = .22 If (BAO11 = 0 & BAO12 = 0), mpBAO = 0 |
| PISO | Floor is not paved (1/0) | .475 |
| ESTGAS | Do not have gas heating system (1/0) | .761 |
| REFRI | Do not have a refrigerator (1/0) | .507 |
| LAVA | Do not have washing machine (1/0) | .127 |
| VEHI | Do not have vehicle (no car nor truck) | .159 |
| RURURB | House is in rural area | .653 |
| REG | Region (19 regions) | Various dummy coefficients |
| CONS | Constant | -1.579 |

Note. The "weights" attached to variables were calculated by the program based on national household survey data using discriminant analysis. The underlying score was used to identify beneficiary households that applied to program offices.

of 2002 by the National Institute of Public Health (INSP). Two surveys were collected: a census survey of all households in a random selection of blocks in participating localities, and a sample survey of a subset of these households. The census sample was selected by first choosing a random sample of eligible localities (Urban Evaluation Survey, INSP, 2002). From this sample of localities, a stratified random sample of 149 blocks was selected in which blocks with a greater number of poor households were given a higher weight. A census survey of all 20,859 households in these blocks was carried out. This study focuses on the evaluation of targeting performance within these very poor urban blocks and is not necessarily representative of other urban populations.

Using the census information, the proxy-means score was calculated for each household and, based on this score, households were classified into three groups: Poor, Quasi-Poor (i.e., those just above the cutoff), and Nonpoor. Households were also asked if they were beneficiaries of the program. Then a stratified random sample of households, based both on these classifications and on self-reported beneficiary status, was chosen. For sampling purposes,

TABLE 2
TRANSFER LEVELS BY GRADE AND GENDER
(PESOS PER MONTH, 2002)

| | Boys | Girls |
|-----------------|------|-------|
| Primary school: | | |
| Grade 3 | 100 | 100 |
| Grade 4 | 115 | 115 |
| Grade 5 | 150 | 150 |
| Grade 6 | 200 | 200 |
| Middle school: | | |
| Grade 7 | 290 | 310 |
| Grade 8 | 310 | 340 |
| Grade 9 | 325 | 375 |
| High school: | | |
| Grade 10 | 490 | 565 |
| Grade 11 | 525 | 600 |
| Grade 12 | 555 | 635 |

Note. Education transfers are conditional on 85% school attendance. There is a cap on the amount households can receive in education grants: 1,680 pesos if the household has children attending high school, 915 otherwise. Households also receive a monthly "food transfer" of 150 pesos, conditional on regular attendance at health centers.

four groups were identified: (i) households reporting being beneficiaries, (ii) Poor nonbeneficiaries, (iii) Quasi-Poor nonbeneficiaries, and (iv) Nonpoor nonbeneficiaries.⁵

III. Methodology

To motivate our approach to evaluating the implications of self-selection and administrative selection for the overall targeting performance of the program, we present a simple model to develop a targeting evaluation statistic.⁶ Consider a government with a fixed social safety-net budget, B , to distribute across households by choosing from a range of alternative targeting approaches for selecting eligible households.⁷ Social welfare is specified as a standard Bergson-Samuelson function:

$$W[V^1(\mathbf{p}, m^1), \dots, V^b(\mathbf{p}, m^b), \dots, V^H(\mathbf{p}, m^H)],$$

⁵ For our analysis we derive the sampling weights by merging all census households with sample households and comparing populations in the sampling categories. We focus on the sample survey since only this collected consumption data and data on the targeting process (i.e., knowledge, application, and acceptance). The sample weights are adjusted for nonresponse in the sample based on the above four stratified sampling categories—of the 10,527 sampled households, only 9,817 households completed the sample questionnaires.

⁶ See Coady and Skoufias (2004) for a more detailed discussion of this approach.

⁷ Fixing the budget enables us to avoid the issue of how the marginal cost of funds changes with

where $V(\mathbf{p}, m)$ is the indirect utility function for households (denoted by superscript b), \mathbf{p} is the vector of commodity and factor prices faced by the household (which are assumed fixed), and m is lump-sum transfers from the government. A transfer program can be characterized by a vector $d\mathbf{m} = \{dm^b\}$ where $dm^b > 0$ for beneficiary households and $dm^b = 0$ for nonbeneficiary households. The social welfare impact of a transfer program is then

$$dW = \sum_b \frac{\partial W}{\partial V^b} \frac{\partial V^b}{\partial m^b} dm^b \equiv \sum_b \beta^b dm^b, \quad (1)$$

where β^b is the social valuation of extra lump-sum income to the household (i.e., the so-called welfare weight of each household).⁸ Multiplying and dividing the right-hand side of (1) by the program budget gives

$$dW = \sum_b \beta^b \frac{dm^b}{\sum_b dm^b} \sum_b dm^b \equiv \sum_b \beta^b \theta^b \sum_b dm^b \equiv \lambda B, \quad (2)$$

where θ^b is the share of the transfer budget going to each household. Note that λ is independent of the size of transfers in that scaling up transfers results in the same value and thus can be interpreted as a pure index of (vertical) targeting performance that increases as transfers are more concentrated on more “deserving” low-income households with higher welfare weights.⁹ Note also that λ is the same regardless of whether one focuses on households or individuals. For the former, one multiplies the transfer to each household in the sample by the population expansion factor for that household. For the latter, one multiplies per capita transfers by the individual expansion factor (i.e., the household expansion factor times household size), and this is equivalent to the former.

the size of the budget. We are essentially assuming that the budget is financed in a given way and is independent of how the budget is spent. For an empirical example of incorporating the cost of public funds across alternative financing instruments, see Coady and Harris (2004).

⁸ Note that from a social welfare perspective, since households are assumed to optimally choose their allocation of time between labor supply (or consumption) and leisure, we are indifferent between how households “consume” the transfer and thus can ignore the income effect on labor supply. We also abstract from the issue of disincentives (e.g., where households change labor supply in order to become eligible for the program or devote resources to increasing their probability of being deemed eligible through deceit or lobbying).

⁹ This is strictly only true for “small” transfers such that welfare weights are approximately constant. More generally, eq. (2) needs to be multiplied by an additional term capturing the impact of changing welfare weights. This term will be maximized for the optimal transfer program that differentiates transfers across households such that incomes are equalized at the bottom of the income distribution (subject to a budget constraint and only households with incomes below this new “minimum” income receiving transfers).

For a given budget, the program with the largest λ also delivers the largest increase in social welfare. Since $B = N \cdot m^*$, where N is the size of the program in terms of total number of beneficiary households and m^* is the average transfer level per beneficiary household, a decrease in the size of the program under a fixed budget requires an equi-proportionate increase in average transfers. Even if λ is constant across programs of different size, one should still compare the implications of program size for the coverage of targeted poor households. In general, for well-targeted programs, one expects λ to decrease with program size but the participation (or coverage) rate of the poor to increase.

In this study, we are concerned with identifying the implications of the different components of the targeting approach adopted for the overall targeting performance of the program. For this purpose, we treat the different stages as different programs. For example, the first stage of the targeting process is knowledge of the existence of the program and how to apply (KNOW), and this involves self-selection by households reflecting a decision to incur the cost of acquiring knowledge. The second stage is applying for entry into the program (APPLY), which involves self-selection by households reflecting a decision to incur application costs. The final stage is gaining entry into the program (ACCEPT), which involves administrative targeting based on a proxy-means test. In addition to these three stages, we also identify a fourth stage called demographic targeting (DEMOG) since, once households have been accepted, rather than getting a uniform transfer, transfers are differentiated according to the schedule set out in table 2. This helps to identify the implications for targeting performance of differentiating transfers according to demographic structure, conditional on the outcome of the first two stages of the process (i.e., knowledge acquisition and application stages).

We evaluate the targeting efficiency of the program at the various stages by comparing λ across these “programs.” So the first program (KNOW) is one where all households with knowledge of the program receive a uniform transfer, the second program (APPLY) is where all households applying for the program receive a uniform transfer, and the third program (ACCEPT) is where all households that were accepted receive a uniform transfer. The actual program is one where transfers are differentiated across accepted households (DEMOG). Each of these programs is compared to a reference program (NEUTRAL), taken here to be one without any targeting, that is, where all households receive a uniform transfer; under this program the “target” population will receive their population share.

Since we expect the combination of different targeting methods to improve targeting performance, the share of transfers accruing to the target population

should exceed their population share. We therefore compare the difference between λ for each program— $\lambda(\text{KNOW})$, $\lambda(\text{APPLY})$, $\lambda(\text{ACCEPT})$, and $\lambda(\text{DEMOG})$ —and that for the reference program, $\lambda(\text{NEUTRAL})$, across programs. We then divide these differences by the total increase in targeting performance, that is, $\lambda(\text{DEMOG}) - \lambda(\text{NEUTRAL})$, to calculate the share of the total improvement in targeting performance accounted for by each stage of the targeting process. For example, the share accounted for by the third stage (ACCEPT) is given by

$$\text{Contribution}(\text{ACCEPT}) = \frac{\lambda(\text{ACCEPT}) - \lambda(\text{APPLY})}{\lambda(\text{DEMOG}) - \lambda(\text{NEUTRAL})},$$

and the share of the other stages is derived by replacing the numerator accordingly.¹⁰ Note that these gains in vertical targeting performance may come at the expense of program coverage of the target poor population since households are eliminated at each stage of the process.¹¹ We therefore also document this trade-off and identify reforms that can reduce this trade-off.

We finish this section with a discussion of welfare weights. Underlying our objective of poverty alleviation must be the view that extra income to low-income households is more socially valuable than extra income to high-income households. Making this view explicit requires the specification of a set of “welfare weights,” and we expect these weights to decrease with the consumption (or welfare) level of the household. For example, based on Atkinson (1970), the welfare weight for each household (β^b) can be derived from a constant elasticity social welfare function as $\beta^b = (y^k/y^b)^\varepsilon$, where y refers to income, b superscript denotes the household in question, and k superscript denotes a reference household, which always has a weight of unity. The term

¹⁰ Note that it is logical to evaluate the stages in sequence (i.e., NEUTRAL then KNOW then APPLY then ACCEPT) given the inherent sequential nature of the participation process and the fact that we expect targeting to improve across stages. But clearly one could give demographically differentiated transfers at all stages instead of uniform transfers. However, this would mean that the transfers under ACCEPT would be the same as those under DEMOG. Since we expect (and find) that moving from uniform to demographically differentiated transfers increases targeting performance, comparing ACCEPT (=DEMOG) to a uniform transfer to accepted households would result in a decrease in targeting performance at this final stage. Therefore, given our focus on participation outcomes and on the relative contribution of each stage to targeting performance, it makes more sense for us to assume uniform transfers for the first three stages (KNOW, APPLY, ACCEPT) followed by a shift to differentiated transfers at the final stage (DEMOG).

¹¹ For uniform transfer programs of equal size in terms of number of beneficiaries, improved targeting performance comes automatically with improved coverage of the target poor population. More generally, for a fixed budget, improved targeting performance can be associated with either lower or higher coverage of the poor since these can be offset by a differences in the structure of transfers.

ε captures one's "aversion to inequality" of income and determines how the welfare weights vary (i.e., decrease) with household income. For example, a value of $\varepsilon = 0$ implies no aversion to inequality and all welfare weights take the value unity, that is, an extra unit of income to households is viewed as being equally socially valuable regardless of initial consumption levels. A value of $\varepsilon = 1$ implies that if household b has twice (half) the income of household k then its welfare weight is 0.5 (2.0) as opposed to unity for k . A value of $\varepsilon = 2$ similarly implies a welfare weight of 0.25 (4.0) for b .

As ε approaches infinity, the impact of the program on the welfare of the lowest income group dominates any evaluation, consistent with a Rawlsian maxi-min social welfare perspective where one cares only about how much of the program benefits are received by the poorest of the poor. For example, if we attach a welfare weight of unity to poor households and zero to nonpoor households, this being consistent with the welfare weights underlying the conventional poverty gap measure, then λ converges to the share of transfers accruing to the poor as ε approaches infinity. We use per capita household consumption as our welfare measure and also present the sensitivity of our findings to different sets of welfare weights based on different degrees of inequality aversion (i.e., different values of ε). We also consider how sensitive our results are to introducing adjustments for household size into our household welfare measure. Consistent with the program objectives, we consider only values of $\varepsilon > 0$.

Although the specific normalization of welfare weights is irrelevant when comparing across programs (since the ratio of λ s across programs is independent of this normalization), it can be useful to base the normalization on the reference program used for the evaluation of targeting performance. For example, if one normalizes welfare weights such that they sum to the number of households (i.e., by dividing welfare weights by the average welfare weight), then the λ for a universal uniform transfer program will be unity and the percentage gain in targeting performance for any specific program relative to no targeting (i.e., relative to a universal uniform transfer program) is simply 100 times $(\lambda - 1)$. Given our focus on the gains from targeting, we adopt this normalization convention.

IV. Results

To inform the presentation of results, it is useful to represent the participation outcome by the following sequential participation process set out by Heckman and Smith (2003):

$$P(\text{PART}) \equiv P(\text{KNOW}) \cdot P(\text{APPLY}|\text{KNOW}) \cdot P(\text{ACCEPT}|\text{KNOW, APPLY}),$$

where $P(\text{PART})$ is the (unconditional) probability of a household in the population participating in the program, $P(\text{KNOW})$ is the (unconditional) probability of a household knowing about the program, $P(\text{APPLY}|\text{KNOW})$ is the probability of a household applying for the program conditional on knowing about the program, and $P(\text{ACCEPT}|\text{KNOW}, \text{APPLY})$ is the probability of a household's being accepted into the program conditional on the household's knowing about the program *and* applying for the program. When evaluating targeting performance of the actual program, we are interested ultimately in the unconditional probabilities, that is, $P(\text{KNOW})$, $P(\text{APPLY})$, and $P(\text{ACCEPT})$. That is, we are solely interested in whether a household in the population is included or not at each stage of the process. However, an analysis of the conditional probabilities can inform us about how the targeting performance of the program is likely to evolve in response to various policy reforms.

Focusing first on the unconditional probabilities, out of the total population of households, 63% report knowing about the program, 47% applying for the program, and 30% being accepted as eligible. To the extent that the actual participation rate (i.e., the percentage of the population accepted) reflects the desired size of the program, we can therefore consider the bottom three income deciles as the target population.¹² For this reason we are particularly interested in the ability of the program to target the bottom of the income distribution and so pay particular attention to targeting performance at higher values of ε . Figure 1 describes how the above unconditional probabilities of knowing, applying, and being accepted vary across the income distribution—the positively sloped line is the cumulative income distribution. The unconditional probabilities of knowing, applying, and being accepted all decrease with household income, especially at higher income levels. This gives a clear indication that targeting is effective, in the sense that the share of transfers accruing to households will be greater than the population share of income groups. The closeness of the three probabilities, especially at the lowest income levels, also indicates that a very high percentage of those households that know about the program actually apply and a high percentage of those applying are accepted.

Figure 2 describes how the pattern of conditional probabilities varies across income groups. Conditional on knowing, a very high proportion of low-income households actually apply. For example, for the bottom income quintile, around 87% of those who know actually apply for the program. However, it is also noticeable that a high proportion of higher-income groups apply, conditional on knowledge. For example, in the top income quintile, around 45% of those

¹² Throughout the paper we will use “income” to denote our adopted household welfare measure (i.e., per capita consumption) and “welfare” to denote social welfare.

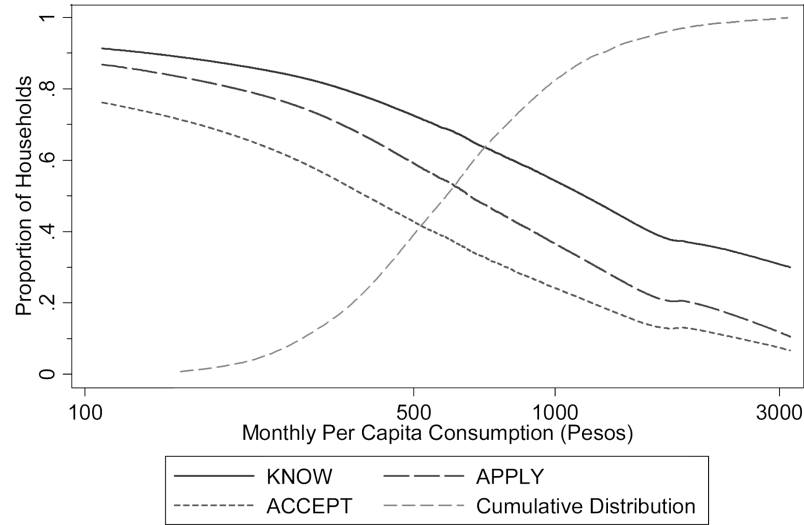


Figure 1. Unconditional probabilities

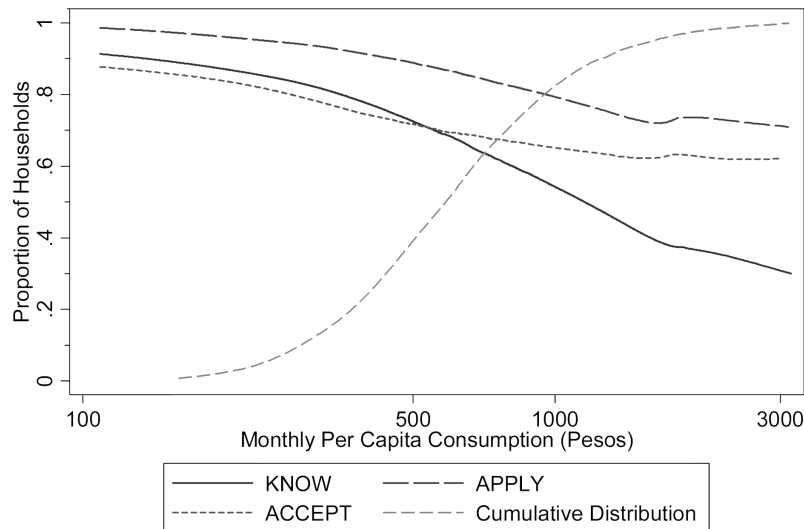


Figure 2. Conditional probabilities

who know apply for the program. A similar pattern holds for the probability of being accepted conditional on applying, which does not decrease much across income levels, being 74% for the bottom quintile and 51% for the top. In other words, the slopes of the unconditional probabilities in figure 1 appear to be driven mainly by the slope of the probability of acquiring knowledge. However, the fact that such a small proportion of higher-income groups actually

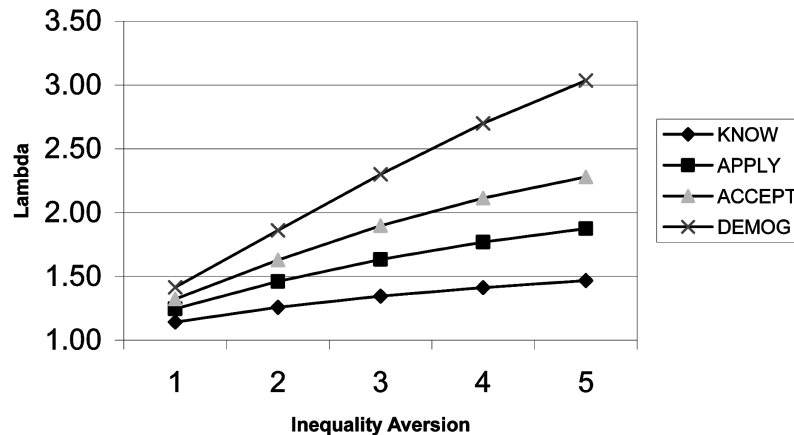


Figure 3. Targeting performance by stage (for different degrees of inequality aversion, ϵ)

make it to applying—only 8% of the top quintile apply—means that these high conditional probabilities of acceptance do not translate into substantial leakage of benefits to these groups. But the high conditional probabilities of applying and being accepted among higher-income groups do raise concerns regarding how targeting performance might be affected by universal knowledge.

A. Targeting Performance across Participation Stages

Each targeting stage reduces program coverage until it reaches the desired 30% coverage as dictated by the program budget. Effective targeting reduces coverage of the nontarget population while simultaneously maintaining high coverage of the target population. Figure 3 presents the targeting performance for each of the program stages for alternative degrees of aversion to inequality (i.e., $1.0 \leq \epsilon \leq 5.0$). For all values of ϵ , targeting performance improves over all stages, that is, the lines in figure 3 shift upward across stages. In other words, the share of transfers going to the target population (as captured by choice of ϵ) increases at each stage. Therefore, for a given target population, the trade-off between decreasing program size (and decreasing leakage) and decreasing coverage of the target population improves across stages.

But one expects that the magnitude of the above trade-offs may vary according to the target group (i.e., the choice of ϵ). Figure 4 presents the share of the total improvement in targeting performance accounted for by each of the stages across various values for ϵ . Since the magnitude of these shares depends in part on the baseline, which here is the NEUTRAL program, we do not attach too much relevance to the level of the share and focus more on

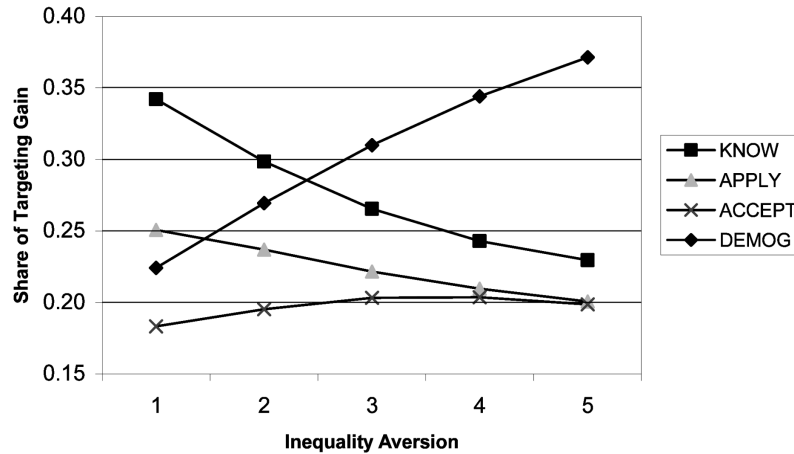


Figure 4. Share of targeting performance by stage (for different degrees of inequality aversion, ϵ)

how shares vary across values of ϵ .¹³ The fact that the share of KNOW decreases rapidly for higher values of ϵ indicates that the distribution of knowledge of the program plays a more important role in reducing leakage to the highest income households and is less important in determining participation in, say, the bottom half of the distribution. This is consistent with figure 1, where we saw that knowledge decreased more rapidly at higher-income groups. More interesting, however, is the increase in the share of ACCEPT as ϵ increases. Leakage to higher-income households decreases over all stages, but the cost of reducing leakage in terms of increased undercoverage of the lowest income households is smaller for ACCEPT than for APPLY. In other words, in the context of *Oportunidades*, administrative selection based on proxy-means testing plays a crucial role in further reducing program coverage without substantially reducing coverage of the lowest income groups.

The share of DEMOG also increases substantially with ϵ and is also always above the share of ACCEPT and even more so at the higher values of ϵ . This indicates that the differentiation of transfers across participating households substantially increases the share of transfers going to the lowest income households compared to uniform transfers. Our analysis of the different components of the transfer (see table 2) showed that primary and middle school transfers are clearly better targeted. Redistributing the budget from high school transfers

¹³ For example, if we choose as our reference a program that was better targeted than NEUTRAL, then the improvement in targeting performance under KNOW would be smaller. It is also the case that targeting based on knowledge may not be desirable in that one may not want participation to be determined by ignorance. Since the qualitative nature of the results does not change when we use KNOW as our reference, we discuss the results when NEUTRAL is our reference because it facilitates a discussion of the implication of knowledge for targeting outcomes.

to middle, or even more so, primary school transfers would increase targeting performance. However, given the broader objectives of the program, the implications of such a redistribution for human capital impacts would need to be considered.

We performed a number of statistical tests using bootstrapping techniques based on generating λ for 2,000 drawings from the sample with replacement. First, we tested whether targeting performance increases across stages for a given value of ε (i.e., whether the lines in fig. 3 shifted upward with each stage) and found that this held true in all of the drawings. Second, we tested whether targeting performance for each of the stages increases with ε (i.e., whether the lines in fig. 3 were always positively sloped) and found that it did so for all of the drawings. Third, we tested whether the increase in targeting performance as ε increases differed across stages (i.e., the lines in fig. 3 fanned out as ε increased). Again we find that this almost always holds true for all comparisons over increasing values of ε .

We also examined the sensitivity of our results to adopting an alternative measure of household welfare that accounted for economies of scale. This is intended to adjust our welfare measure for the possibility that the cost of achieving a given welfare level differs by household size. The presence of such economies can reflect, for example, the fact that household members benefit from each other's consumption or that there are public goods that can be used by all household members at zero additional cost. In particular, such an adjustment can be expected to affect the contribution of demographically structured transfers to the targeting performance of the program since this results in higher transfers to households with children, and larger households tend to reflect the presence of children.

The standard way of adjusting for economies of scale is to use as one's welfare measure total household consumption divided by household size raised to the power of γ , where γ is a measure of economies of scale (Lanjouw and Ravallion 1995; Deaton 1997). A value of $\gamma = 1$ results in consumption per capita as the welfare measure (i.e., no economies of scale), with $\gamma < 1$ implying positive economies of scale. Therefore, if one doubles household size and also doubles household resources, then households would experience a 100 times $(1 - \gamma)$ percent increase in effective per capita resources (or welfare). Deaton and Paxton (1998) find that most estimates of γ fall in the range 0.6–0.8. We redid the above analysis using a value of $\gamma = 0.7$ and found that the contribution of demographically structured transfers to targeting performance decreased substantially from the earlier 22%–37% for various values of ε to 9%–18% (fig. 5). However, the results regarding the increasing importance of demographic transfers at higher values of ε still hold so that switching from uniform to

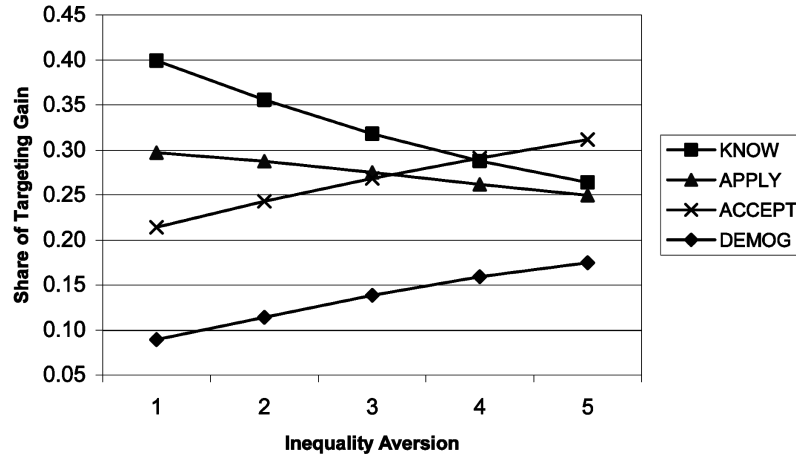


Figure 5. Share of targeting performance by stage (for different degrees of inequality aversion, ϵ , and allowing for economies of scale).

demographic transfers is particularly effective at targeting transfers at the poorest households. But now the acceptance stage seems equally important at achieving this end.

B. Policy Reform Simulations

Although vertical targeting performance improves across stages, this is also obviously associated with a decrease in population coverage, including coverage of poor households. The top panel of table 3 shows the trade-off across program stages between coverage of households in the bottom three income deciles (designated as “Poor” for presentational purposes) and vertical targeting performance (as measured by the share of transfers accruing to these same deciles). Although coverage of these poor households (the “Poor” participation rate, PPR) does not decrease as quickly as the total participation rate, it still falls substantially from 100% to 47%, while the poor’s share of transfers increases simultaneously from 30% to 52%. Out of the total undercoverage rate of 53%, 22% occurs at the knowledge acquisition stage, a further 12% at the application stage, and a further 19% at the acceptance stage.

To further reinforce the importance of including undercoverage when evaluating the welfare impact of a transfer program, we include two additional welfare indices in table 3. First, based on equation (2), the welfare impact of a program can be calculated as $W = \lambda \cdot B = \lambda \cdot N \cdot m^*$, where B is the size of the program budget, N is the number of program beneficiaries, and m^* is the average transfer level per beneficiary. If the {“Poor,” “Nonpoor”} are assigned welfare weights of {1, 0}, then the transfer share of the poor is a measure of

TABLE 3
TRADE-OFF BETWEEN VERTICAL TARGETING PERFORMANCE AND PROGRAM COVERAGE (%)

| | Alternative Programs | | | |
|--|----------------------|------|-------|--------|
| | NEUTRAL | KNOW | APPLY | ACCEPT |
| Current program: | | | | |
| Total participation rate (TPR) | 100 | 63 | 47 | 30 |
| "Poor" participation rate (PPR) | 100 | 78 | 66 | 47 |
| "Nonpoor" participation rate (NPR) | 100 | 54 | 36 | 21 |
| Average beneficiary transfer (pesos per month) | 145 | 246 | 344 | 534 |
| Transfer share of "Poor" | 30 | 40 | 47 | 52 |
| Welfare impact (transfer share \times TPR) | 30 | 25 | 22 | 16 |
| Targeting differential (TD = PPR – NPR) | .00 | .24 | .30 | .26 |
| Universal knowledge: | | | | |
| Total participation rate (TPR) | 100 | 100 | 79 | 30 |
| "Poor" participation rate (PPR) | 100 | 100 | 97 | 52 |
| "Nonpoor" participation rate (NPR) | 100 | 100 | 68 | 18 |
| Average beneficiary transfer (pesos per month) | 145 | 145 | 199 | 534 |
| Transfer share of "Poor" | 30 | 30 | 40 | 57 |
| Welfare impact (transfer share \times TPR) | 30 | 30 | 32 | 17 |
| Targeting differential (TD = PPR – NPR) | .00 | .00 | .29 | .34 |

Note. The focus on the poorest 30% of households reflects the actual size of the program; for presentational purposes, the bottom 30% are referred to as "Poor" and the top 70% as "Nonpoor."

targeting performance, which when multiplied by the program budget and assuming transfers are never large enough to pull any household above the poverty line, is equivalent to the decrease in the absolute poverty gap. If the program budget remains constant, then increasing targeting performance will always increase the welfare impact of the program. However, if the budget falls in line with the number of program beneficiaries (i.e., N), the welfare impact could decrease. From table 3 we can see that the overall welfare impact decreases at the APPLY and ACCEPT stages compared to previous stages, reflecting the fact that the increase in targeting performance is more than offset by the decrease in program size.

The second welfare index we use is the targeting differential, defined as the difference between the program participation rate of the poor (PPR) and the nonpoor (NPR).¹⁴ This index takes the value +1 when all of the transfers go to the poor and all the poor receive transfers, –1 when all of it goes to the nonpoor and all the nonpoor receive transfers, and 0 for a universal transfer. It can be interpreted as capturing both targeting performance as well as pro-

¹⁴ See Ravallion (forthcoming) for a discussion of the targeting differential and how it relates to other indicators, and Coady and Skoufias (2004) for a comparison of the index used in this paper with other similar indicators.

gram coverage and attaches equal weight to a percentage increase in the PPR and a percentage decrease in the NPR.¹⁵ From table 3 we see that this index also decreases at the final acceptance stage, reflecting the fact that the decrease in the PPR is greater than the decrease in the NPR.

Increasing knowledge of the program is obviously a key policy issue regarding increasing coverage of poor households. However, the implications for vertical targeting performance are less clear since this will depend on how increases in application and acceptance rates vary across income groups. To get a sense of the trade-off involved, for each household we estimate the conditional probabilities of each household applying as well as the conditional probabilities of acceptance. Because these conditional probabilities need to be estimated on the selected samples of those households that reported knowing (for estimating conditional application probabilities) and applying (for conditional acceptance probabilities), our estimation strategy needs to correct for the potential presence of sample selection bias. A more detailed account of how we address this issue and the regression results are presented in the appendix.

The bottom panel of table 3 presents the results from our simulations, which can be compared to the results under existing knowledge levels presented in the top panel. Under universal knowledge, the total proportion of the population applying for the program increases from 47% to 79%. Although coverage of the poorest three income deciles increases from 66% to 97%, the share of this group in total transfers decreases from 47% to 40%, reflecting a proportionally higher increase in application rates among higher-income groups. This decrease in targeting performance is also captured in the values of λ , which fall from 1.25–1.88 to 1.17–1.34 (over various values of ε). Because the size of the program is kept fixed, at the acceptance stage the participation rate stays at 30%. However, the composition of this 30% changes, with the coverage of the poorest three deciles increasing from 47% to 52% and their share of transfers (which for uniform transfers is equivalent to their share of beneficiaries) increasing from 52% to 57%. This is also reflected in the targeting performance with the range for λ increasing from 1.32–2.28 to 1.39–2.44 across various values of ε . With fixed program size, improved vertical targeting performance is mirrored by improved coverage of poor households.

¹⁵ For a constant program size, N , there is obviously a well-defined relationship between the PPR and the NPR, with a unit increase in the PPR implying a decrease in NPR equal to the ratio of the share of the poor in the population to the share of the nonpoor (or, equivalently, the poverty rate divided by one minus the poverty rate). Therefore, for such a program, welfare will always increase with PPR.

While increased application rates (reflecting increased knowledge) facilitate both improved vertical targeting performance and improved coverage of the poor, still substantial undercoverage of the poorest three deciles remains, with nearly half of these households excluded from the program. This partly reflects the fact that the program threshold in terms of the underlying proxy-means score was set too high relative to the program budget and that program officials, as indicated by our regression analysis, were clearly using household characteristics other than their proxy-means score to ration program places. One approach to improving targeting performance may then be to ensure a more strict application of the proxy-means score to ration places. However, applying this rule to those households that applied would increase targeting performance only marginally, with λ increasing from 1.39–2.44 to 1.41–2.57, with this approach only performing better in terms of getting some more of the very poorest households into the program. For example, the coverage of the poorest three deciles increases only from 52% to 53%. This, in turn, partly reflects the imperfect correlation of this proxy-means score with our measure of household welfare (i.e., per capita consumption).

To get a sense of the potential for improving this correlation, we derived a score for each household based on regressing per capita consumption on the same household variables used to derive the actual proxy means score. Under this approach, λ increases more substantially to 1.52–3.00 and the coverage rate of the poorest three deciles increases to 60%. This could possibly be further improved through including household characteristics that are better correlated with welfare in such a proxy-means algorithm.

Another possible approach would be to change the structure of the benefits so that they are more negatively correlated with income, for example, by linking them directly to the proxy-means scoring system. Such an adjustment can be expected to improve targeting performance through two separate channels. First, for a given pattern of participation across income groups, it increases transfers to lower-income households and decreases transfers to the higher-income households, thus decreasing the welfare cost of erroneously accepting the latter. Second, it increases the probability of receiving applications from lower-income households and decreases it for higher-income households. This makes it more likely that lower-income households push out higher-income households at the acceptance stage.

To illustrate the magnitudes of this second channel in the context of the current program, we estimated the reduced form participation regression including benefit levels (t , in thousand pesos) and benefits squared along with the other variables included in the conditional probability regressions described

in the appendix. The estimated relationship between the participation probability and these variables was

$$P(\text{PART}) = 0.1195 \, t - 0.0649 \, t^2,$$

where the coefficients on t and t^2 were both significant at the 5% significance level. The estimated impact of a 100 peso per month adjustment in transfers for households with transfer levels initially at 450 pesos (coinciding approximately with the median transfer level) is 0.6 of a percent. Reducing the transfer for such a nonpoor household and switching it to such a poor household would thus increase the gap in participation probability between these households from zero to 1.2%. This gain would be in addition to the targeting gains from differentiating transfers across beneficiaries. In addition, since these gains arise through the application stage, the cost associated with processing applications from nonpoor households would be reduced.

However, it is important to recognize that the linking of transfers to poverty scores can provide an additional incentive for households to misreport their socioeconomic characteristics in an attempt to both increase their probability of being eligible and the level of benefit received if deemed eligible. In the context of the *Oportunidades* program, Martinelli and Parker (forthcoming) find that misreporting was widespread. Although the probability of a household misreporting increased with the level of transfer they would receive if deemed eligible, the magnitude of this relationship was small. This could reflect the belief by households that while misreporting would enhance the probability that they would be deemed eligible, it would not affect the level of transfer received. Clearly introducing such a link between scores and transfers could provide further incentives for misreporting. Within a fixed budget, increasing transfers for households with high poverty scores increases the probability of misreporting both due to its positive impact on the probability of being deemed eligible and its positive impact on the level of transfer received. The net impact of lowering transfers for those with lower poverty scores could in principle go either way; introducing differentiation increases the incentives to misreport while the lower benefit received decreases the incentive. But the possibility of exacerbating the misreporting problem further reinforces the importance of having an effective verification process, for example, both through home visits and the use of easily verifiable socioeconomic characteristics in the proxy-means algorithm.

Further reductions in undercoverage of poor households could be achieved through expansion in the size of the program. But this is likely to be associated with a decrease in the targeting performance of the program. However, adjusting the benefits structure by linking benefit levels directly to proxy-means

scores will help to reduce the adverse trade-off between higher coverage of the poor and lower targeting performance. The corollary of this is that the budget cost of a given reduction in undercoverage is reduced.

V. Summary

The unique information collected in the evaluation of the *Oportunidades* program enables us to evaluate the implications of the different components of the targeting strategy, more specifically, the relative contributions of self-selection by households at the knowledge acquisition and application stages and of administrative selection from households at the acceptance stage. The results of our evaluation are also likely to have valid lessons for similar programs being implemented throughout the region as well as for enhancing the potential role for self-selection in programs with strong administrative selection components.

A key finding is that, relative to self-selection at the household application stage, administrative selection based on proxy-means testing is particularly effective at reducing overall program coverage while maintaining high coverage of the lowest welfare households. While the information strategy used to disseminate knowledge of the program was effective at reducing coverage among the highest welfare households, it had little role in increasing the concentration of transfers on the lowest welfare households. In any case, reducing program coverage through ignorance is unlikely to be a desirable feature of targeted safety net programs. Our results also indicate that the differentiation of transfers across households based on demographic characteristics is also effective at concentrating program resources on the poorest households, reflecting the fact that these households typically have a relatively larger number of children and young adults.

Reducing the undercoverage of poor households requires, first and foremost, that the issue of knowledge constraints be addressed, which in turn will involve increased knowledge among the nonpoor population. Our results show that the importance of proxy-means targeting is reinforced in the context of universal knowledge, reflecting the high application rates among high-income households conditional on knowledge. Improvements in the proxy-means algorithm to increase its correlation with welfare can help to further decrease undercoverage and leakage. Tying transfer levels directly to these scores (e.g., by increasing the benefits for those with lower scores and decreasing them for those with higher scores) can similarly improve targeting performance both by reducing the leakage of benefits to higher-income groups wrongly deemed eligible by the proxy-means algorithm and by improving the contribution of self-selection at the application stage to overall targeting performance. This

could also save on resources allocated to processing applications by higher-income households. However, since the linking of transfer levels to poverty scores may provide additional incentives to households to misreport their socioeconomic characteristics, the importance of having an effective verification process (e.g., via home visits and including easily verified characteristics in the proxy-means algorithm) is reinforced.

Finally, addressing any remaining undercoverage of poor households will require an expansion of the program, which is likely to be associated with an adverse trade-off in terms of targeting performance. This trade-off can also be reduced by improving the underlying proxy-means algorithm and by linking transfers directly to the proxy-means score (assuming an effective verification process). Maintaining the average transfer levels for those with higher poverty scores and lowering the transfers for those with lower scores can generate budgetary savings to expand the program and possibly reduce incentives for misreporting.

Appendix

Details of Simulations Estimating Targeting Implications of Universal Knowledge

In this appendix we describe the methodology used to simulate the implications of universal knowledge for application and acceptance outcomes. In the text we represented the outcomes observed in our data through the following participation process:

$$P(\text{PART}) = P(\text{KNOW}) \cdot P(\text{APPLY}|\text{KNOW}) \cdot P(\text{ACCEPT}|\text{KNOW}, \text{APPLY}),$$

where $P(\text{PART})$ is the unconditional probability of being observed participating in the program, $P(\text{KNOW})$ is the unconditional probability of knowing about the program, $P(\text{APPLY}|\text{KNOW})$ is the probability of applying conditional on knowing, and $P(\text{ACCEPT}|\text{KNOW}, \text{APPLY})$ is the probability of being accepted into the program conditional on knowing about the program and applying for the program.

We are interested in estimating the probability that a household that knows about the program will apply, that is, $P(\text{APPLY}|\text{KNOW})$. One approach would be to regress the observed application outcome on household characteristics for those households that reported knowing, and then applying this model to households that did not know. However, to the extent that knowledge acquisition reflects household decisions, one expects that those households that report knowing about the program are a selected sample, for example, that they have unobserved characteristics that make them more likely to apply. Therefore, estimating $P(\text{APPLY}|\text{KNOW})$ on the sample of those reporting

knowing is likely to give an upwardly biased estimate of $P(\text{APPLY} | \text{KNOW})$ so that some correction for this bias is required.

One can think of the knowledge acquisition stage as being determined by household decisions based on the perceived costs and benefits of acquiring knowledge as well as by the program implementation strategy, for example, the advertising strategy, the latter being exogenous to household decisions.¹⁶ Similarly, one can think of a household's decision to apply as reflecting the perceived costs and benefits of applying. If one controls for the benefits a household would receive if accepted, then the coefficients on the remaining variables can be interpreted as capturing differential costs of applying. However, the problem is that the variables that determine the cost of acquiring knowledge are likely to also determine the cost associated with applying. To correct for possible sample selection bias one needs to have some independent variation in the knowledge (selection) equation, that is, a variable that affects the probability of applying only through its impact on the probability of acquiring knowledge. Information on the intensity of advertising in the block in which a household resides (e.g., resources devoted to advertising) would be an obvious candidate for such a variable. Although we do not have such information, we do know that advertising was concentrated in blocks with the highest poverty rates as estimated by the program's proxy-means algorithm. Therefore, we include the block poverty rate as an explanatory variable in the knowledge regression equation and expect higher poverty to be associated with higher knowledge (via a greater concentration of advertising resources on these blocks).

The first four columns of results in table A1 present the regression results of a Heckman sample selection estimation approach. The results are from a two-step estimation of a linear probability model. This provides consistent estimates and facilitates easy interpretation of results, but it was also the case that more complex full-information models based on this model and a probit model failed to converge, presumably reflecting the low percentage of households that do not apply conditional on having knowledge. As expected, the coefficient on the block poverty rate in the knowledge equation is positive and highly significant, indicating that poor households living in nonpoor blocks are less likely to know about the program. The significant coefficient on the inverse Mills ratio reinforces our concerns regarding the need to correct for sample selection bias.

The predicted probabilities from this model are used to identify households

¹⁶ For more detailed discussions on modeling program participation decisions within a cost-benefit framework, see Moffit (1983), Cowell (1986), Blundell, Fry, and Walker (1988), Duclos (1995), and Pudney, Hernandez, and Hancock (2002).

TABLE A1

| | Heckman Application Regression | | | | Heckman Acceptance Regression | | | | Consumption Regression | |
|----------------------------|--------------------------------|-------------|--------------------|-------------|-------------------------------|-------------|---------------------|-------------|------------------------|-------------|
| | P(APPLY KNOW) | | Selection: P(KNOW) | | P(ACCEPT APPLY) | | Selection: P(APPLY) | | Coefficient | t-Statistic |
| | Coefficient | t-Statistic | Coefficient | t-Statistic | Coefficient | t-Statistic | Coefficient | t-Statistic | | |
| Head of household: | | | | | | | | | | |
| Education (years) | .0007 | 1.98 | .0030 | 2.62 | -.0001 | -.14 | .0042 | 3.74 | -.630 | -2.31 |
| Gender | .0037 | .37 | -.1483 | -4.26 | .0464 | 3.29 | -.1319 | -3.86 | -11.49 | -1.39 |
| Age (years) | -.0007 | -1.43 | -.0064 | -3.93 | .0006 | .82 | -.0074 | -4.54 | 1.34 | 3.40 |
| Indigenous | .0003 | .03 | .2066 | 4.26 | .0238 | 1.31 | .2007 | 4.30 | -35.31 | -3.18 |
| Household characteristics: | | | | | | | | | | |
| Number of rooms | | | | | -.0364 | -4.55 | | | 28.98 | 7.50 |
| Children 0–5 years | .0182 | 2.68 | .1205 | 5.19 | .0021 | .22 | .1520 | 6.63 | -30.55 | -5.80 |
| Children 6–11 years | .0165 | 2.85 | .1945 | 10.32 | .0267 | 3.11 | .2199 | 11.79 | -28.79 | -6.60 |
| Children 12–17 years | .0094 | 1.27 | .1087 | 4.37 | .0104 | 1.11 | .1178 | 4.79 | -6.81 | -1.32 |
| Family size | -.0144 | -3.22 | -.0808 | -5.63 | .0098 | -1.55 | -.1047 | -7.32 | -74.92 | -22.92 |
| Disabled | .0190 | .61 | .0282 | .28 | -.0102 | -.23 | .0541 | .53 | -103.08 | -4.21 |
| No bath | | | | | .0102 | .54 | | | -78.53 | -8.51 |
| Bath without water | | | | | .0128 | .53 | | | -25.04 | -1.95 |
| Health insurance | | | | | -.0625 | -1.32 | | | 144.40 | 4.52 |
| Dirt floor | | | | | .0417 | 3.19 | | | -22.13 | -2.84 |
| Vehicle | .1744 | 6.04 | .5134 | 7.64 | .0084 | .17 | .7308 | 9.66 | -210.77 | -13.00 |
| Fridge | | | | | -.0136 | -.98 | | | 62.35 | 8.28 |
| Washing machine | | | | | .0225 | 1.16 | | | 96.53 | 9.61 |
| Gas heater | | | | | .0117 | .79 | | | 48.39 | 5.43 |

TABLE A2
APPLICATION OUTCOMES UNDER UNIVERSAL KNOWLEDGE

| | Prereform | | Universal Knowledge: % Apply (Threshold for Application Decision) | | |
|--------------|-------------------|-----------|--|----------|----------|
| | Sample Households | Share (%) | t = .5 | t = .886 | t = .908 |
| Full sample: | | | | | |
| KNOW = 0 | 3,649 | 37.8 | 100 | 80.5 | 68.3 |
| KNOW = 1 | 6,017 | 62.2 | 100 | 94.5 | 88.6 |
| All | 9,666 | 100.0 | 100 | 89.2 | 81.0 |
| Know = 1: | | | | | |
| APPLY = 0 | 685 | 11.4 | 100 | 83.7 | 72.6 |
| APPLY = 1 | 5,332 | 88.6 | 100 | 95.8 | 90.8 |

Note. KNOW = 1 refers to those households that in the prereform situation knew about the program. APPLY = 1 refers to those households that in the prereform situation both knew of and applied for the program.

that are likely to apply for the program conditional on universal knowledge. Table A2 presents the implications for application outcomes. When we use the traditional 0.5 probability as a threshold for identifying those households that apply we find that all households apply, reflecting the fact that all households have a predicted probability above 0.6. This is not unusual in samples with a small percentage of households not applying. One approach in such circumstances is to use the average application rate for the selected sample as the cutoff threshold (see Wooldridge [2005] for further discussion); the second column of table A2 shows that 88.6% of households that knew actually applied. Using 0.886 as our threshold, we find that 89.2% of the sample apply. However, one expects that those that previously did not acquire knowledge are less likely to apply compared to those that did. This pattern is consistent with our estimates, which show that 81% of the latter applied compared to 95% of the former. Thus, the fact that the proportion of applying across all households increases from 87% to 89% is worrying. We therefore use an alternative approach that chooses the threshold such that the proportion of households that previously knew (i.e., for which KNOW = 1 prereform) that applied postreform is the same as the proportion that applied prereform (i.e., 88.6%). This threshold was derived as 0.908 by trial and error. Under this approach, the proportion of all households that apply under universal knowledge is 81%, and the pattern across subgroups is as expected, with nearly 91% of those that previously applied now applying and only 68% of those who previously did not acquire knowledge now applying.

The issue of sample selection bias also arises when estimating the acceptance probability. We view acceptance as being based on administrative selection. We know from discussions with program officials that since a larger number of households that turned up at the program offices were classified as poor

under the proxy-means test than was budgeted for, program officials had to use additional criteria to ration program places. One obvious candidate for rationing is the score itself—among those classified as poor under the proxy-means system, one selects households with the highest poverty score. We therefore include quintile dummies based on this score, classifying households as extremely poor, moderately poor, quasi-poor, quasi-nonpoor, and nonpoor to allow for strong nonlinearities. Another possibility is that program officials attach a higher weight to components of the score than is implicit in the score itself. To capture this we include the variables included in the score. Other household characteristics associated with the overall motivation of the program (e.g., having infants or children of school-going age or being classified as indigenous) are also included. In addition, we include as an additional regressor a variable capturing the proportion of poor households visiting the program office that received a verification visit. According to program officials, once the quota for beneficiaries was reached, verification visits were suspended, so this variable should pick up the intensity of the budget constraint. In the unconditional application selection equation we include the same variables as before but now also include the instrumental variable used in the knowledge selection equation above since the unconditional application outcome depends on both the knowledge acquisition outcome and the conditional application outcome, that is, $P(\text{APPLY}) = P(\text{KNOW}) \cdot P(\text{APPLY}|\text{KNOW})$. Thus $P(\text{APPLY})$ is estimated on the full sample to get an unbiased estimate of $P(\text{ACCEPT}|\text{APPLY})$ for all households. Note also that the level of benefits a household would receive if accepted is included in the selection (application) equation but not the acceptance equation, reflecting the fact that acceptance is based on administrative selection and not on household decisions. This variable thus also acts as an additional instrumental variable in the application selection equation.

Columns 5–8 in table A1 present the results of a Heckman two-step linear probability acceptance regression. In the application selection equation the coefficient on the block poverty rate is again positive and highly significant. Similarly, the effect of transfers is significantly positive and decreasing. The probability of being accepted is much higher for those with lower compared to higher proxy-means scores consistent with the score being used to ration program places in the presence of a tighter budget constraint. The coefficient on the number of children of primary school age is also significantly positive, and the probability of being accepted is also negatively associated with housing conditions as captured by the number of rooms suggesting a higher weight being given to this by program officials than under the score. Note, however, that the inverse Mills ratio is significant only at around the 90% level.

To identify households that are accepted under the reformed universal knowl-

edge program, we assume that the program size is still binding at 30% of the population. The 30% of households with the highest probability of being selected (conditional on applying, which determined by the earlier simulation of $P(\text{APPLY}|\text{KNOW})$) are assumed to participate so that newly applying households can now “bump out” households that were previously selected. We also compare this to a situation in which only the proxy-means score was used to select households. Finally, we compare both of these to selection based on an alternative proxy-means algorithm based on the consumption regression presented in the final two columns of table A2.

References

- Atkinson, Anthony. 1970. “On the Measurement of Inequality.” *Journal of Economic Theory* 2:244–63.
- Blundell, Richard, Victoria Fry, and Ian Walker. 1988. “Modelling the Take-up of Means-Tested Benefits: The Case of Housing Benefit in the United Kingdom.” *Economic Journal* 98:58–74.
- Coady, David. 2006. “The Welfare Returns to Finer Targeting: The Case of the PROGRESA Program in Mexico.” *International Tax and Public Finance* 13:217–39.
- Coady, David, and Rebecca Harris. 2004. “Evaluating Transfer Programs within a General Equilibrium Framework.” *Economic Journal* 114:778–99.
- Coady, David, Margaret Grosh, and John Hoddinott. 2004a. “Targeting Outcomes, Redux.” *World Bank Research Observer* 19, no. 1:61–85.
- . 2004b. *The Targeting of Transfers in Developing Countries: Review of Lessons and Experiences*. Washington, DC: IFPRI and World Bank.
- Coady, David, and Emmanuel Skoufias. 2004. “On the Targeting and Redistributive Efficiencies of Alternative Transfer Instruments.” *Review of Income and Wealth* 50, no. 1:11–27.
- Cowell, Frank. 1986. “Welfare Benefits and the Economics of Take-up.” TIDI Discussion Paper no. 89, STICERD, London School of Economics.
- Deaton, Angus. 1997. *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*. Baltimore: Johns Hopkins University Press.
- Deaton, Angus, and Cristina Paxton. 1998. “Economies of Scale, Household Size and the Demand for Food.” *Journal of Political Economy* 106, no. 5:897–930.
- Duclos, Jean-Yves. 1995. “Modelling the Take-up of State Support.” *Journal of Public Economics* 58:391–415.
- Heckman, James, and James Smith. 2004. “The Determinants of Participation in a Social Program: Evidence from a Prototypical Job Training Program.” *Journal of Labor Economics* 22:243–98.
- Lanjouw, Peter, and Martin Ravallion. 1995. “Poverty and Household Size.” *Economic Journal* 105:1415–34.
- Martinelli, Cesar, and Susan Parker. Forthcoming. “Deception and Misreporting in a Social Program.” *Journal of the European Economic Association*.
- Micklewright, John, Aline Coudouel, and Susan Marnie. 2004. “Targeting and Self-

- Targeting in a New Social Assistance Scheme." Discussion Paper no. 1112, Institute for the Study of Labor (IZA), Bonn.
- Moffit, Roger. 1983. "An Economic Model of Welfare Stigma." *American Economic Review* 73, no. 5:1023–35.
- Pudney, Stephen, Maria Hernandez, and Ruth Hancock. 2002. "The Welfare Cost of Means-Testing: Pensioner Participation in Income Support." Unpublished manuscript, Department of Economics, University of Leicester.
- Ravallion, Martin. Forthcoming. "How Relevant Is Targeting to the Success of an Antipoverty Program?" *World Bank Research Observer*, forthcoming.
- Skoufias, Emmanuel, Benjamin Davis, and S. de la Vega. 2001. "Targeting the Poor in Mexico: An Evaluation of the Selection of Households into PROGRESA." *World Development* 29, no. 10:1769–84.
- Wooldridge, John. 2005. *Introductory Econometrics: A Modern Approach*. Cincinnati, OH: South-Western College Publications.