FCND DISCUSSION PAPER NO. 117

AN EVALUATION OF THE DISTRIBUTIONAL POWER OF *PROGRESA*'S CASH TRANSFERS IN MEXICO

David P. Coady

Food Consumption and Nutrition Division

International Food Policy Research Institute
2033 K Street, N.W.
Washington, D.C. 20006 U.S.A.
(202) 862–5600
Fax: (202) 467–4439

July 2001

ABSTRACT

Using both national-sample and program-level census survey data, we evaluate the distributional power of Mexico's Programa Nacional de Educacion, Salud y Alimentacion (PROGRESA) transfers using the so-called distributional characteristic. These transfers are targeted both geographically at marginal localities and at poor households within these localities. Transfers are also conditioned on household members attending school and health clinics. We show that the program has a relatively high distributional power compared to a range of alternatives considered. Although geographic targeting has a relatively large effect on the distributional power of the program, the demographic structure of transfers is more important than household targeting. However, the gains from household targeting increase as the program expands into less marginal localities. Within the structure of transfers, the education component is distributionally more powerful than the food component, reflecting the fact that the former is based on household demographics, while the latter is uniform across households. Restructuring education grants towards secondary schooling in order to generate higher education impacts does not appear to affect the distributional power of the program. In any case, any adverse impact could be offset by increasing the cap on transfers, which is regressive. Take-up of the program is high but relatively higher among the poorest households, thus increasing distributional power. However, this effect is mitigated by the fact that, conditional on program take-up, the poorest households take up a relatively lower proportion of potential transfers.

CONTENTS

TABLES

FIGURES

ACKNOWLEDGMENTS

David P. Coady
International Food Policy Research Institute

1. INTRODUCTION

In response to the emphasis on cutting budget deficits inherent in the structural adjustment programs of the mid 1980s onwards, there has been much debate, both in academic and policy circles, regarding the desirability of targeting transfer programs. On the one hand, well-targeted transfers can achieve a greater reduction in poverty for a fixed budget. On the other, targeting involves costs related to the collection, processing and analysis of survey data, and also involves excluding poor households solely on the basis of geographic location.[1] Motivated by this debate, in August 1997, the Mexican government introduced the Programa Nacional de Educacion, Salud, y Alimentacion (PROGRESA) with dual objectives, namely (1) the alleviation of current poverty through targeted cash transfers, and (2) the generation of a sustained decrease in poverty by conditioning these transfers on the accumulation of human capital (i.e., education and health status). In a sense, targeting is meant to increase the effectiveness of the program at ensuring that a large fraction of the budget gets to the poorest households (i.e., to increase the *distributional power* of the program), while conditioning is meant to introduce incentives for beneficial household responses. Since the introduction of PROGRESA, many others countries in Latin America have introduced similar programs.[2]

---

[1] The exclusion of poor households based on geographic location is seen as undesirable from the perspective of horizontal equity, which requires equal treatment of equals based on "relevant" characteristics. It is therefore desirable that other components of a more comprehensive poverty alleviation strategy address this issue. Note that targeting may also involve non-economic costs (see Adato 2000).

[2] Similar programs have already been introduced in Honduras and Nicaragua, while Argentina, Brazil, and Colombia are currently at the design stage.

In this paper we are interested in evaluating the program from the perspective of the first of the stated objectives, i.e., we wish to evaluate the distributional power of the program as reflected in its ability to get a relatively large proportion of the budget to the poorest households. As pointed out above, the targeted dimension of the program is obviously motivated by such an objective. The program is targeted in two respects. First, it is targeted to the poorest (or most marginal) rural localities, i.e., it is geographically targeted. Second, it is targeted at "poor" households within these localities. The conditioning of transfers in order to achieve the human capital objectives may also have important implications for its (short-term) distributional power. First, the linking of transfer levels to the demographic structure of households (i.e., transfers to children of school-going age) will, in general, affect its distributional power. Second, the conditioning of transfers to the accumulation of human capital involves households incurring private costs (e.g., those associated with schooling and health visits), which will generally affect the pattern of program take-up and thus its distributional power.

The objectives of this paper are threefold. We wish to (1) determine how the existing structure of the transfers compares to a range of alternatives, (2) understand how the different components of the transfer system contribute to or detract from the distributional power of the program, and (3) understand any trade-offs that exist between the poverty alleviation and human capital accumulation objectives of the program.

The format of the paper is as follows. In the next section, we set out the methodology used in the paper. In Section 3, we describe the program. Section 4 addresses the issue of the relative gains from geographic and household targeting. Section

5 presents an analysis of the distributional power of the program, conditional on locality targeting. Section 6 summarizes.

2. METHODOLOGY

In this section, we motivate and discuss the methodology employed to evaluate the distributional power of the program. For this purpose, it is useful to set out a very simple model of an economy with two groups, namely, households and the government.[3] The objective of public policy is taken to be to increase social welfare, which, in turn, depends on household welfare. For our purposes, the objective of the "social planner" may then be specified as choosing the size of the transfer *to* or *from* each household so as to maximize social welfare subject to the government budget constraint that the sum of transfers *to* household is less than or equal to the sum of transfers *from* households plus some exogenous budget, *F* (e.g., a predetermined level of foreign aid). Specifically, social welfare is specified as a function of household welfare, $V(\mathbf{p},m)$, where **p** is the vector of commodity and factor prices faced by the household and *m* is lump-sum transfers to or from the government. The Lagrangean function for the planner's problem can thus be written as choosing a set of values m^h for each household *h* so as to:

$$\max \Psi = W(...,V^h(\mathrm{p},m^h),...) + \lambda (F - \sum_h m^h),$$

---

[3] This section draws directly on Coady and Skoufias (2001) and Skoufias and Coady (2001). See Drèze and Stern (1987) and Coady and Drèze (2001) for a more rigorous discussion of the model.

where $W(.)$ is the social welfare function and λ is the Lagrange multiplier associated with the budget constraint.[4] This specification is essentially the specification for the determination of the optimal levels of cash transfers that maximize social welfare. As is well known (Atkinson and Stiglitz 1980; Stiglitz 1988), the solution to this optimization problem is determined from the first-order necessary conditions:

$$d\Psi = \frac{\partial W}{\partial V^h}\frac{\partial V^h}{\partial m^h}dm^h - \lambda dm^h = \beta^h dm^h - \lambda dm^h = 0, \; \forall h,$$

which implies $\beta^h = \lambda^*$, for all h, where β^h is the social valuation of an extra unit of income to household h, the so-called "welfare weight" of household h, and λ^* is the marginal social value of government revenue at the optimum. In other words, at the optimum, the pattern of transfers must be such that the social valuation of income at the margin is constant across all households. If all households are modeled as having the same utility function, then the optimum is characterized by an equal distribution of income.

By summing across all households and rearranging, the above first-order conditions can be re-written as

$$\lambda^* = \frac{\sum_h \beta^h dm^h}{\sum_h dm^h} \; .$$

---

[4] This formulation of the problem essentially assumes that cash transfers are nondistortionary lump-sum transfers and that no other distortions exist in the economy. Although restrictive, this simple formulation is adequate for our purposes. See Drèze and Stern (1987) for details and Coady and Harris (2000) for an application where tax distortions already exist and must be manipulated to finance the program.

One can interpret alternative income vectors $d\mathbf{m} = \{\ldots, dm^h, \ldots\}$ as representing alternative targeting schemes for a given budget. This statistic is simply the benefit derived from the transfers divided by the program budget, i.e., the social welfare impact per unit of funds transferred or the benefit-cost ratio of a transfer program.[5] At the optimum this is constant across all programs. Optimality essentially assumes that the poverty-alleviation budget is endogenously determined. In practice, however, there are economic, social, and political constraints both on the size of budgets and their distribution. In the absence of an optimal distribution of income, β^h will, in general, differ across households. Therefore, λ will differ across alternative transfer programs both because β^h differs across households and the structure of $d\mathbf{m}$ differs across alternative programs.

EVALUATION OF ALTERNATIVE TRANSFER PROGRAMS

Based on the above, for each program j, we have an associated distributional characteristic defined as

$$\lambda_j \equiv \frac{\sum_h \beta^h dm_j^h}{\sum_h dm_j^h} \quad , \tag{1}$$

---

[5] It is analogous to the so-called *distributional characteristic* found in the literature on indirect taxation. For a more detailed discussion of this literature see, for examples, Feldstein (1972), Sandmo (1976), Auerbach (1985), Newbery and Stern (1987), and Myles (1995).

which can be interpreted as the marginal social value of a unit of revenue transferred to households through the program in question. It can be used to compare the relative welfare impact of alternative programs with a common budget or of reallocations of a budget between different programs. Notice that it is independent of the size of the budget, i.e., scaling up or down benefits and the budget of a program proportionally will not change its distributional characteristic.

The distributional characteristic of a program can also be usefully re-written as

$$\lambda_j \equiv \sum_h \beta^h \theta_j^h \,, \tag{2}$$

where θ_j^h is the share of household h in the total budget of program j. Programs in which those receiving relatively high shares of the budget have relatively high welfare weights (i.e., are relatively needy) will obviously have relatively high welfare impacts. In this sense, it captures the distributional power of a program, i.e., how effective a given program is at getting the transfer budget to the most needy households.

ALLOWING FOR DIFFERENT PROGRAM BUDGETS

As pointed out above, the distributional characteristic of a program is scale neutral in that it is independent of the size of the budget (or level of transfers) and depends only on the distribution of transfers across households. Therefore, it serves as a sufficient statistic only for evaluations comparing alternative distributions of a fixed

budget or for the reallocation of funds across budgets. When program budgets (*B*) differ,

then it is also useful to think of the welfare impact as

$$dW = \lambda B , \tag{3}$$

so that

$$dW^* = \lambda^* + B^* , \tag{4}$$

where an asterisk denotes a proportional change. In other words, the proportional

difference in the impact between two programs can be seen as the sum of the proportional

difference in the welfare impact per unit of budget expenditure (i.e., the distributional

characteristic) plus the proportionate difference in the total budget. For well-targeted

programs, one expects λ^* to be negative in the face of program expansion to include more

households (i.e., extensive expansion due, for example, to an increase in the poverty line

being used to select households). For intensive expansion, i.e., an increase in the transfer

levels to existing beneficiaries, $\lambda^* = 0$, so that $W^* = B^*$.[6] In other words, the proportional

difference in program budgets can be taken as the proportional difference in their welfare

impacts only for an intensive expansion of the program.

COMPARISON ACROSS PROGRAM COMPONENTS

   Where a program consists of a combination of separately identifiable transfer

components (e.g., education subsidies and food transfers for the present program), it is

---

[6] Strictly speaking this is only true for marginal increases in transfer levels since otherwise welfare weights are endogenous and decreasing in transfer levels.

useful to decompose the distributional characteristic by component. By viewing *dm* as the
sum of a number of components, *i*, it is easy to show that

$$\lambda = \sum_i \lambda_i \sigma_i \ , \tag{5}$$

where λ is the distributional characteristic for the full program, λ_i is the distribution
characteristic for component *i* in isolation, and σ_i is the share of component *i* in the total
program budget. By comparing across λ_i, one can determine the relative distributional
power of the various components.

TARGETING EFFICIENCY VERSUS REDISTRIBUTIVE EFFICIENCY

It is also often useful to decompose the λ for each program into the sum of its
targeting efficiency (λ_T) and its *redistributive efficiency* (λ_R) by adding and subtracting
the *average* level of the transfer across all *beneficiaries* (i.e., across households with
$dm^h > 0$) to get (Coady and Skoufias 2001)

$$\lambda = \frac{\sum_h \beta^h dm^*}{\sum_h dm^h} + \frac{\sum_h \beta^h (dm^h - dm^*)}{\sum_h dm^h} = \lambda_T + \lambda_R \ , \tag{6}$$

where, for beneficiaries, $dm^* > 0$ is the average level of the transfer across beneficiaries
while, for nonbeneficiaries, $dm^* = 0$. One can interpret λ_T as the welfare impact of a
program that transfers the poverty alleviation budget to the same beneficiary households
but in equal amounts, and λ_R as the adjustment that needs to be made to allow for the

differentiation of the transfers across households in a more progressive ($\lambda_R > 0$) or

regressive ($\lambda_R < 0$) manner.[7]

WELFARE GAINS FROM HOUSEHOLD TARGETING

One of the contentious issues related to the program has been the decision to

target households within localities. Evaluating the gains from targeting within the present

framework is straightforward; one just treats the targeted (T) and nontargeted (N)

programs as separate programs with different patterns of transfers across households. But

one may also be interested in evaluating how the gains from targeting are distributed

across different localities. For any given program, it is then useful to write

$$\lambda = \frac{\sum_h (\beta^h dm_1^h + \beta^h dm_2^h + ... + \beta^h dm_L^h)}{\sum_h dm^h} \ , \qquad (7)$$

where $d\mathbf{m}_i = d\mathbf{m}$ for households living in locality i and equals zero otherwise, with L

localities. Multiplying and dividing through by $\sum_h dm_i^h$, this can be rewritten as

$$\lambda = \sum_i \lambda_i \theta_i \ , \qquad (8)$$

where subscript i refers to localities. Therefore, the distributional characteristic of a

program can be seen as a weighted average of the distributional characteristics of

---

[7] This decomposition essentially views a uniform transfer as distributionally "neutral." A similar decomposition can be derived by replacing the uniform transfer with a proportional transfer. The former has been termed *rightist* by Kolm (1976), and the latter is often termed leftist, since it is a more demanding definition of neutrality. Intermediate versions, termed μ-variance are discussed in Eichhorn (1988).

subprograms in each locality, with the share of the total budget going to each locality (θ_i) as weights. One can then write the welfare gains due to targeting as

$$\lambda^T - \lambda^N = \sum_i \theta_i (\lambda_i^T - \lambda_i^N) = \sum_i \theta_i g_i \quad , \tag{9}$$

where superscripts *T* and *N* refer to the targeted and nontargeted programs and *g* refers to the gain from targeting.[8] So the total welfare gains from targeting can be seen as a weighted average of the welfare gains from targeting across localities with locality budget shares as weights.

INTRODUCING PROGRAM COSTS

So far, we have ignored the fact that different programs may involve different levels of administrative costs. For example, targeting will generally require the collection, processing, and analysis of extra survey information. Higher administrative costs mean that a smaller proportion of the poverty alleviation budget is available to be transferred to households. If administrative costs increase with the degree of targeting, then there will be a trade-off between the increased distributional power due to finer targeting (or the marginal benefit from finer targeting, *MB*) and the decreased budget available for transfers (or the marginal cost of finer targeting, *MC*). The magnitude of this trade-off, or where *MB = MC*, is an empirical issue.

---

[8] In deriving this, we make use of the assumption that locality budget shares are kept constant so that the nontargeted transfers can be viewed as being scaled down appropriately.

The adjustment needed to be made to equation (1) for the presence of program

costs is straightforward. The equation for the distributional characteristic becomes

$$\lambda_j \equiv \frac{\sum_h \beta^h dm_j^h}{\sum_h dm_j^h + C} \quad , \tag{10}$$

where C is total program costs.[9] Finer targeting essentially involves (1) increasing the

numerator through increasing the correlation between welfare weights and transfer levels,

and (2) increasing the denominator by increasing C. We make use of this definition of the

distributional characteristic below when we compare targeted versus nontargeted

programs or when comparing programs with different degrees of targeting (e.g.,

geographic versus geographic plus household).

WELFARE WEIGHTS

Underlying our objective of poverty alleviation must be the view that extra

income to low-income (or poor) households is more socially valuable than extra income

to high-income (or nonpoor) households. Making this view explicit essentially requires

the specification of a set of "welfare weights" and we expect this weight to decrease with

the (initial) consumption (or welfare) level of the household. The welfare weight for each

household (β^h) can be derived from a constant elasticity social welfare function as

follows (Atkinson 1970):

---

[9] If C is interpreted as (incremental) private costs that can be attributed differently to specific groups of households, then one could instead include the term (dm^h - c^h) in the numerator.

$$\beta^{h} = (y^{k} / y^{h})^{e} \ ,$$

where y refers to consumption (or "permanent income"), h superscript denotes the household in question, and k superscript denotes a reference household, which always has a weight of unity (e.g., the household just on the poverty line, in which case $y^{k} = z$, where z is the poverty line).[10] The term e captures one's "aversion to inequality" of income or consumption and determines how the welfare weights vary (i.e., decrease) with household income. For example, a value of $e = 0$ implies no aversion to inequality and all welfare weights take the value unity, i.e., an extra unit of income to households is viewed as being equally socially valuable regardless of initial consumption levels. A value of $e = 1$ implies that if household h 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 $e = 2$ similarly implies a welfare weight of 0.25 (4.0) for h. As e 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 in which one cares only about how much of the program benefits are received by the poorest of the poor. The welfare weights used in our simulations presented below use initial consumption as their welfare reference and we also evaluate the sensitivity of our findings to different sets of welfare weights based on different degrees of aversion to inequality of initial consumption (i.e., different values of e). Consistent with the program objectives, we consider only values of $e > 0$.

---

[10] Which household we use as the reference household to normalize welfare weights is irrelevant to our analysis. See, for example, Ahmad and Stern (1984; 1991, 129) for discussion on the choice of welfare weights.

3. A DESCRIPTION OF *PROGRESA*

The targeting process adopted by PROGRESA is essentially a two-stage process.[11] Using the national census data, an index of marginality (IML) is constructed for each locality. Based on this index, the most marginal localities are chosen to participate in the program. Once participating localities are identified, PROGRESA then undertook a locality census (Encuesta de Características Socioeconómicas de los Hogares [ENCASEH]) that includes data on household demographics, income, and assets. Households are categorized as "poor" and "nonpoor" based on income with reference to a standard food basket. Households are then reclassified using discriminant analysis and household characteristics other than income, e.g., dependency ratio, characteristics of household head (i.e., age, sex, occupation, and schooling), and dwelling characteristics. This classification process appears to have changed over time: the initial classification (PRO) had just over 50 percent of households classified as poor, but this increased to nearly 80 percent with the "densification" process (PROD), which used a higher poverty line. The increase in the percentage of households classified as poor came about essentially due to a community participation process that suggested that the program selection mechanism had led to a substantial underestimate of the poverty rate.

Once households are deemed eligible for the program, they receive benefits according to the structure set out in Table 1. There are two types of benefits. First, there is the education subsidy, which starts in grade 3 in primary school continuing until grade

---

[11] See Skoufias, Davis, and Behrman (1999) for details.

9 at the end of junior secondary school. To receive the subsidy, children must attend

school at least 85 percent of the time. Subsidies increase with age and are also higher for

girls in secondary school.[12] Second, households receive a fixed transfer made conditional

on household members making trips to health clinics for preventative check-ups.

Transfers are given to mothers every two months and, in principle, eligibility is to be

reviewed every three years. By applying the grants structure set out in Table 1 to the data,

Table 1—Benefit structure of PROGRESA, 1997 (pesos/month)

	Boys	**Girls**
Education Scholarships		
Primary		
8 years old	60	60
9 years old	70	70
10 years old	90	90
11-12 years old	120	120
Materials (annual)	110	110
Secondary		
13-14 years old	175	185
15 years old	185	205
16-17 years old	195	225
Materials (annual)	140	140
Food Transfer	90 per family	
Benefit Cap	550 per family	

Note: The benefits structure is meant to mimic as closely as possible that of the actual program that is linked to grades and not age: we choose age since this is consistent with an unconditional transfer program and, in any case, the data on maximum grade achieved is not very reliable. To be consistent with the figures for monthly per adult equivalent consumption, these numbers are deflated by a factor of 2.2868 to bring them to 1994 prices. The structure of scholarships is applied by age group but, in practice, is applied by grade and conditional on enrollment and attendance. All families receive the food transfer but, in practice, this is made conditional on regular visits to a health clinic. The cap on the total benefits a household can receive is applied only to the sum of the education scholarships and food transfer.

---

[12] In reality, subsidies are linked to grade, not age. Typically the ages of children over the grades 3-9 range from 8-17 years old, although there is also evidence of slower progression over the early years, particularly for boys. Throughout our analysis, we link subsidies to age as indicated in the table, partly because the information on maximum grade achieved is not very clean.

we are essentially assuming 100 percent take-up of benefits, which is probably more likely to reflect the situation if the transfers were unconditional. We use this structure of transfers as our reference, but we also compare these hypothetical transfers to actual transfers.

By the end of 1999, the program was being implemented in nearly 50,000 rural localities in over 2,000 municipalities in 31 states. In all, approximately 2.6 million families, equivalent to 40 percent of all rural families and one-ninth of all families, were receiving benefits. The total budget of the program was just under 20 percent of the federal poverty alleviation budget or 0.2 percent of GDP.

The analysis below is split into two parts. First, using household survey data, we analyze the relative welfare gains from geographic and household-level targeting. Second, conditional on geographic targeting, we use program data to further analyze the welfare gains from household-level targeting as well as the implications of the demographic structure of transfers and of take-up for the distributional power of the program.

4. COMPARING GEOGRAPHIC AND HOUSEHOLD-LEVEL TARGETING

As indicated above, the program targets in two dimensions. It first targets the most marginal localities (i.e., geographic targeting) and then targets poor households within these localities (i.e., household targeting). While such targeting is obviously expected to increase the distributional power of transfers, one also needs to recognize that

targeting requires resources. In this section, we use nationally representative household survey data to simulate the *potential* welfare gains from locality and household targeting.

We base our simulations on the 1996 Mexican National Survey of Income and Expenditures (ENIGH), which contains 13,649 households in 33 states, 416 municipalities, and 708 localities.[13] The welfare index we use is consumption per adult equivalent (CPAE). For the purposes of simulation, we assume that the bottom 30 percent of households according to CPAE are "poor" and the structure of transfers is based on that presented in Table 1. We evaluate three programs:

1. *universal eligibility*, where all households receive benefits (i.e., no targeting),
2. *locality targeting*, where all households in the poorest localities receive transfers (i.e., first-stage targeting), and
3. *household targeting*, where only poor households in the poorest localities receive transfers (i.e., first- and second-stage targeting).

The program budget is determined by (3), which differs from (2) in that more localities can be included when nonpoor households are excluded. Note that (3) can be viewed as "perfect targeting" since all poor households are correctly identified and included.

In Figure 1, we compare the percentage increase in the distributional characteristic in moving from (1) to (2) and (3), i.e., from introducing finer levels of

---

[13] In our analysis we ignore the fact that the survey is only representative at the state and urban-rural level. This may have important implications for the standard errors around our estimates.

targeting. Since we are assuming perfect knowledge of household or mean locality welfare, welfare gains should be viewed as upper bounds. As one expects, the gains from targeting increase the greater one's concern for the poorest of the poor. For moderate aversion to inequality (e.g., $e = 2$), the gains from locality targeting are 67 percent compared to 85 percent for household targeting. Therefore, the gains from the first-stage locality targeting account for between 76-88 percent of the gains from household targeting, this increasing with inequality aversion. Therefore, the gains from second-stage household targeting appear relatively small, even without any targeting errors.

The above results ignore the costs of targeting. In Coady, Perez, and Vera-Llamas (2000), we calculate that first-stage locality targeting costs 0.004 pesos for every 100 pesos transferred to households, while second-stage household targeting costs an extra 2.7 pesos per 100 pesos transferred. The relatively low cost of first-stage targeting reflects the fact that it was based on an existing census and therefore simply involved processing and analyzing these data.[14] These fixed costs are also spread over a large number of households receiving relatively large transfers. Household-level targeting requires the collection, processing, and analysis of new household survey data (i.e., the ENCASEH data used below). Even then, the associated costs are not exorbitant, reflecting a combination of management efficiency and relatively high transfers. When these costs are included, they result in a scaling down of the distributional characteristics by factors 0.996 (for locality targeting) and 0.970 (for household targeting). Including

---

[14] In reality, of course, locality marginality was based on locality characteristics other than consumption, since the latter is not available in the census.

costs, then, the proportion of the total gains from targeting accounted for by first-stage locality targeting increases to between 77-90 percent.

We also carried out a similar analysis replacing locality targeting with municipality targeting and got very similar results. As expected, the percentage of the targeting gains accounted for by first-stage municipality targeting was lower, but still relatively high, at between 69-82 percent when targeting costs were ignored and between 70-83 percent when they were incorporated into the analysis.

Finally, it is important to point out that one of the shortcomings of geographical targeting is the fact that some poor households are left out of the program based solely on the fact that they live in the "wrong" locality. For example, under locality targeting, nearly 39 percent of poor households are left out of the program, essentially being replaced by nonpoor households. This outcome is clearly undesirable from the perspective of horizontal equity so that it is important that other components of a more comprehensive poverty alleviation strategy address this issue.

5. THE RELATIVE DISTRIBUTIONAL POWER OF *PROGRESA*

In this section, we evaluate the program conditional on geographic targeting using the 1997 baseline ENCASEH household census survey of 24,077 households in 506 localities. For the purposes of evaluating the various dimensions of the program (i.e., education and health impacts), these localities were randomly allocated to "control" (186 localities) and "treatment" (320 localities) groups. The analysis below is based on data

for the 14,856 households in the "treatment" sample. Using data on household

composition, we first estimate the benefits received by households based on the payment

schedule set out in Table 1. For all households identified as "poor" by PROGRESA,

payments are linked to the number, age, and gender of children. Therefore, one can view

the program as involving a combination of household targeting and a demographic

structure of subsidies. Table 2 presents average transfers by component separately for

control and treatment localities. In treatment localities, the transfers account for, on

average, nearly 29 percent of total household consumption.

Table 2—Distribution of benefits across component for treatment and control beneficiary households

	Treatment Sample	Control Sample
Primary scholarships	76.1	75.9
Secondary scholarships	130.0	132.6
School materials	15.5	15.5
Food transfer	90.0	90.0
Total transfer	295.5	298.6
Number of households	14,856	9,221
Proportion poor:		
Pre-densification (PRO)	0.528	0.508
Post-densification (PROD)	0.782	0.778

Note: Transfers are in 1997 prices based on Table 1. The total transfer reflects the benefit cap.

The estimated benefits received by households are essentially *hypothetical*

transfers, i.e., the transfers that would exist if there was 100 percent take-up by all

eligible households. This hypothetical program acts as our reference for evaluation

purposes and one would expect its benefits structure to closely resemble that of

PROGRESA if the transfers were unconditional. We compare the welfare impact of such a program with the following alternatives:

1. *Pre-Densification Transfers* (PR): We compare the present post-densification pattern of transfers (PRD) with that which existed prior to the increase of the poverty line;

2. *Targeted Uniform Transfers* (TU): Instead of poor households receiving transfers linked to demographic characteristics, one can consider a uniform transfer to these households;

3. *Non-Targeted Uniform Transfers* (NU): Same as (2) but now all (i.e., poor and nonpoor) households receive a uniform transfer;

4. *Non-Targeted Demographic Transfers* (ND): A program without household targeting, where all households in the selected localities receive the benefits based on demographic structure.

Comparisons across (1)-(4) help to identify the relative gains from household and demographic targeting. However, for the actual targeted demographic program (PRD), it is also useful to compare the existing structure of transfers with various alternatives. We therefore carry out the following comparisons:

5. *Transfer Components*: We decompose the welfare impact of each program component (i.e. primary scholarships, secondary scholarships, school materials,

and food transfer) in order to identify the contribution of each to the total welfare impact. This analysis will inform the issue of the welfare impact of a change in the structure of the transfers (e.g., reducing food transfers or primary scholarship levels to finance an increase in secondary scholarships in order to get a greater education effect).

6. *Actual Transfers*: Some households do not receive the theoretical transfers because they decide not to take up certain benefits or do not satisfy certain conditions. Households that do not undertake their scheduled visits to the health center do not receive the food transfer. Neither do households in which children do not meet the 85 percent school attendance criterion receive transfers for these children. Actual will also deviate from hypothetical in so far as ages are not synonymous with school grades.

Note that the reference program is the post-densification program. Where the total budgets of the programs differ from the actual post-densification budget, the benefits are effectively scaled up or down appropriately.

COMPARISON OF PRE- AND POST-DENSIFICATION *PROGRESA*

As indicated above, the classification of households as "poor" and "nonpoor" appears to have changed over time as a result of a "densification process." In order to facilitate comparison to the earlier evaluation of PROGRESA's targeting (Skoufias, Davis, and Behrman 1999), for both stages of the process, we first compare the leakage

(*L*) and undercoverage (*U*) rates, defined, respectively, as the percentage of poor

households wrongly left out of the program (i.e., errors of exclusion) and the percentage

of beneficiary households that are wrongly included (i.e., errors of inclusion). This

requires that we establish an "ideal" welfare indicator and, in line with convention in

economics (Ravallion 1993; Deaton and Zaidi 1999), we choose household per adult

equivalent consumption.[15] The distribution of this variable (henceforth referred to simply

as consumption) is presented in Figure 2 together with the pre- and post-densification

poverty lines.

Using consumption as the reference welfare measure, we classify households as

poor and nonpoor and compare this with PROGRESA's classification, which was based

on income. We find that pre-densification *U* and *L* were both 27 percent compared to

post-densification, where both were around 16 percent. This difference, in part, just

reflects the fact that a higher percentage of households were included post-densification

thus leaving less room for *U* and *L* errors. For example, even using random selection of

households, *U* and *L* would still decrease from 47 percent to 22 percent. But notice that

using performance relative to random allocation, the pre-densification program performs

---

[15] We use an updated version of the measure used in Skoufias, Davis, and Behrman (1999), which was provided by Emmanuel Skoufias to whom I are very grateful. The welfare indicator is, in fact, predicted consumption based on coefficients from a consumption model estimated using quantile regression analysis of data from the 1996 nationally representative household survey (ENIGH), which returned a pseudo-R^2 of 0.35. However, here we ignore the error inherent in this prediction, so that it is important that care be exercised in interpreting the magnitude of differences across comparison programs. But since we are often aggregating over large numbers of households, this can substantially reduce the associated standard errors (see Elbers, Lanjouw, and Lanjouw 2001 for details). Based on this consumption measure and the transfer schedule in Table 1, the program budget was 22.5 percent of the poverty gap and transfer levels were, on average, equal to 21.1 percent of total household consumption.

better than the post-densification program: whereas the former reduces undercoverage by 50 percent more than random allocation, the latter achieves only a 20 percent reduction.

However, what matters more for the welfare impact of each program is where these errors occur in the distribution of consumption (e.g., are they concentrated around the poverty line or spread out). For example, the welfare losses from mis-targeting will be relatively high if a high proportion of the poorest households are wrongly excluded and/or if a high proportion of the richest households are wrongly included. We can examine this by looking at "predicted error probability" curve used in Skoufias and Coady (2001). We construct a binary variable where each household that is misclassified as poor or nonpoor by PROGRESA's methodology is assigned a value of unity with all other (correctly classified) households being assigned a value of zero. We then simply plot the averages for the various consumption 5-percentiles (Figure 3). Notice that although with pre-densification the curve is bell-shaped, with the percentage error decreasing the further one gets from the poverty line; with post-densification, this is not the case, with well over 50 percent of nonpoor households being wrongly included in the program. This is suggestive of substantial welfare losses due to mis-targeting.[16]

We finish this section by plotting the average percentage of households classified as poor in each locality ordered by marginality index and grouped into 5-percentiles

---

[16] This comparison also highlights the fact that comparing leakage rates across programs of different sizes (as above) can be very misleading; they are relatively low post-densification because, although a relatively high proportion of the nonpoor are wrongly classified as poor, these are a much lower percentage of the total households receiving benefits, given that nearly 80 percent receive benefits. Similarly with undercoverage, the greater the proportion of total households included in the program, the lower the potential for undercoverage.

(Figure 4). We do this for both the post- and pre-densification programs, but also for the

post-densification program under the assumption of perfect targeting. Comparing the pre-

and post-densification relationship between locality coverage and marginality, we see

that the biggest increases in coverage occurred in the least marginal localities, i.e., those

localities in the bottom quartile according to the marginality index. Since mean income is

negatively correlated with marginality and given the observed high levels of leakage in

the highest income groups, one expects that much of this increase is due to mis-targeting.

Comparing PROGRESA's post-densification coverage levels with those based on

consumption (i.e., perfect targeting), it is clear that the inclusion of nonpoor households

in the least marginal localities came mainly at the expense of poor households in

localities with marginality indices just above the least marginal localities, i.e., in the next

to bottom quartile according to the marginality index.

PROGRESA VERSUS ALTERNATIVES

In this section, we compare the welfare impact of PROGRESA with that of the

alternatives (1)-(4) identified earlier. The welfare impacts of all these programs are

presented in Table 3 and their performance relative to PROGRESA post-densification

plotted in Figure 5. Focusing on Figure 5, the first thing to notice is that the distributional

efficiency is higher pre-densification compared to post-densification. At a relatively

moderate aversion to inequality ($e = 2$), the welfare impact per unit expenditure is over 12

percent higher pre-densification compared to post-densification. This, of course, is not

surprising since, for the most part, the densification process is about incorporating

Table 3—Distributional characteristics of alternative transfer programs

	PROGRESA (Post-densification)	PROGRESA (Pre-densification)	Consumption (Perfect)	No Target (ND)	PROGRESA Uniform (TU)	No Target Uniform (NU)
$e = 0.5$	1.42	1.48	1.44	1.38	1.30	1.25
$e = 1.0$	2.11	2.26	2.13	1.98	1.77	1.65
$e = 2.0$	5.05	5.66	5.04	4.56	3.74	3.32
$e = 3.0$	13.27	15.25	13.09	11.71	8.99	7.76
$e = 4.0$	37.23	43.41	36.33	32.41	23.76	20.16
$e = 5.0$	109.32	128.66	105.92	94.38	67.05	56.34
Average transfer	113	122	125	110	113	110
Number beneficiaries	11,623	7,837	11,623	14,856	11,623	14,856
Budget Scale Factor	-	0.73	1.11	1.25	1.00	1.00

Note: The budget scale factor is the factor is the ratio of the alternative program budgets to the post-densification budget; its inverse is the amount by which the post-densification transfers need to be scaled down or up in order to keep the total budget constant. Note that the distributional characteristic is independent of the size of the budget and is interpreted as the welfare impact of a unit of income being transferred to households through the various programs; in other words, the distributional power of the alternative programs.

households that were previously deemed to lie above the poverty line and, for well-targeted programs, the distributional characteristic will always decrease as the program expands to include extra households. But in moving to the post-densification program, the budget also increases by 37.5 percent, which, using (4) above, implies an overall welfare increase of 25.5 percent. The question then becomes whether the decrease in the distributional power of the program could have been lower. For this one needs

comparisons with other potential programs, a natural one being a program that targets

perfectly using household consumption levels.[17]

What is somewhat surprising is that the post-densification program compares very

favorably to the situation with perfect targeting. In fact, it dominates the latter for higher

levels of inequality aversion. The intuition behind this at first counterintuitive result lies

in the realization that while perfect targeting correctly identifies poor households, the

transfer *levels* are not optimal across households; this would have included higher

transfers to the poorer households. What is happening is that those wrongly incorporated

under the densification process are moderately poor and have fewer children than the

moderately poor households that they replace. Therefore, as well as having relatively low

welfare weights, they also receive relatively low transfers, thus increasing the relative

share of the total transfer budget received by poorer households, which have higher

welfare weights. With consumption targeting, the households correctly included are

moderately poor with more children and therefore receive relatively high transfers, thus

decreasing the proportion of the budget going to the severely poor. Thus, the

distributional efficiency of the program, which can be seen as a weighted average of the

welfare weights of beneficiaries with the weights being the share of the overall budget

they receive, decreases. In the terminology of Coady and Skoufias (2001), although the

---

[17] One can distinguish between "perfect" targeting and "optimal" targeting. Under the former, the poverty headcount is predetermined and perfect targeting involves identifying correctly those households that lie below this welfare level. In a sense, for a given structure of transfers, the transfer budget becomes endogenous. Under optimal targeting with a fixed budget, on the other hand, the cutoff point is determined endogenously as part of determining optimal transfers. Transfers are distributed so as to equalize the incomes of the poorest households until the budget is exhausted.

"targeting efficiency" (i.e., who you hit) of the densified program is lower than that of a

perfectly targeted program, its "redistributive efficiency" is higher. Using the

decomposition described in equation (6), we can see from Table 4 that for $e = 1$, although

the targeting efficiency is higher for consumption targeting ($\lambda_T = 1.90$ as against 1.77),

the redistributive efficiency is relatively low ($\lambda_R = 0.23$ as against 0.34).

Table 4—Decomposition of distributional characteristics into its targeting and redistributive efficiencies

	Post-Densification			Consumption			Pre-Densification		
	λ	λ_T	λ_R	λ	λ_T	λ_R	λ	λ_T	λ_R
e=0.5	1.43	1.30	0.13	1.44	1.36	0.08	1.48	1.38	0.10
e=1.0	2.11	1.77	0.34	2.13	1.90	0.23	2.26	1.98	0.28
e=2.0	5.05	3.74	1.31	5.04	4.07	0.97	5.66	4.45	1.21
e=5.0	109.32	67.05	42.27	105.92	71.89	34.03	128.66	85.50	43.16

Note: λ is the distributional characteristic, λ_T is the targeting efficiency, and λ_R the redistributive efficiency. The targeting efficiency pre-densification is higher than that for consumption because it uses a lower poverty line to select households.

The results in Table 3 and Figure 5 also indicate that the (average) gains from

targeting households within localities is in the range 3 - 14 percent, depending on one's

aversion to inequality—below we look at the distribution of this gain across localities

with different characteristics. There are also very sizeable gains both from differentiating

payments by demographic composition as opposed to uniform transfers ranging from

9 - 39 percent when the program targets poor households and 12 – 49 percent without

household targeting. Both of these gains increase with the level of aversion to inequality.

So, in conclusion, it is clear that in spite of the mis-targeting inherent in the densification

phase of the program, the distributional power of the program is still relatively high.

PROGRESA is an extremely effective program in terms of getting transfers into the

hands of the most needy. However, it is also clear that in terms of relative contribution to

the distributional power of the program, it is the demographic structure of transfers that is

important, rather than the existence of household targeting.

DISTRIBUTIONAL EFFICIENCY OF INDIVIDUAL PROGRAM COMPONENTS

The structure of PROGRESA's transfers reflects its underlying objectives of

improving the current welfare of poor households while simultaneously encouraging

households to invest in their human capital. One of the issues being discussed by

policymakers is whether or not the structure of benefits should be changed, in particular,

whether a restructuring of scholarships so as to give higher grants to secondary school

children and lower grants to primary school children is desirable. From an education

point of view, this would appear desirable (Schultz 2001), since the biggest enrollment

impact comes from this older age group, where enrollment is still relatively low

compared to primary enrollment levels.[18] Here we are concerned with the trade-off in

terms of current welfare inherent in such a restructuring of benefits.

To address the above issue, we can interpret each of the separate components of

the program as alternative transfer instruments. By calculating a distributional

characteristic for each component using equation (5), we can identify the welfare impact

---

[18] However, it is important to note that although enrollment rates are high in primary school, at around 95 percent, the program seems to have important impacts in terms of improving progression rates and reducing drop-out rates (Behrman and Todd 2001).

of reallocating a unit of the budget between components. A crucial feature of the program, however, is that total monthly payments to households are capped at 550 pesos per household. Therefore, transfers of the budget between program components may have very little effect on the net transfers to capped households compared to uncapped households. This would appear to be particularly important for the distributional impact of the program, given that from Figure 6, we can see that it is the poorest of the poor who are relatively constrained by the cap, with 70 percent of the poorest income group being capped. The difference between capped and uncapped transfers is also greatest for these households (Figure 7). We address this issue by analyzing the distributional impact of components with and without a transfer cap. We finish this section by analyzing the welfare impact of scaling down primary school transfers to finance a scaling up of secondary scholarships.

Table 5 and Figures 8a and 8b present the relative welfare impacts of the various program components ignoring capping. It is clear that the more concerned we are about

Table 5—Welfare impact of program components

| | Uncapped Transfers | | | | | Capped Transfers | | | |
	Total	Primary	Secondary	Utilities	Food	Total Cap	Primary Cap	Secondary Cap	Food Cap
$e = 0.5$	1.45	1.52	1.50	1.51	1.30	1.43	1.45	1.38	1.24
$e = 1$	2.17	2.36	2.32	2.34	1.77	2.12	2.16	1.96	1.61
$e = 2$	5.34	6.01	5.96	5.98	3.74	5.12	5.05	4.31	3.05
$e = 3$	14.29	16.45	16.44	16.40	8.99	13.51	12.71	10.44	6.61
$e = 4$	40.66	47.16	47.76	47.28	23.76	37.96	33.95	27.26	15.92
$e = 5$	120.64	140.29	143.77	141.38	67.05	111.62	95.27	75.53	41.61

those suffering from severe poverty, the more attractive are the educational components compared to the food (or health) component from a distributional perspective. This reflects the fact that the former payments are linked to the number of children in a household, whereas the latter is a uniform transfer across households, and the number of children in a household is positively correlated with household welfare. It is also clear that the capping of transfers reduces the redistributive power of transfers, consistent with the poorest households being more likely to be capped since they have more children. Also, in the absence of capping, primary and secondary transfers are equally redistributive. However, when transfers are capped, primary transfers have a greater distributional impact compared to secondary transfers. Also, the greater our concern for the poorest of the poor, the greater the relative distributional power of primary transfers: the welfare impact from increasing primary transfers compared to secondary transfers is nearly 17 percent at moderate levels of inequality aversion ($e = 2$), increasing to just over 26 percent for higher levels (e.g., $e = 5$).[19] This suggests, therefore, that in the presence of capping, restructuring transfers in favor of secondary school children (in an attempt to increase the overall enrollment impact) may involve a trade-off in terms of a lower impact on current welfare for the poorest households.

The above analysis is strictly only valid for very small (i.e., marginal) changes in the payments structure, since it assumes that those households that are capped do not

---

[19] In the presence of capping, we assume that capped households do not receive any extra funds allocated to either the primary or secondary budgets, while other households receive amounts that keep their relative shares of the budget constant.

become capped due the restructuring of the transfer system. For larger changes, this will presumably not hold, since households are capped to different degrees and a crude categorization of household into capped and uncapped may be misleading. We therefore conclude this section by evaluating the welfare impact of a specific restructuring of the transfer scheme: a 10 percent increase in secondary scholarships, with the budget held constant by an appropriate decrease in primary grants. We also replicate a modified version of this with the cap also being rescaled up by 10 percent. Our results indicate that although welfare decreases when we increase secondary grants at the expense of primary grants, these welfare losses are always less than 1 percent. This lower welfare loss indicates that the restructuring enables the poorest households to get a greater share of the budget than is suggested by the marginal analysis, consistent with some of the higher income households becoming capped under the restructured transfer system. When we also scale up the cap by 10 percent, welfare increases by a high of 1.9 percent when $e = 5$ (Table 6). These results confirm that the welfare losses from restructuring scholarships

Table 6—Distributional characteristic for restructured benefits

Inequality Aversion	Cap = 550	Cap = 605
$e = 0.5$	1.43	1.43
$e = 1.0$	2.12	2.13
$e = 2.0$	5.11	5.17
$e = 3.0$	13.46	13.68
$e = 4.0$	37.85	38.59
$e = 5.0$	111.32	113.75

Note: The benefits are restructured by scaling up secondary scholarships by 10 percent and scaling down primary scholarships by a factor of 0.83, the latter keeping the budget constant. This is motivated by a desire to get a larger enrolment impact. The first set of results keeps the cap at 550 pesos, while the second also scales this up by 10 percent.

with the objective of enhancing the enrollment effect of the program are relatively small, but also that these could, if required, be offset by an increase in the maximum transfer allowed per household.

THE GAINS FROM HOUSEHOLD TARGETING

An important policy issue concerns the magnitude of the welfare gains from targeting poor households within localities. Our earlier discussion concentrated on the overall, or average, welfare gains from targeting; from Table 7 we can see that, on average, the welfare impact per unit expenditure increases by 2.9 percent for $e = 0.5$ to 10.7 percent for $e = 2$ to 15.8 percent for $e = 5$. In general, one expects the gains from targeting to depend on

- *The proportion of households that are to be included in the program*: The greater the proportion of households to be included, the lower the gains from targeting. For example, the gains are obviously zero if all households are eventually classified as poor and included. This explains the fairly modest average gains from targeting indicated above, given the very high coverage rate of 78 percent.

- *The targeting efficiency of the program*: The better the targeting efficiency of the program, the greater the gains from targeting. In general, for well-targeted programs, there are always gains from targeting. However, if programs use inefficient targeting mechanisms, then the gains from targeting will be relatively

small or even negative (e.g., if the poorest households were incorrectly left out of the program).

Table 7—Average percentage gains from household targeting

Inequality Aversion	Targeting	No Targeting	Targeting Gains
$e = 0.5$	1.42	1.38	2.9%
$e = 2.0$	5.05	4.56	10.7%
$e = 5.0$	109.32	94.38	15.8%

The above factors imply that although the average gains from targeting are modest, these gains may differ substantially across localities both because the percentage of household included in the program differs across localities and because there is evidence that the degree of mis-targeting also varies across localities. We therefore analyze the distribution of the gains from targeting for the current post-densification program, but also the potential distribution for the program if the current mis-targeting were eliminated.

The distributions of the gains from targeting, relative to the mean gain, both for the current post-densification program and for this program corrected for mistargeting, are presented in Figures 9 and 10 for $e = 2$. Notice first that with the actual program, the introduction of targeting *decreases* the distributional efficiency of the program in some localities, i.e., the gain in moving to targeting is negative (for 17 localities)—45 localities show zero gains. This is in part due to the fact that the actual targeting mechanism makes errors of inclusion and exclusion and dropping targeting may include some poor households previously excluded. As expected, these negatives do not exist for the

perfectly targeted program. The maximum gains over both programs are 7.7 and 9.2, respectively (relative to a common mean), suggesting that the gains from targeting are relatively substantial in some localities and that these gains can be increased through improved targeting.

Understanding the pattern of gains across localities can further help to identify when gains are likely to be substantial. We therefore examine the relationship between the magnitude of targeting gains and the marginality index. Figure 11 plots the relationship between absolute targeting gains and the marginality index. For the actual post-densification program, the gains first increase as one moves down the marginality index from the most marginal localities, but they suddenly begin to decrease again for localities in the top marginality quartile. This reflects the high degree of mis-targeting at this part of the marginality distribution. However, if this mis-targeting were to be reduced, one would find a consistently negative relationship between targeting gains and marginality. This helps to highlight the fact that efficient targeting becomes especially important as the program expands to less marginal rural and semi-urban localities if one is to capture the potential gains from targeting poor households.

So far we have ignored the cost of targeting households within localities. In Coady, Perez, and Vera-Llamas (2000), we calculated that (incremental) household targeting costs were 2.7 percent of the transfer budget: these costs include the cost of collecting the ENCASEH survey as well as processing the data and applying discriminant analysis to select households. For every $100 allocated to the program, $2.7 is absorbed in targeting costs, with $97.3 available to be transferred to households. These costs

therefore need to be offset against the improved distributional power of the program due to targeting. Applying equation (10), this is done by scaling up the denominator of the distributional characteristic (i.e., the program budget) by a factor of 1.027 or, equivalently, scaling down the distributional characteristic by a factor of 0.974.

The distributional characteristic of the targeted program decreases from 5.05 when program costs are excluded to 4.64 when these costs are included. That for the untargeted program decreases from 4.56 to 4.34. Therefore, the average percentage gains from targeting decrease from around 11 percent to around 7 percent. But, as with our earlier results, this modest gain hides variation across localities. Figure 12 compares the distribution of gains for the actual program across localities according to their marginality index, this time including program costs. The top line presents the distribution of gains when program costs are ignored; the bottom, the gains when program costs are included. The latter indicates that the gains from targeting are near zero both for the most marginal localities (reflecting very high coverage rates) and the least marginal localities (reflecting high mis-targeting). Figure 13 presents the corresponding pattern for a perfectly targeted program and helps to highlight that, even when targeting costs are included, the gains from targeting can be substantial on the margin but that this is dependent on cleaning up the mis-targeting that occurred during the densification process. But the gains from targeting are still negligible among those localities in the bottom 15 percent of the distribution of the marginality index.

In terms of the present program, targeting costs are already sunk, so cannot be recovered, and are therefore irrelevant to the issue of whether to continue targeting or not.

But if the recertification process requires incurring similar targeting costs to determine which households remain eligible, then, from a purely economic perspective, it appears that in these localities the returns to targeting are negligible. On the other hand, one could argue that because the net effect on the welfare impact of the program is negligible among these highly marginal localities and substantial in others, combined with the fact that the former tend to be relatively small in terms of number of households, it is not worthwhile deviating from the general principle of targeting. In this case, it is then probably desirable to apply a simpler targeting rule based on, for example, occupation (i.e., teacher, doctor, government official, etc.) or landholdings, at least in the most marginal localities. However, the analysis of Adato (2000) suggests that there may be important social costs associated with targeting and, if so, the nature and magnitude of these need to be considered. For example, if the related "social conflicts" are more important in localities where only a relatively small proportion of households are excluded, then it may be that the decision to target in the most marginal localities needs to be reconsidered. Whether or not to target in such localities is then more a socio-political decision rather than an economic one. Since one of the common complaints from households has been that the lack of transparency in the selection of beneficiaries is unfair, it may be that the simpler rules are more acceptable and sufficiently precise in the most marginal localities. But there is no escaping the fact that the gains from targeting are relatively substantial on the margin and will presumably increase as the program expands into even less marginal localities.

THE IMPACT OF CONDITIONING

The above analysis is based on hypothetical transfers that involve an implicit assumption of 100 percent take-up, an assumption consistent with the program being unconditional. However, the conditioning of transfers (e.g., on households sending their children to school or on making regular scheduled visits to health clinics) means that households incur private costs, and this may result in some households not taking up the program or not taking up all of the benefits potentially available. We address two dimensions that can lead to the levels and distribution of transfers differing between the actual and hypothetical programs.

- *Program Take-Up*: Some eligible households may not take up the program at all. For example, those without children may decide that the relatively low transfer levels are not worthwhile, given the costs involved in take-up.

- *Partial Take-Up*: Households may only take up some of the transfers available. For example, given the relatively lower level of the food transfer, some households may think it is not worth taking up this component. Or some children eligible for transfers may nevertheless not attain sufficient attendance records or may not even enroll in school.

We consider the effect of each of these on the distributional power of the program.

In order to evaluate the impact of program take-up rates for the program, we need to adjust the sample of households to allow for operational problems experienced during

the course of the program. In a number of localities, households that were supposed to be

incorporated during the densification phase of the program were not. Therefore, we need

to compare the distributional power of program with and without full take-up using a

sample of households excluding those wrongly left out. This reduces the sample from

11,761 to 9,598 poor treatment households.[20] To calculate the take-up rate, we use a

snapshot of the program over the months of September-December 1999. Any household

for whom a transfer was sent out and collected for this period is deemed to have taken up

the program. Under this definition of take-up, we find an 87.6 percent take-up rate. This

varies slightly across households incorporated before and after the densification phase,

with the take-up rate for the former being 87.9 percent and 86.2 percent for the latter.

Table 8 describes the distributional power of the program with and without the

operation errors at the incorporation stage, as well as the impact of take-up. The

distributional power of the program is substantially higher with the errors accounted for

than without, and the difference between the two increases, the greater our concern for

Table 8—Impact of take-up on distributional power of program

	Operational Errors		Definition of Take-Up		
	Without	With	Sent Out	Collected	Actual
$e = 0.5$	1.42	1.465	1.471	1.474	1.475
$e = 1.0$	2.11	2.217	2.234	2.243	2.244
$e = 2.0$	5.05	5.495	5.571	5.605	5.588
$e = 3.0$	13.27	14.769	15.038	15.155	15.035
$e = 4.0$	37.23	42.06	42.968	43.345	42.782
$e = 5.0$	109.32	124.82	127.80	129.00	126.76

[20] See the Appendix for more detailed discussion.

the poorest of the poor. This is consistent with the fact that those not incorporated come from the densification stage, which incorporated the moderately poor. Leaving out some of the moderately poor results in a higher proportion of the actual transfer budget being concentrated in the hands of the poorest households. The final three columns present the distributional power of the program once we allow for some households not taking up the program (Figure 14). The appropriate comparison is with column two, i.e., assuming full take-up in the sample of households that were actually incorporated into the program (i.e., with operation errors accounted for). In all cases, adjusting for take-up increases the distributional power of the program. However, the fact that the magnitude of the differences between the distributional power of the program are low, e.g., never being more than 3.5 percent, reflects the very high take-up rates. Using transfers "sent-out" as the basis for determining take-up, the increase in the distributional power of the program is positively related to our concern for the poorest households, indicating that take-up is highest among this group. Adjusting further for collection of transfers, we see an even higher increase in the distributional power of the program, especially for high aversions to inequality, indicating further that collection is higher among the poorest households. But when we adjust even further by replacing hypothetical transfers with actual transfers, we find that, for high degrees of aversion to inequality, the increase in distributional power is lower. This suggests that although a relatively high proportion of the poorest households participate in the program, conditional on participation, they take up a lower proportion of their potential benefits (e.g., because not all eligible children attend school or mothers do not make all of their scheduled health visits). Since this also has important

implications for human capital impacts, this aspect of the program warrants further attention.

6. CONCLUSIONS

In this paper we have been concerned with evaluating the distributional power of PROGRESA, i.e., its ability to get transfers to the most needy households in the program localities, relative to other potential transfer schemes. The program delivers transfers based on household demographic composition (i.e., based on the number of children, their ages, and gender) targeted both at the most marginal localities and at poor households within these localities. Our results suggest the following.

- The gains from geographic targeting are substantially larger than those from household targeting. In fact, the gains from the demographic structure of transfers are also substantially larger than the gains from household targeting.
- In spite of substantial leakage during the densification phase of the program, the distributional power of the program is still very high relative to alternatives. This reflects its effectiveness at identifying poor households, but particularly its effectiveness at getting a relatively high proportion of total transfers to the poorest of the poor. The latter, in turn, operates through the demographic structure of education transfers.

- Restructuring education transfers towards higher grants for secondary schooling in order to try to enhance the educational impact of the program has little effect on the distributional power of the program. Any adverse effect it has can be reversed through simultaneously adjusting the cap on transfers, which is relatively more binding for the poorest of the poor.

- Although the average gains from household targeting are modest, these vary inversely with locality marginality. But to reap the gains from targeting as the program expands to include less marginal rural and urban localities, it is important that the targeting errors that occurred during the densification process be avoided.

- The impact of program take-up is to increase the distributional efficiency of the program, reflecting the relative higher take-up rates among the poorest households. In other words, relatively more moderately poor households select themselves out of the program. But these gains are small because of the very high take-up of the program. However, conditional on take-up, the poorest households take up a relatively lower percentage of the full transfers, e.g., due to lower enrollment rates. This aspect of the program is particularly important, given that it affects both the distributional power of the transfers and the human capital impacts for the poorest households. It therefore warrants further analysis.

APPENDIX: CALCULATING TAKE-UP OF THE PROGRAM

To identify program take-up, we make use of a dataset that contains information on all the payments sent out to beneficiary households, which we merge with the baseline household dataset (i.e., ENCASEH 1997). We consider two different definitions of take-up:

- households that have received some payments for any bimester up to and including November-December 1999;
- households that received a payment for any one of the final two bimesters of 1999, i.e., for the months September-December.

Given that for the first few months of the program households received payments that were conditional only on enrollment at schools and registration at health clinics, regardless of attendance, the first definition can be interpreted as identifying households that decided to take up the program at its inception. But some households will have dropped out of the program over time. Thus, the second definition treats the final two bimesters of 1999 as providing a snapshot of take-up some two years after the program was introduced in the treatment localities. Note that the first definition is subsumed within the second so that the latter is stricter and will therefore result in lower take-up rates. In our empirical analysis, for the most part, we focus on the second definition.

The evaluation baseline dataset contains 24,407 households, which includes 14,994 households in "treatment" localities. We find that around 60 percent of treatment households had some money sent out to them since the start of the program, implying that these households were deemed to have met the conditions required for receiving transfers.[21] Focusing on the 14,994 treatment households, we find that 78 percent (i.e., 11,761 households) of these were classified as poor post-densification, 53 percent being incorporated pre-densification and the remaining 25 percent being incorporated during the densification process. Out of the 11,761 beneficiary households, 77 percent had at least one positive payment sent out to them since the start of the program. These constitute 95 percent of the pre-densification poor but only 40 percent of the poor households incorporated during the densification process. So, 2,681 poor households (i.e., 23 percent of poor households) never had a transfer sent out to them. If we add in the 170 households that did not have a transfer sent out to them for at least one of bimesters five or six in 1999, this increases to 2,851 households. To this we add 43 households that did not receive a transfer in bimester six and for which the data indicating whether or not they picked up their payment for bimester five are missing. These adjustments bring to 2,894 (i.e., nearly 25 percent of poor households) the number of households classified as not having transfers sent out (576, i.e., nearly 20 percent, that were incorporated pre-densification).

---

[21] Conditional on registering in school and a health clinic, households can receive up to three bimester payments (i.e., 6 months) without further conditioning. Nonreceipt of any transfers thus suggests that a household decided up-front not to participate. Below we will discuss how this number needs to be adjusted for mistakes during the incorporation process. The corresponding figure for the control group was 78 percent.

So, in our data, it appears that whereas just over 7 percent of the poor households that were incorporated pre-densification did not receive transfers for the last two bimesters of 1999, the corresponding number increases dramatically to just over 60 percent for households incorporated during the densification process. However, this number includes a group of households that were identified as poor during the densification process but are from localities that were never, in fact, incorporated. For our purposes, we view these households as not being part of the program and eliminate them from the analysis completely. To identify these households, we take all the households that were initially meant to have been introduced into the program during the densification process, that reside in localities where no such households ever received a transfer. This is consistent with the incorporation error being locality specific. In all, we identify 2,163 such households (i.e., nearly 42 percent of the households that were meant to be incorporated during the densification process). When these households are dropped from the sample, we are left with 9,598 households in the treatment sample, of which 731 households never had a payment sent out since the start of the program *nor* did they receive a transfer in either bimester five or six. So nearly 8 percent of poor households that were incorporated can be deemed not to have taken up the program under our second measure of take-up.

As well as those households that by choice or default have not taken up the program, there are households for whom transfers were sent out but were never collected. This provides us with a third definition of take-up, i.e., those households that both received and collected transfers for the months September-December 1999.

In our empirical analysis we focus on the second and third definitions of take-up. The percentages of households not taking up the program under both definitions are 7.6 percent and 12.4 percent, respectively. The corresponding numbers for pre- and post-densified households are 7.3 percent and 12.1 percent, and 9.4 percent and 13.8 percent, respectively. So, even after adjusting for incorporation problems, take-up of the program appears to be substantially lower among those poor households incorporated during the densification process.

FIGURES

Figure 1—Relative gains from locality and household targeting
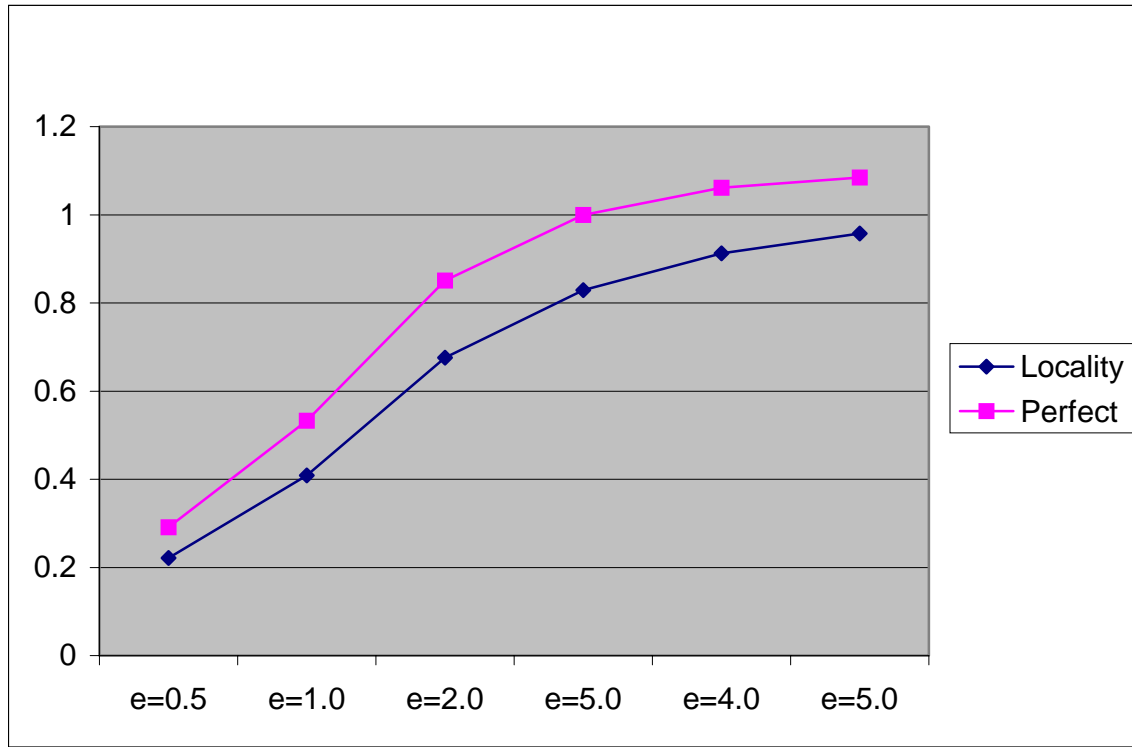
Figure 2—Kernel density for monthly per adult equivalent consumption (1994 prices)
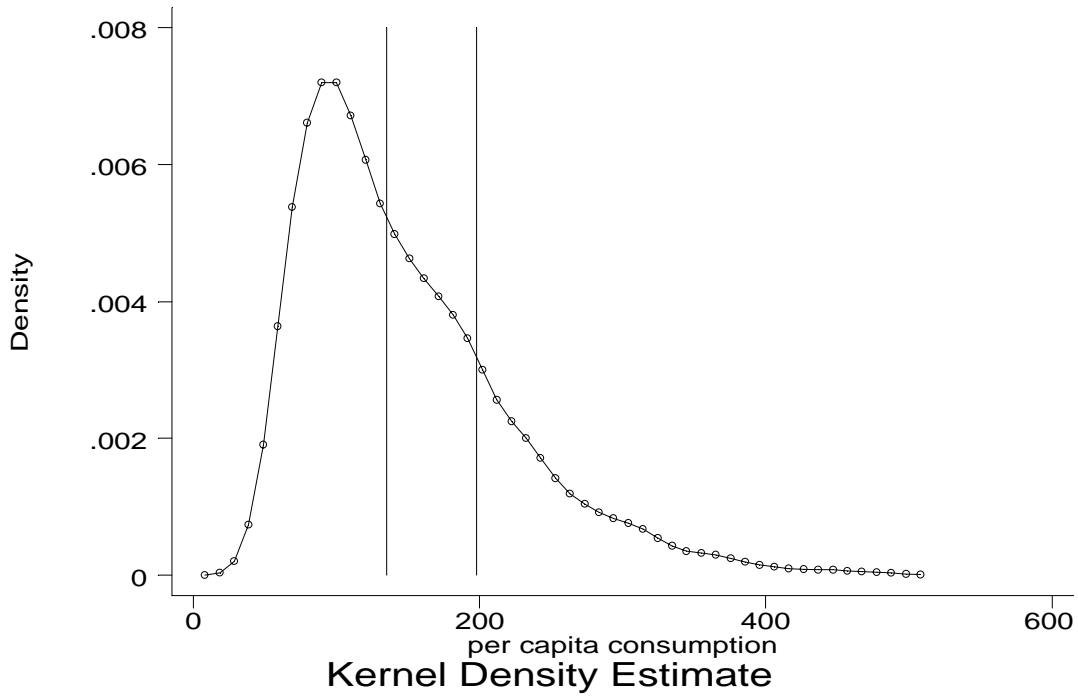
Figure 3—Pre- and post-densification targeting errors
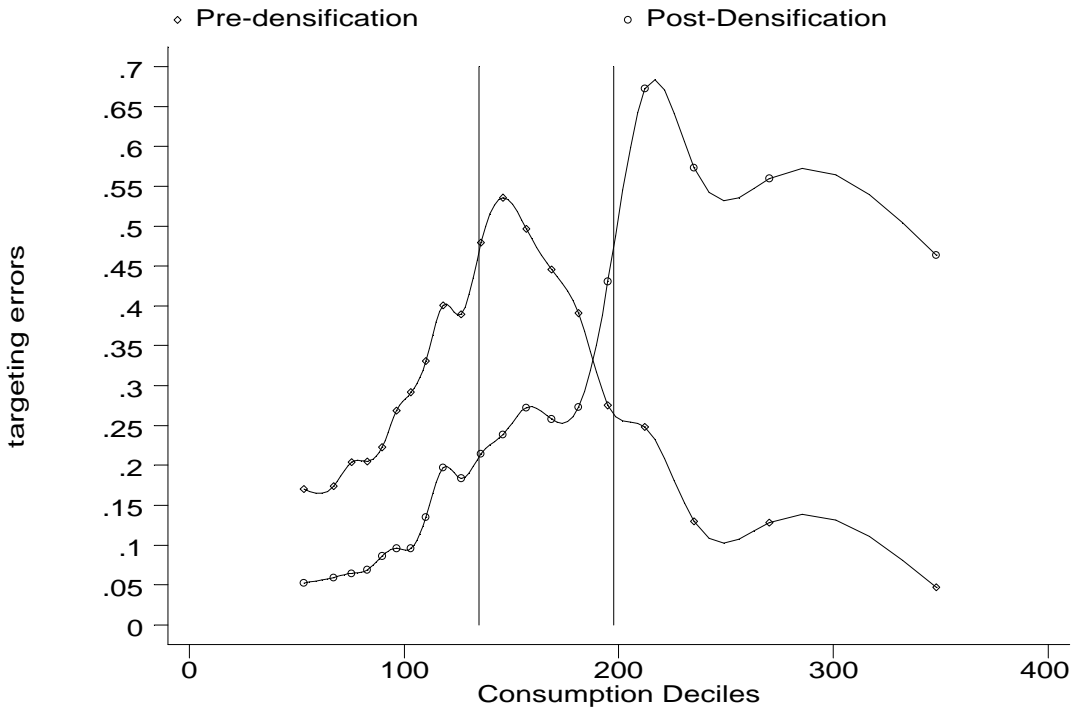
Figure 4—Poverty headcount rates across localities

49
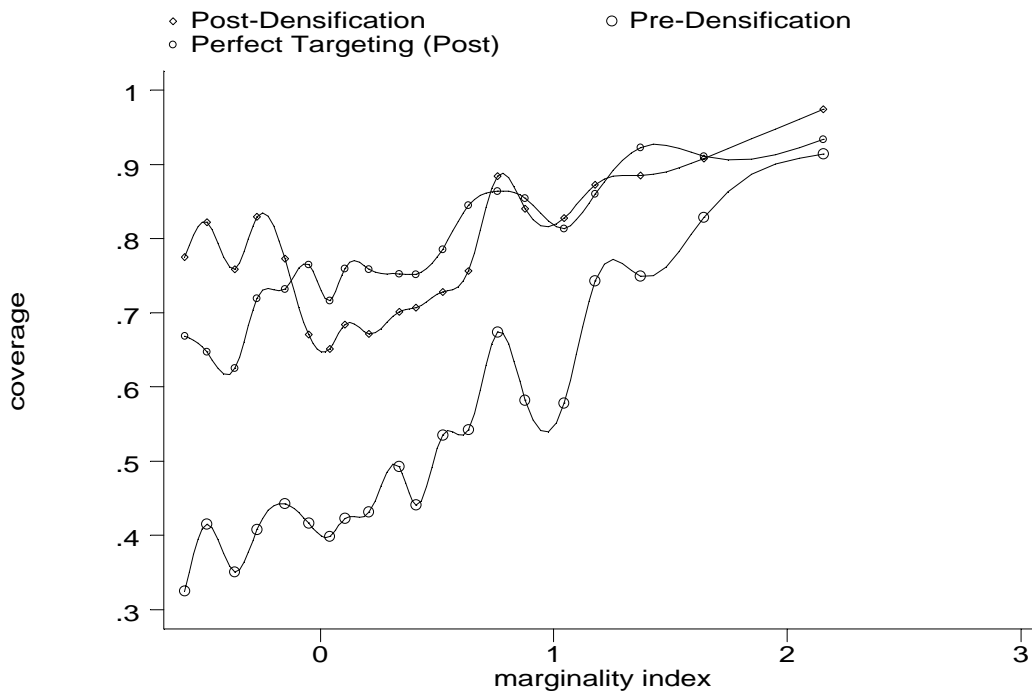
Figure 5—Relative welfare impact of alternative programs
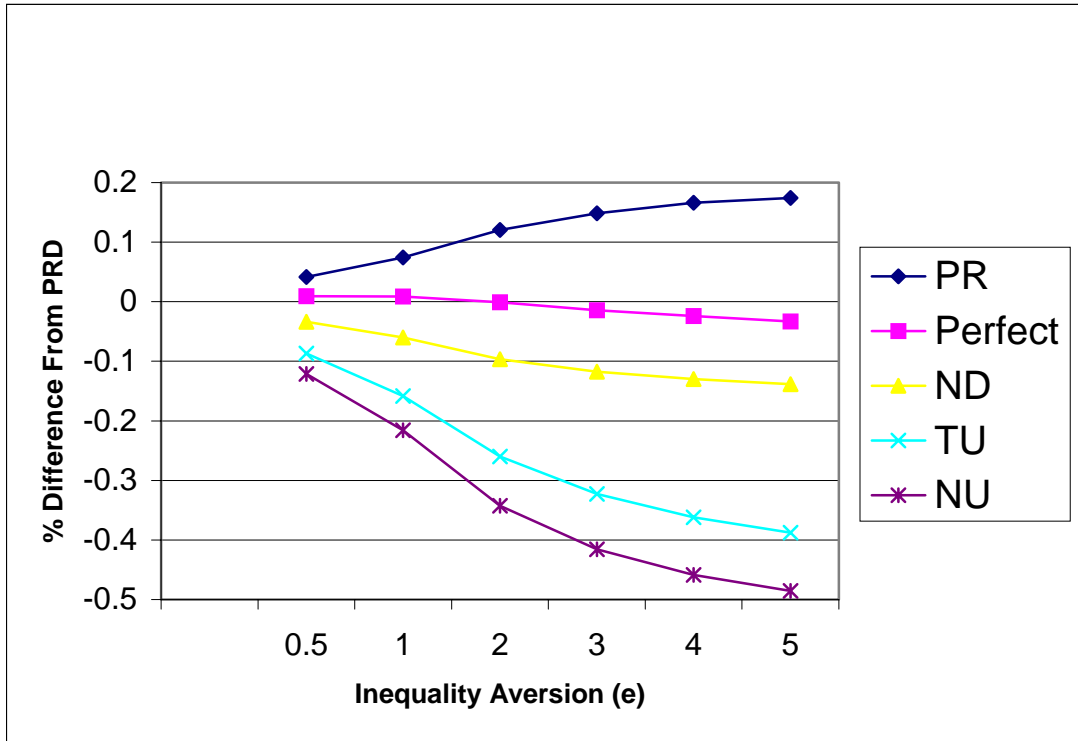
Note: Here we compare a range of program alternatives to the post-densification PROGRESA program that targets demographic transfers, including: the pre-densification program (PR), a perfectly targeted post-densification program (Perfect), nontargeted demographic transfers (ND), targeted uniform transfers (TU) and nontargeted uniform transfers (NU).

Figure 6—Proportion of capped households, by consumption group
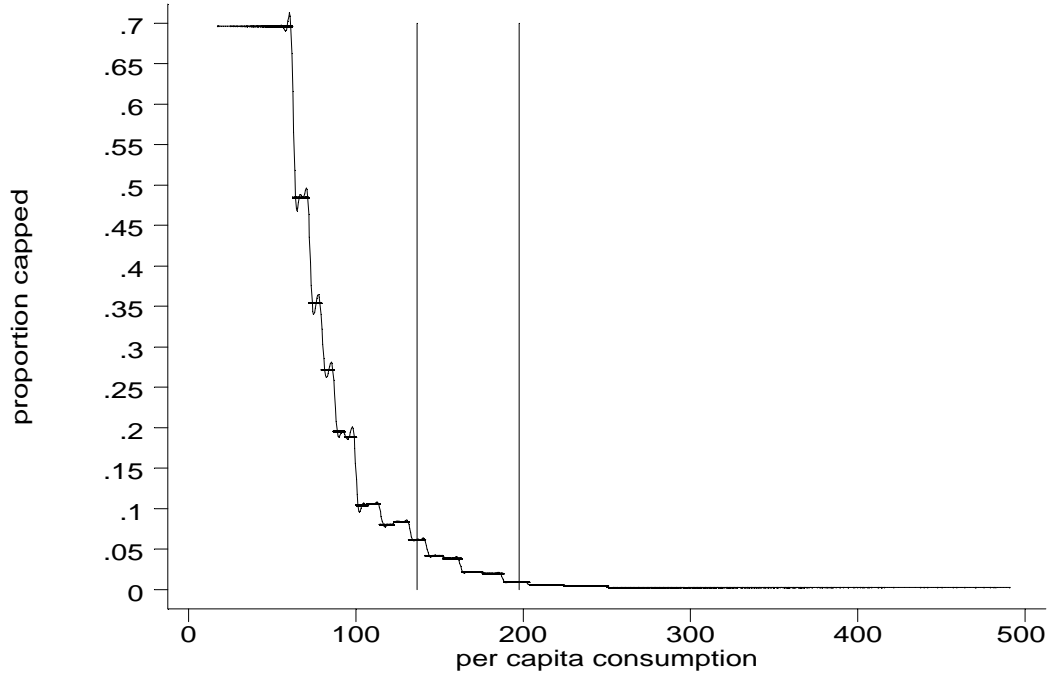
Figure 7—Average level of capped and uncapped benefits, by consumption group
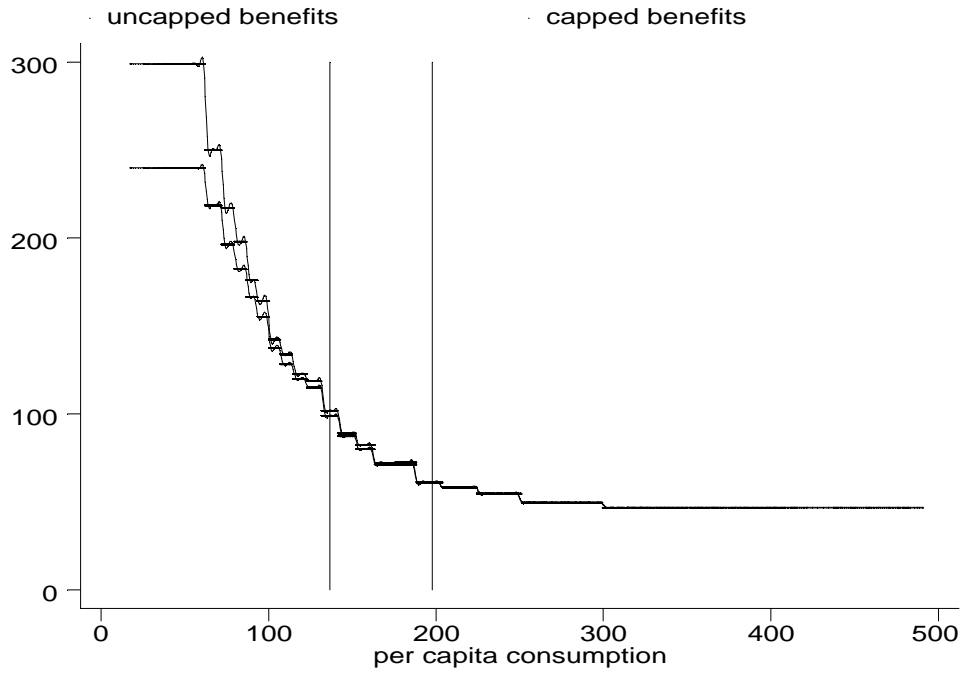
Figure 8a—Lambdas for uncapped components
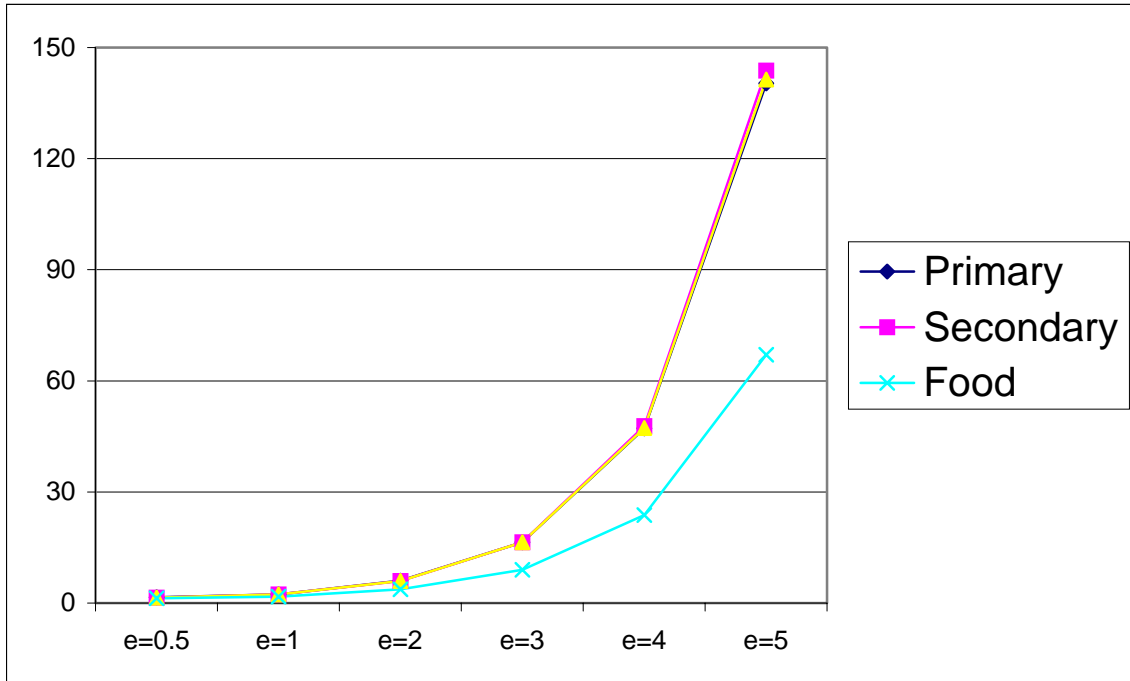
Figure 8b—Lambdas for capped components
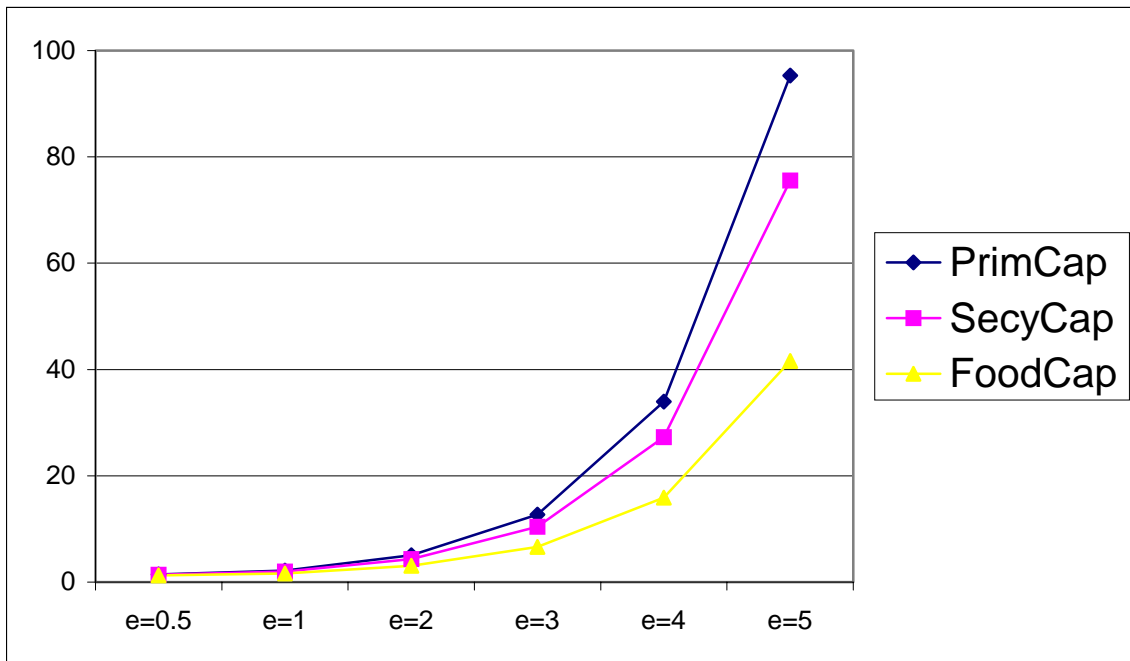
Figure 9—Kernel density of gains from targeting across localities (actual, e = 2)

dgain2

Kernel Density Estimate
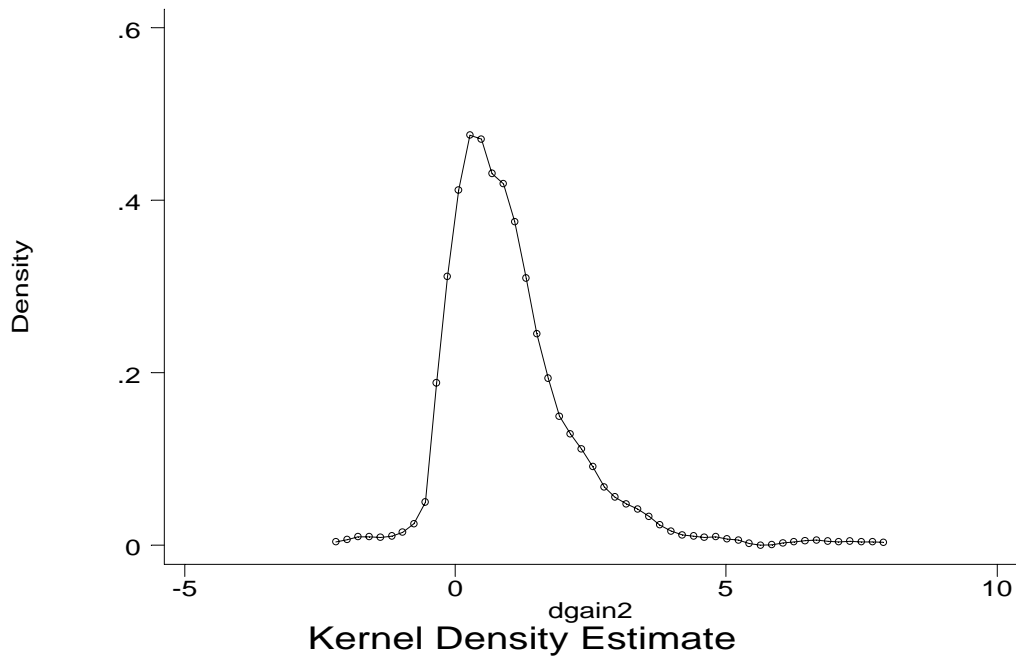
Figure 10—Kernel density of gains from targeting across localities (consumption, e = 2)

dgainC2

Kernel Density Estimate
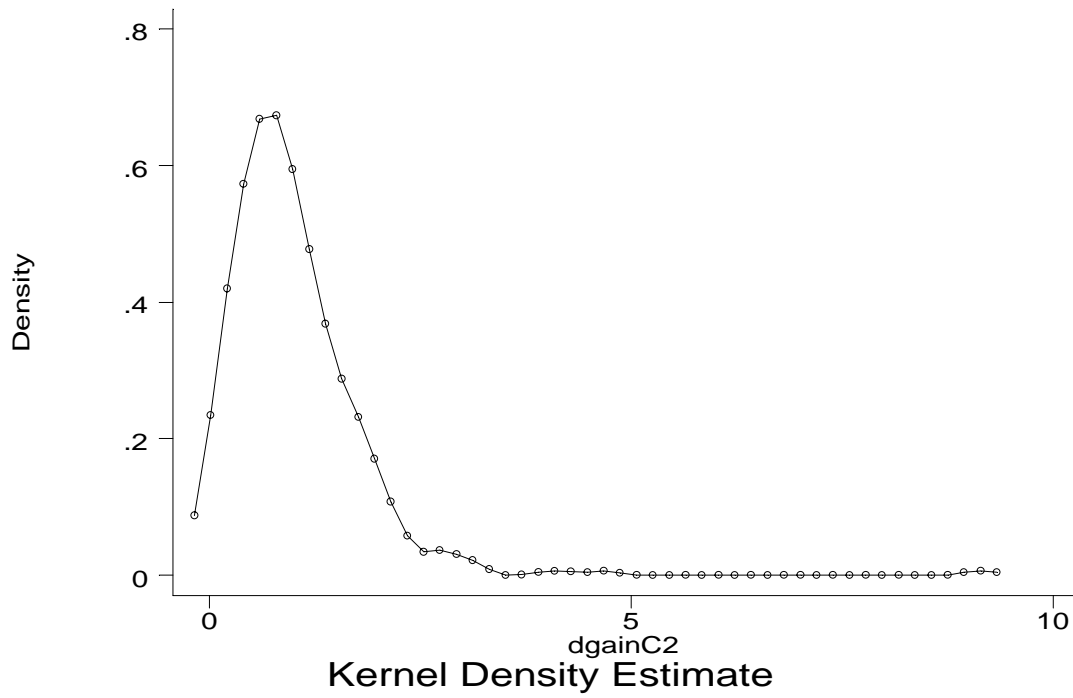
Figure 11—Welfare gains from household targeting across locality marginality (e = 2)

Figure 12—Welfare gains from targeting with/out program costs (actual program)
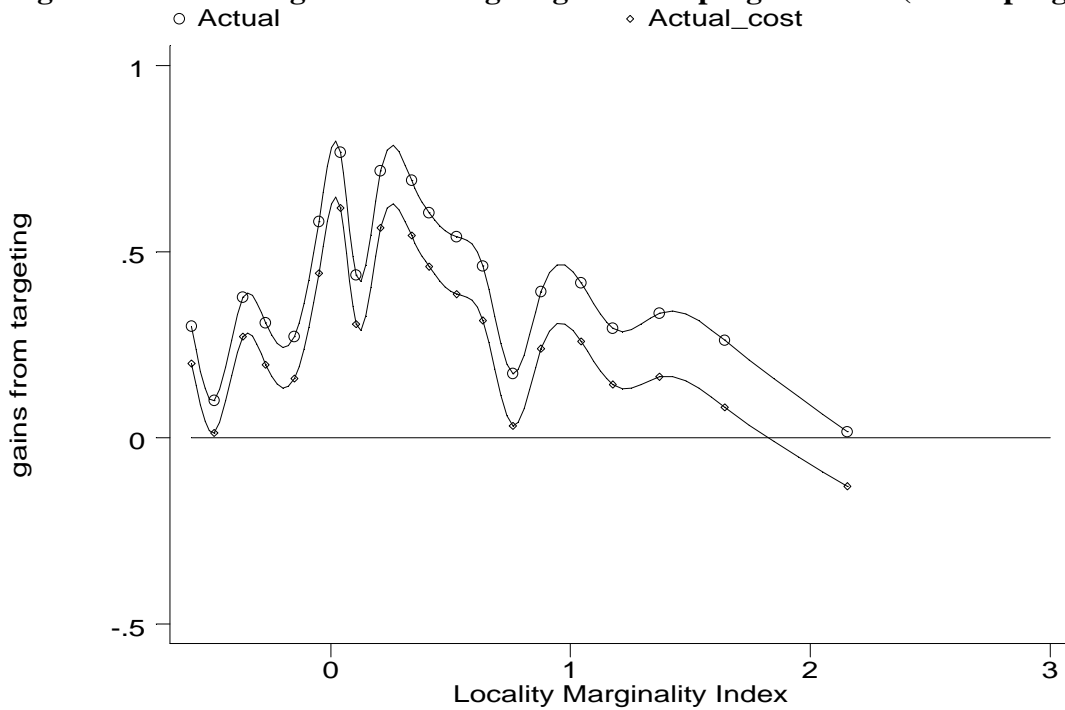
Figure 13—Welfare gains from targeting with/out program costs (perfect targeting)

Figure 14—Impact of take-up on distributional power of program
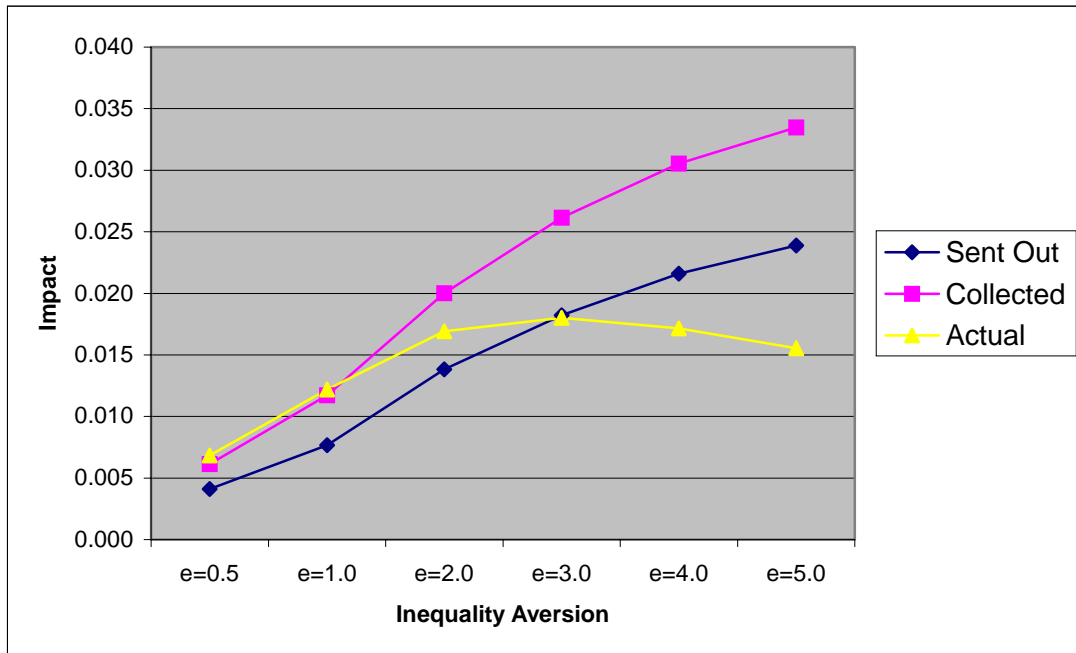
REFERENCES

Adato, M. 2000. The impact of PROGRESA on community social relationships. Report submitted to PROGRESA. International Food Policy Research Institute, Washington, D.C.

Ahmad, E., and N. Stern. 1984. The theory of reform and Indian indirect taxes. *Journal of Public Economics* 25: 259-298.

Ahmad, E., and Stern, N. 1991. *The theory and practice of tax reform in developing countries*. Cambridge: Cambridge University Press.

Atkinson, A. 1970. On the measurement of inequality. *Journal of Economic Theory* 2: 244-263.

Atkinson, A., and J. Stiglitz. 1980. *Lectures in public economics*. New York: McGraw-Hill.

Auerbach, A. 1985. The theory of excess burden and optimal taxation. In *Handbook of public economics*, vol. 1, ed. A. Auerbach and M. Feldstein. North-Holland: Amsterdam.

Behrman, J., and P. Todd. 2001. Progressing through PROGRESA: An impact assessment of Mexico's school subsidy experiment. International Food Policy Research Institute, Washington, D.C. Photocopy.

Coady, D. 2000. The application of social cost-benefit analysis to the evaluation of PROGRESA. Report submitted to PROGRESA. International Food Policy Research Institute, Washington, D.C.

Coady, D., and J. Drèze. 2001. Commodity taxation and social welfare: The generalized Ramsey rule. *International Tax and Public Finance*, forthcoming.

Coady, D., and R. Harris. 2000. General equilibrium analysis of the welfare impact of PROGRESA transfers. 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 transfers: Illustrations for Mexico. Food Consumption and Nutrition Division Discussion Paper 100. International Food Policy Research Institute, Washington, D.C.

Coady, D., R. Perez, and H. Vera-Llamas. 2000. An analysis of PROGRESA's costs: Alternative program designs and policy issues. In The application of social cost-benefit analysis to the evaluation of PROGRESA, by D. Coady. Report submitted to PROGRESA. International Food Policy Research Institute, Washington, D.C.

Deaton, A., and S. Zaidi. 1999. *Guidelines for constructing consumption aggregates for welfare analysis*. Washington, D.C.: World Bank.

Drèze, J., and N. Stern. 1987. The theory of cost-benefit analysis. In *Handbook of public economics*, Vol. II, ed. A. Auerbach and M. Feldstein. Amsterdam: North Holland.

Eichhorn, W. 1988. On a class of inequality measures. *Social Choice and Welfare* 5: 171-177.

Elbers, C., J. Lanjouw, and P. Lanjouw. 2001. Welfare in villages and towns: Micro-level estimation of poverty and equality. Photocopy.

Feldstein, M. 1972. Distributional equity and the optimal structure of public prices. *American Economic Review* 62 (1): 32-36.

Kolm, S.-C. 1976. Unequal inequalities I and II. *Journal of Economic Theory* 12: 416-442 and 13: 82-111.

Myles, G. 1995. *Public economics*. Cambridge: Cambridge University Press.

Newbery, D., and N. Stern, eds. 1987. *The theory of taxation for developing countries*. Oxford: Oxford University Press.

Ravallion, M. 1993: *Poverty comparisons: A guide to concepts and methods*. Living Standards Measurement Study Working Paper 88. Washington, D.C.: World Bank.

Sandmo, A. 1976. Optimal taxation: An introduction to the literature. *Journal of Public Economics* 6: 37-54.

Schultz, T. P. 2001. School subsidies for the poor: Evaluating a Mexican strategy for reducing poverty. Food Consumption and Nutrition Division Discussion Paper 102. International Food Policy Research Institute, Washington, D.C.

Skoufias, E., and D. Coady. 2001. Are the welfare losses from imperfect targeting important? Food Consumption and Nutrition Division discussion paper. International Food Policy Research Institute, Washington, D.C. Forthcoming.

Skoufias, E., B. Davis, and J. Behrman. 1999. The evaluation of the selection of beneficiary households in the education, health, and nutrition program (PROGRESA) of Mexico. International Food Policy Research Institute, Washington, D.C. Photocopy.

Stiglitz, J. 1988. *The economics of the public sector*, 2nd edition. New York: Norton.

FCND DISCUSSION PAPERS

FCND DISCUSSION PAPERS

FCND DISCUSSION PAPERS

FCND DISCUSSION PAPERS

FCND DISCUSSION PAPERS