IFPRI

FCND DP No. 125

FCND DISCUSSION PAPER NO. 125

ARE THE WELFARE LOSSES FROM IMPERFECT TARGETING IMPORTANT?

Emmanuel Skoufias and David 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

January 2002

ABSTRACT

We evaluate the size of the welfare losses from using alternative "imperfect" welfare indicators as substitutes for the conventionally preferred consumption indicator. We find that whereas the undercoverage and leakage welfare indices always suggest substantial losses, and the poverty indices suggest substantial losses for the worst performing indices, our preferred welfare index based on standard welfare theory suggests much smaller welfare losses. We also find that we cannot reject the hypothesis that the welfare losses associated with using the better performing alternative indicators are zero. In the case of our preferred welfare index, this reflects the fact that most of the targeting errors, i.e., exclusion and inclusion errors, are highly concentrated around the poverty line so that the differences in welfare weights between those receiving and not receiving the transfers are insufficient to make a difference to the overall welfare impact. Our results appear to be robust to the aversion to inequality assumed, as well as across the various welfare indices.

CONTENTS

TABLES

FIGURES

ACKNOWLEDGMENTS

1. INTRODUCTION

Over the last decade or more, developing countries have placed much more emphasis on targeted programs as part of their overall poverty alleviation strategy (Grosh 1994). This preference for targeting reflects an increased bias against "universal" programs, e.g., food subsidies that, because they are badly targeted, result in substantial "leakage" of the poverty budget to nonpoor households (Cornia and Stewart 1995). Poor targeting may result in a much smaller impact on the welfare of low-income (poor) households. In countries where poverty alleviation budgets, and social expenditures in general, are being cut in response to budgetary crises, these concerns tend to be magnified.

In order to efficiently target transfers to households, one needs an observable indicator that is highly correlated with program objectives, in this case, household welfare (Besley and Kanbur 1993). This paper evaluates the range of indicators that has been used in practice or suggested in the literature. As a starting point—and consistent with the literature, we take household consumption as our "ideal" indicator of household welfare (Ravallion 1992; Deaton 1997; Deaton and Zaidi 1999).[1] This standard is perceived in the economics literature as a better indicator of lifetime welfare, i.e., "permanent income" or persistent poverty. However, household consumption is also

---

[1] Household consumption captures only one of a number of important dimensions of welfare, namely the ability of households to purchase goods through markets. But it is an important dimension that is commonly focused on in both policy analysis and the relevant literature. For a more complete welfare analysis, one may wish to supplement such information with data on access to public goods that cannot be purchased through markets (especially where access is not highly correlated with income or consumption), or even with indices of "capability" (Sen 1992).

perceived as time-consuming and expensive to measure. Thus, it is not always available from household surveys and one is often forced to rely on alternative indicators that can be constructed with data that already exist or can be more easily or cheaply collected. When evaluating these alternative indicators as targeting variables, one needs to address the trade-off between the inevitable targeting errors that will result and the cost savings. In this paper, we are concerned with the former.

The structure of the paper is as follows. In the next section, we briefly discuss the range of *indicators* that we consider in this paper as alternatives to consumption as a basis for targeting program benefits. These are motivated by their greater availability, by their ease and lower cost of collection, or simply by the fact that they are commonly used or suggested for use in the literature or by policymakers. In Section 3, we set out the methodology used to evaluate the welfare losses resulting from use of these alternative, "imperfect" welfare indicators. Our preferred approach is firmly based within standard welfare theory, but we also incorporate other welfare *indices* that are commonly used in the literature, namely, various poverty indices and indices of "undercoverage" and "leakage." The data used to simulate program interventions are based on 4,378 rural and 9,001 households in Mexico surveyed as part of the 1996 National Survey of Household Income and Consumption (ENIGH). Our results are presented in Section 4, while Section 5 summarizes and concludes.

2. ALTERNATIVE INDICATORS

In this section, we set out the range of indicators that we will evaluate in Section 3. To focus attention, and to relate to a real-life poverty alleviation program, we use the design of the main poverty program, the Programa Nacional de Educación, Salud y Alimentación (PROGRESA), recently implemented by the Mexican government. This program involves cash transfers to households, with the amount of the transfer depending on certain demographic characteristics of the household, i.e., the age and gender of each child. Cash transfer programs similar to PROGRESA, which are conditioned on regular school attendance and visits to health centers, are finding much favor in other Latin American countries. Examples include the Programa de Asignación Familiar (PRAF) in Honduras, and *Red de Protección Social* in Nicaragua, the *Familias en Acción* in Colombia, and the *Bolsa Alimentacao* and *Bolsa Escola* in Brazil.[2]

We consider a range of alternative indicators that can be used in practice to target or select households for participation in the program. The "gold standard" household welfare indicator, against which all other indicators are compared, is reported total consumption per adult equivalent. Consumption is widely used in the literature reflecting, in part, both the theoretical and empirical support for household consumption as a better

---

[2] In reality, receipt of transfers is also conditional on enrollment and attendance of the child at school. However, in this paper, we ignore the human capital objectives of the program and focus solely on the objective of alleviating "current poverty." The presence of such "multiple objectives" often lies behind the use of transfer rules that are not optimal from a more narrowly defined income perspective of welfare. But such transfer schemes may also reflect the recognition of the measurement error present in all indicators as well as of social, political, and administrative constraints. In any case, the approach outlined below could be very easily extended to include comparisons with such optimal transfer schemes.

measure of household welfare over the life cycle, or a better indicator of persistent poverty. This reflects the fact that it is thought to be less susceptible to seasonal (or inter-temporal) variation. It is also thought to be less susceptible to underreporting by households. However, consumption data are also more expensive and time-consuming to collect, which explains the wider availability of the alternative indicators discussed below. In calculating it, we include both food and nonfood expenditures as well as the value of food items consumed out of own production.[3] These values are adjusted for inflation over the six-month frame of reference for the survey.[4] The measure also excludes loans made and durable assets purchased by the household. Since there is no regional price deflator available for Mexico, we do not account for cost-of-living differences across regions.

The indicators evaluated are:

1. *Reported expenditures*: Very often data on consumption out of own production (autoconsumption) are not available, so one has to fall back on this proxy.
2. *Food share*: This is derived as the share of total consumption accounted for by various foods. The food share occupies a prominent place in the economic literature as an indicator of household welfare (e.g., see Deaton 1997).

---

[3] We have also eliminated extreme values in per-adult equivalent expenditures by dropping values less than the 1 percentile and greater than the 99th percentiles of the distribution. This reduced our rural sample from 4,466 to 4,378 households and our urban sample from 9,183 to 9,001 households.

[4] Since substantial price variation typically exists between urban and rural areas, we focus on these subsamples separately in our empirical analysis presented later. See Hentschel and Lanjouw (1996) and Deaton and Zaidi (1999) for more detailed discussion on the construction of an appropriate consumption variable.

3. *Reported income:* Many surveys only collect income, as opposed to consumption, data. But the collection of income data has its own problems, including seasonal variability and its myriad of sources in developing countries, and especially in agriculture.

4. *Probability of being poor*: Policymakers responsible for Mexico's PROGRESA program used the following alternative to reported income. They first constructed a binary variable that took the value 1 if the household fell below a certain income poverty line. Then discriminant analysis was carried out with this binary variable as the dependent variable and variables describing the age and gender composition of the household, the level of schooling and occupation of the household head, and the housing and asset holdings.[5] The discriminant score was then used as the basis for categorizing households as poor and nonpoor, with the poor receiving the transfers. Since logit analysis is more widely used in the economics literature when dealing with binary variables, we construct a poverty indicator using the same procedure as above but using logit and the predicted probability of being poor in place of discriminant analysis and the discriminant score.[6]

5. *Asset index*: Most surveys collect information on the range of assets held by households, this reflecting their wealth stock (as opposed to the income flow from

---

[5] It should be noted that PROGRESA actually first applied geographic targeting by identifying marginal localities where poor households were more likely to be located. See Skoufias, Davis, and Behrman (1999).

[6] A comparison of the logit versus discriminant analysis revealed that two methods yielded practically identical results in the selection of beneficiary households.

that wealth). Arguably, ownership of assets is also more easily verified. Including such "assets" as access to public services also gives the index a multidimensional flavor. Using principal components analysis, we derive the first principal component that allows us to construct an asset index (Sharma 1994). This approach essentially determines the weights for a composite index of household assets and characteristics. An asset index *AI* is built for each household denoted by the subscript *j* from the first principal component using the formula:

$$AI_j = f_1 \times \left([X_{j1} - X_1] \Big/ S_1 \right) + \ldots + f_N \times \left([X_{jN} - X_N] \Big/ S_N \right),$$

where f_1 is the scoring or weighting factor for the first of the *N* variables, X_{j1} is the *j*th household's value for the first variable, and X_1 and S_1 are the mean and standard deviation of the first variable over all households. It should be noted that the first principal component explained only 14 percent of the total variance in the data. Although we include a very wide range of assets, the absence of information on landholdings means that our assessment of this indicator can be viewed as incomplete.[7]

All the above welfare indicators are viewed as imperfect proxies for the "ideal" consumption indicator. Since these are commonly used or suggested as possibilities in the

---

[7] However, a number of studies have found landholdings to be a poor indicator of welfare (see, for example, Ravallion 1989).

literature or in policy discussions, and are more widely available, it is important to have information on the trade-off in terms of the welfare losses associated with using these imperfect indicators in the place of consumption.[8]

3. METHODOLOGY AND DATA

In this section we set out the methodology employed in the paper to evaluate the alternative indicators described above. We start by assuming a given poverty-alleviation budget that is to be disbursed to "poor" households. Ideally we would like to use household consumption information to choose which households are to be classified as poor and thus receive the transfer. However, in practice, such information is often not available or is deemed too costly to collect, so we need to use another indicator that is only imperfectly correlated with household consumption. We can thus view the use of these alternative indicators as different (competing) programs and evaluate them accordingly.

In order to motivate the preferred welfare index used in our empirical analysis, we first present a very simple model of social welfare maximization by the government. In an ideal world with perfect knowledge of the welfare function of each household, the objective of the "social planner" may be specified as choosing the size and value of the transfer to each household so as to maximize social welfare subject to the constraint that

---

[8] For example, Filmer and Pritchett (1998) conduct a detailed study of the usefulness of an asset-based index in place of consumption expenditures.

the total amount to be disbursed, i.e., budget available for fighting poverty, is fixed at an

exogenously determined level B.[9] Specifically, social welfare is specified as a function of

household welfare, $V(\mathbf{p}, y)$, where $\mathbf{p}$ is the vector of commodity and factor prices faced by

the household and y is lump-sum transfers from the government. The Lagrangean

function for the planner's problem can thus be written as choosing a set of values y^h for

each household h so as to

$$\max \Psi = W(..., V^h(\mathbf{p}, y^h),...) + \lambda \left[B - \sum_h y^h \right], \tag{1}$$

where $W(.)$ is the concave social welfare function and λ is the Lagrange multiplier

associated with the budget constraint. This specification is essentially the specification

for the determination of the optimal cash transfer that maximizes social welfare.[10] As is

well known, the solution to this optimization problem is determined from the first order

necessary conditions:

$$d\Psi = \frac{\partial W}{\partial V^h} \frac{\partial V^h}{\partial y^h} dy^h - \lambda \, dy^h = \beta^h dy^h - \lambda \, dy^h = 0, \; \forall h, \tag{2}$$

which implies $\beta^h = \lambda^*$, for all h, where β^h is the social valuation of extra income to

household h, the so-called "welfare weight" of household h, and λ^* is the marginal social

---

[9] In this paper we are not concerned about the source of funds, e.g., taxing the richer households, for the budget allocated to poverty alleviation. See Coady and Skoufias (2000) for an example that includes these considerations into the model.

[10] The literature on optimal taxation is too large to even mention here. See Drèze and Stern (1987) and the references therein for more details.

value of budget at the optimum. In other words, at the optimum, the budget must be distributed such that the social valuation of income at the margin is constant across all households.[11] By summing across all households, the first-order conditions above can be rewritten as

$$l^* = \frac{\sum_h b^h dy^h}{\sum_h dy^h}.$$ (3)

Away from the optimum, β^h will, in general, differ across households. Also, one can interpret alternative income vectors $d\mathbf{y} = \{\ldots, dy^h, \ldots\}$ as representing alternative targeting schemes (denoted by j) for a given budget and calculate an associated l_j. Therefore, l_j will differ across targeting schemes both because the welfare weights β^h differ across households and the structure of income transfers $d\mathbf{y}$ differs across alternative targeting schemes.

Leaving aside, for the moment, the specification of the welfare weights, we can evaluate the welfare impact of the program (dW) as

$$dW = \sum_h b^h dy^h,$$

where h refers to households who receive transfers and the level of transfers is dy^h for household h. A program that transfers more of the budget to poor households, i.e.,

---

[11] Strictly speaking, this condition must hold only for the poorest households that receive transfers. It will hold for all households if the budget is endogenous.

households with relatively high β^h, will exhibit a higher dW, and thus will look

increasingly attractive as an income redistribution mechanism the greater one's concern

for the poorest of the poor. As indicated above, one can transform this statistic into a

more conventional benefit-cost ratio by dividing by the overall poverty budget to get, for

each program j:

$$\lambda_j \equiv \frac{\sum\limits_h b^h dy^h}{\sum\limits_h dy^h},$$

which can be interpreted as the marginal social value of a unit of revenue transferred to

households through the program in question. In our empirical analysis presented below,

we focus on λ_j as opposed to dW but, as indicated above, given the assumption of a fixed

budget, both are equivalent for evaluation purposes.[12]

Our first task then is to evaluate the welfare benefit or impact of each targeting

scheme. 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 rich) households. Making this view explicit essentially requires the

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

---

[12] Note, however, that we are not comparing alternative programs to an "optimal transfer scheme" as in, for example, Chaudhuri and Ravallion (1994). The λ for such a program will obviously be the highest attainable, but is hypothetical in so far as other factors, e.g., the existence of multiple objectives, in practice determine the structure of benefits. Both Chaudhuri and Ravallion (1994) and Schady (1999) focus on the minimum cost of achieving a given poverty impact across a range of transfer schemes, including an optimal transfer scheme. In any case, our approach can easily incorporate such a comparison.

(initial) consumption level of the household. The welfare weight for each household (β^h) can be derived as follows:

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

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).[13] The term ε 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 $\varepsilon = 0$ implies no aversion to inequality and all welfare weights take the value unity, i.e., extra income to households is viewed as being equally socially valuable regardless of initial consumption level. A value of $\varepsilon = 1$ 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 $\varepsilon = 2$ similarly implies a welfare weight of 0.25 (4.0) for h. 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 we care only about how the program benefits the poorest of the poor. The welfare weights used in our simulations presented below use

---

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

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 ε. Consistent with the program

objectives, we consider only values of $\varepsilon > 0$.

Alternatively, one can use conventional poverty measures as a measure of the

welfare impact, as captured by the FGT indices (Foster, Greer, and Thorbecke 1984):

$$P(\boldsymbol{a}) = \frac{1}{N} \sum_{h=1}^{q} (\frac{z - y^{h}}{z})^{\boldsymbol{a}} ,$$

where N is the number of households, y^{ih} is the per adult equivalent consumption of

household h, z is the poverty line, q is the number of poor households, and α is the weight

attached to the severity of household poverty (or the "poverty aversion parameter").

When $\alpha = 0$, the FGT measure collapses to the "poverty headcount index," i.e., the

percentage of the population falling below the poverty line. Since this measure tells us

nothing about how far below the poverty line these households are, it is common to focus

on $\alpha = 1$ or $\alpha = 2$, the "poverty gap index" and "severity of poverty index," respectively.

The former captures the average depth of poverty, while the latter attaches a higher

weight to transfers to households the further they fall below the poverty line (Besley and

Kanbur 1993; Atkinson 1995; Deaton 1997).

The poverty approach can be interpreted within the above framework as attaching a welfare weight of zero to everybody above the poverty line (i.e., regardless of distance above the poverty line) and also essentially to transfer amounts in excess of the minimum required to take a poor household up to the poverty line. The extent to which the welfare weights increase as income falls further below the poverty line is determined by the parameter α. The headcount index has a number of shortcomings as a welfare index, not least the fact that it suggests that resources should be concentrated on those just below the poverty line. Values of $\alpha \geq 1$ try to address this issue by attaching a greater weight to extra income to households the further they fall below the poverty line. However, a shortcoming still present in the $P(1)$ and $P(2)$ measures is that they do not differentiate between leakage (undercoverage) of the program to households that are just above (below) the poverty line and households that that are far above (below) the poverty line. Most would subscribe to the view that the social value of income to someone just above the poverty line is very close to that for someone just below it. Our specification of welfare weights as above takes into consideration this issue. Higher values of the parameter ε can adequately capture our greater concerns for the poorest of low-income households. Using the welfare theoretic approach we essentially use the concept of a poverty line solely as the basis of a targeting rule. But, for the sake of completeness, we also evaluate the alternative programs (or targeting indicators) from the perspective of poverty reduction.

Yet another approach commonly found in the literature is that of the extent of "leakage" and "undercoverage" associated with a transfer program (Cornia and Stewart 1995). These concepts are defined as:

Leakage: The percentage of the total beneficiary population that is wrongly classified as poor (i.e., errors of inclusion).

Undercoverage: The percentage of the poor population wrongly classified as nonpoor (i.e., errors of exclusion).

Although these measures have a number of shortcomings (e.g., they ignore how the levels of transfers vary across households), we also evaluate the programs from the perspective of these measures since they are so commonly presented in the literature.[14]

In our empirical analysis we wish to address the following questions: (1) How much do the welfare impacts of the various programs (i.e., based on alternative welfare indicators) vary? (2) How sensitive is this variation to the degree, or nature, of our aversion to inequality (i.e., different values of ε or α)? (3) How different are the results across the various welfare indices (i.e., welfare, poverty, and undercoverage/leakage indices)? (4) Are the welfare losses identified statistically significantly different from zero? (5) What are the features of the various indicators that make them more or less

---

[14] For an interpretation of these concepts within a welfare-theoretic framework, see Coady and Skoufias (2001).

attractive as targeting devices? (6) How sensitive are our results to the size of the program budget or to an urban-rural focus?

In order to evaluate the performance of the alternative indicators, we classify households as being (actually) poor, using our "gold standard" of household consumption and a (relative) poverty line drawn at the median of the rural sample (i.e., 50 percent of the rural population is assumed to be poor). The welfare weights are also calculated based on household per adult equivalent consumption. For each indicator, we then determine which households receive the transfers by taking those falling into the bottom 50th-percentile according to this indicator. The amount of transfer received is then determined by household composition according to the scheme explained in Appendix 1.[15] Using the various welfare indices, we then compare the welfare impact of the various programs (i.e., alternative indicators) with that which would result from "perfect targeting," i.e., using consumption.

4. SIMULATION RESULTS

In order to get some feel for the imperfect nature of these alternative welfare indicators, we start by constructing a graph that captures both the magnitude of the targeting errors associated with each indicator as well as the nature of this error, i.e., where in the distribution of consumption these errors manifest themselves. For each

---

[15] As noted earlier, for any set of welfare weights, these transfers are unlikely to be "optimal." The transfer scheme used reflects the multiple objectives of the Mexican PROGRESA program, which "conditions" transfers in order to provide incentives for the accumulation of human capital.

indicator we classify a household as "poor" if it is classified as a poor household given the fixed budget of the program, and as "nonpoor" if it is left out of the program. Assuming the same fixed budget, this classification is compared with that suggested by our reference indicator, i.e., consumption. [16] We construct a new variable that takes the value one when households that are classified as "poor" and "nonpoor" according to consumption are classified incorrectly as "nonpoor" and "poor," respectively, according to the alternative indicator (i.e., identifying errors of exclusion and inclusion, respectively). Otherwise, this variable takes the value zero. Using nonparametric methods we then plot the mean of this variable against the log of reported consumption per adult equivalent (Figures 1 and 2). The value on the y-axis can be seen as the "predicted error-probability" (PEP). The height of the curve captures the extent of the targeting errors being made at various points in the distribution. The shape captures where in the distribution these errors are being made. For example, a bell-shaped curve concentrated around the poverty line indicates that most of the error involves a misclassification of households that lies just above and below the poverty line.

Figure 1 plots the PEP curve associated with reported expenditure that excludes consumption out of own production, reported income, and the regression-based poverty indicator. From Figure 1 we can see that the most efficient targeting indicator is "reported expenditure." The fact that the curves for reported income and "the probability of being

---

[16] See Appendix 1 for a more detailed discussion of our simulations.

Figure 1
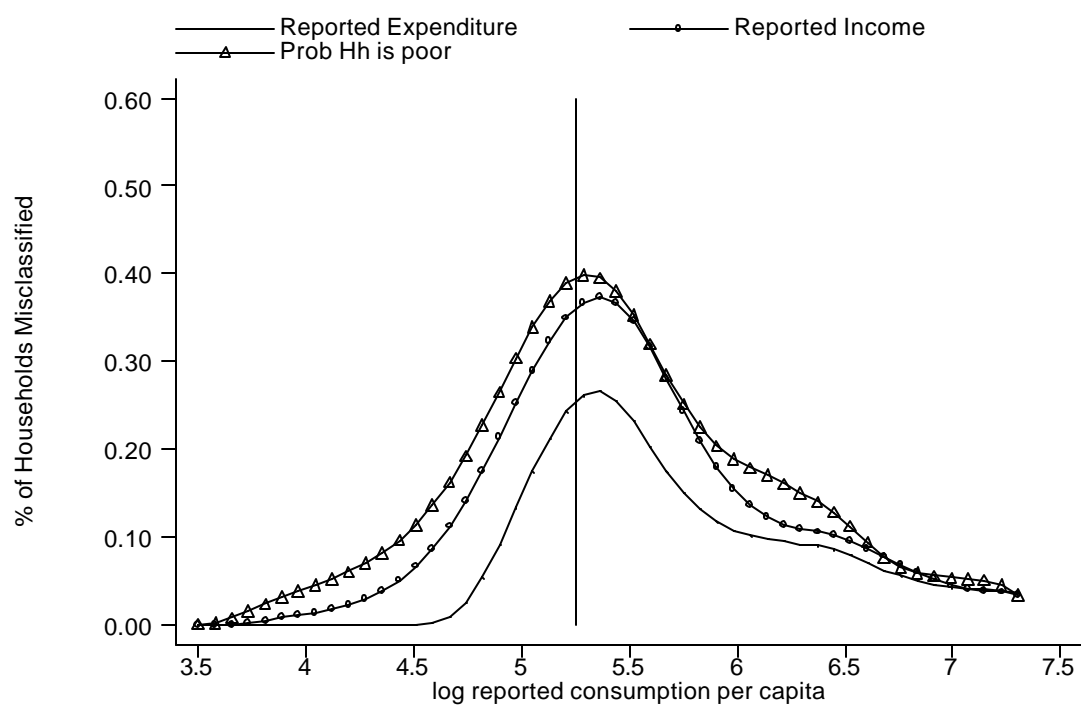
Figure 2
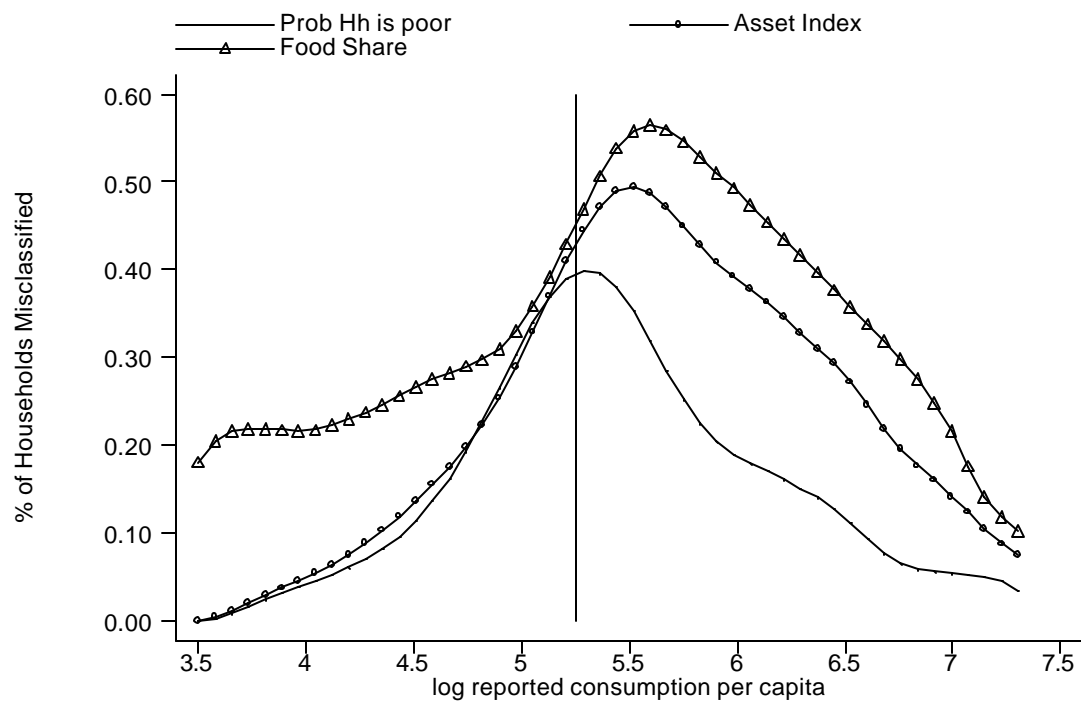
poor" indicator lie everywhere above that for reported expenditure tells us that not only is the proportion of targeting errors greater using the former, but they are also more costly from a welfare perspective, since a larger proportion of the poorest households are excluded and a larger proportion of the richest are included. Figure 2 plots the PEP curve associated with the "probability of being poor" indicator, which had the highest PEP curve in Figure 1, against the PEP curve for the asset index and the food share. From Figure 2 we can see that the most inefficient targeting indicator is the food share. Notice, also, that for the two worst-performing indicators, the probability of misclassifying households is highest at a level of consumption slightly higher than (or to the right of) the poverty line.

The preceding insights are reinforced by our undercoverage (U) and leakage (L) measures (see Table 1). Using both measures, the worst performers are food share and the asset index. For these two targeting indicators, leakage lies in the range of 35.1–41.7 percent, while undercoverage lies in the range of 21.5–28.9 percent. Even the best indicator, reported consumption expenditure, has an undercoverage rate of 9 percent and leakage slightly above 15 percent. These levels of mistargeting, especially the latter, are suggestive of large welfare losses from not being able to use reported total consumption to identify which households receive transfers.[17]

---

[17] If one includes that income "wasted" through excessively high levels of transfers, this would be even higher.

Table 1—Welfare and poverty indices under various targeting/transfer schemes
Sample: All households in RURAL areas
Poverty line: 50th percentile of reported consumption in RURAL areas

(percent difference from value of index using reported consumption)

	Lambda(5)	Lambda(2)	Lambda(1)	P(0) Headcount Index	P(1) Poverty Gap	P(2) Severity Index	U in percent	L in percent
No transfer (no anti-poverty program)				0.5000	0.1731	0.0798		
Targeting based on reported (total) consumption	59.897	3.639	1.786	0.368	0.091	0.032	0.00	0.00
Targeting based on reported expenditures	59.772 -(0.21)	3.583 -(1.56)	1.754 -(1.77)	0.407 (10.69)	0.095 (4.29)	0.032 (1.71)	9.00	15.02
Targeting based on reported income	58.334 -(2.61)	3.443 -(5.41)	1.698 -(4.93)	0.411 (11.81)	0.102 (11.90)	0.035 (11.39)	17.31	22.65
Targeting based on the probability of being poor	57.014 -(4.81)	3.3644 -(7.55)	1.6706 -(6.46)	0.4194 (14.11)	0.1057 (16.01)	0.0372 (17.82)	22.34	24.88
Targeting based on asset index	55.549 -(7.26)	3.237 -(11.06)	1.618 -(9.43)	0.415 (12.93)	0.107 (16.92)	0.038 (20.65)	21.47	35.11
Targeting based on the share of food in total consumption	47.368 -(20.92)	2.879 -(20.88)	1.505 -(15.72)	0.412 (11.99)	0.114 (25.43)	0.044 (40.37)	28.92	41.66

To examine the welfare losses due to inefficient targeting, in Figure 3 we graph for selected indices (i.e., $\lambda(5)$, $\lambda(2)$, $P(2)$, the undercoverage rate U, and the leakage rate L), the percentage difference from the value of the corresponding index with "perfect" targeting. Upon inspection of Figure 3, one can draw the following inferences.

The preferred welfare indices $\lambda(\varepsilon)$ suggest very small welfare losses, less than 5 percent for the two best targeting indicators, i.e., reported expenditures and reported

Figure 3
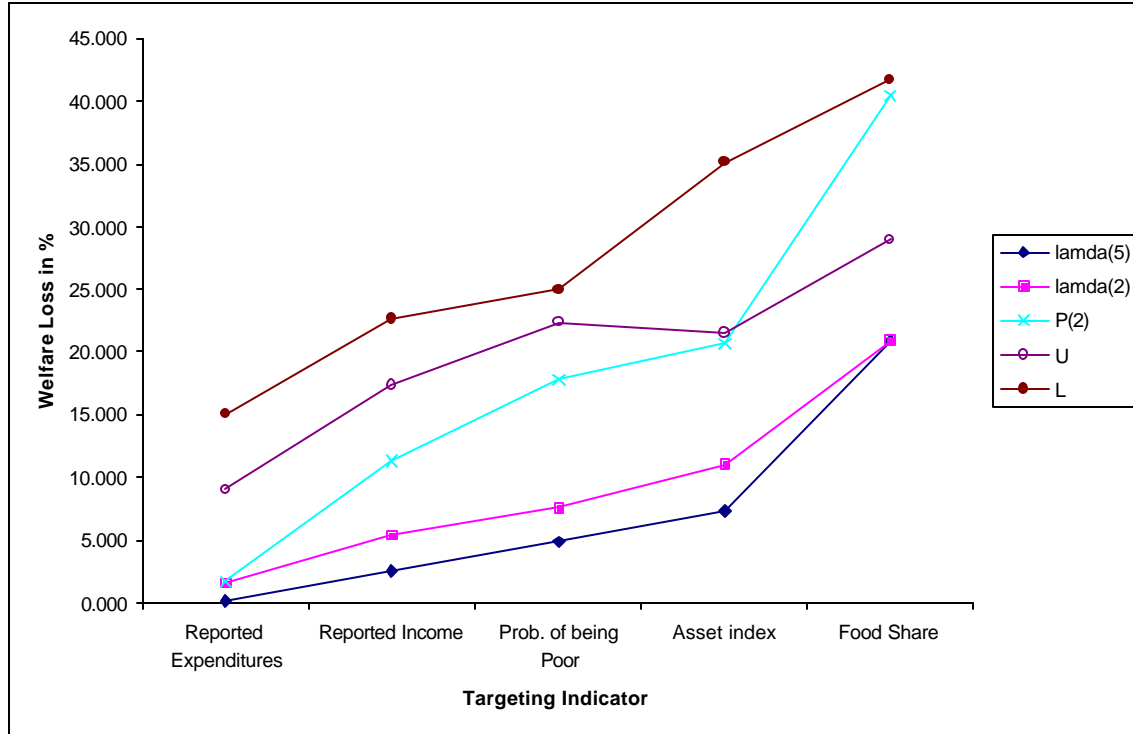
income. In particular, using $\lambda(5)$, it appears that the welfare losses from the best

performer among the alternative indicators are quite low, of the order of less than 1

percent. Considering the implications of this finding for data collection needs for

targeting purposes, we conduct further sensitivity tests and discuss our findings below.

The undercoverage and leakage indices consistently suggest substantially higher

welfare losses from imperfect targeting relative to the welfare losses suggested by our

preferred welfare index $\lambda(\varepsilon)$. For example, when reported expenditures are used as a

targeting indicator, the undercoverage (*U*) or leakage (*L*) welfare indices suggest that the

welfare losses are between 9 and 15 percent, respectively. In contrast, the size of the

welfare losses associated with $\lambda(2)$ and $\lambda(5)$ is 1.6 percent and less than 1 percent, respectively. This reflects the pattern observed earlier in Figure 1, which showed that, for this targeting indicator, the targeting errors were highly concentrated around the poverty line. In other words, the mistargeting involves transferring income between households with very similar welfare weights even for a very high degree of inequality aversion. Given that the undercoverage and leakage indices do not distinguish whether the incorrectly excluded or included households were close or far away from the cutoff point, it is no surprise that the welfare losses associated with these welfare indices appear to be much higher than the welfare losses associated with our preferred welfare indices.

The severity of poverty index $P(2)$ suggests lower welfare losses due to imperfect targeting relative to undercoverage and leakage and higher welfare losses relative to the preferred welfare index $\lambda(\varepsilon)$. Interestingly, the welfare losses suggested by $P(2)$ are especially high for the worst performing indicator, i.e., food share, and especially low (or practically the same as the low welfare losses suggested by the $\lambda(2)$ welfare index) for the best performing indicator, i.e., reported expenditures. Of course, the welfare losses suggested by any $P(\alpha)$ will be even higher, ceteris paribus, the larger the transfer to households since transfers higher than the poverty gap do not contribute to lowering $P(\alpha)$.

Irrespective of the welfare index used, the ranking of the welfare losses is insensitive to the targeting indicator used. For example, using $P(2)$ as an index of welfare implies that the lowest welfare losses are associated with the use of reported

expenditures. Using reported income yields a higher level of welfare losses than reported expenditures, while the probability of being poor yields even higher welfare losses, followed by the asset index and the food share. The same general ranking is obtained using leakage (*L*) as a welfare measure or $\lambda(2)$ and $\lambda(5)$. The use of undercoverage (*U*) as a welfare indicator also yields a similar ranking except for the way in which it ranks the welfare losses associated with the asset index relative to the probability of being poor.

In order to examine the extent to which these patterns hold when we use an urban sample of households, we have also estimated the welfare indicators of Table 1, using the sample of urban households and, as a poverty line, the 25th percentile of reported consumption of urban households (see Appendix 2). As can be easily seen, the same patterns observed for the rural sample also hold here. The two best-performing indicators are reported expenditures and reported income, and for these two indicators the welfare losses in the urban sample are as low as those observed for the same targeting indicators in the rural sample. For example, using $\lambda(5)$, the welfare losses are less than 5 percent as for the rural sample. Additional sensitivity tests were conducted by focusing on the sample of rural households and decreasing the budget used for our simulations to 75 and 50 percent of that used in the earlier simulations in rural areas. Since the patterns observed resembled those in Table 1, we chose not to present them.

Given our findings that the welfare losses across alternative indicators tend to cluster within small ranges for the best (or worst) performers, and since for our preferred welfare indices the losses for the best performers appear quite low, we investigated a

number of related issues in further detail. Specifically, we examined whether the welfare

loss associated with each imperfect targeting indicator are statistically different from zero

for any given welfare index, and whether the answers are sensitive to our aversion to

inequality. As a means of addressing these questions, we used 1,000 bootstrap samples

(each sample randomly selected with replacement). For each bootstrap sample, we

calculated the index value for each targeting indicator and, based on the 1,000 different

values of the indices, we estimated standard errors for each index. For any given targeting

indicator, our primary objective is to examine whether the value of the index is

significantly different from that of the welfare impact achieved using the preferred

indicator, i.e., reported consumption.

The mean values of the estimated indices along with their associated standard

errors are reported in columns 1 and 2 of Table 2. In column 3, we also report the p-

values obtained from formal (two-tailed) t-tests of the null hypothesis that value of the

index is equal to that using reported consumption. Focusing on $\lambda(5)$, we find that the

value for this index using reported expenditures is practically identical with the index

value of the preferred indicator. This implies that the welfare losses associated with using

expenditures that exclude auto-consumption are not significantly different from zero.[18]

This finding is especially interesting in consideration of the fact that consumption out of

own production is both prevalent and difficult and expensive to collect in rural areas.

---

[18] It is important to note, however, that any hypothesis test regarding specific values of the indices obtained
with these two targeting indicators will also have very low power.

Table 2—Mean values and standard errors for selected indices with alternative targeting indicators, based on 1,000 bootstrap rural samples

	1	2	3	4
			Ho: diff = 0	
Targeting indicator	**L(5)**	**Standard Error**	**Ha: diff > 0**	**Ha: diff = 0**
Reported consumption	**59.897**	**3.097**		
Reported expenditure	59.772	3.091	P > t = 0.183	P > t = 0.366
Reported income	58.334	2.978	P > t = 0.000	P > t = 0.000
PROGRESA-style logit	57.014	2.991	P > t = 0.000	P > t = 0.000
Asset index	55.549	2.977	P > t = 0.000	P > t = 0.000
Food share	47.368	2.725	P > t = 0.000	P > t = 0.000
			Ho: diff = 0	
Targeting indicator	**L(2)**	**Standard Error**	**Ha: diff > 0**	**Ha: diff = 0**
Reported consumption	**3.639**	**0.090**		
Reported expenditure	3.583	0.088	P > t = 0.000	P > t = 0.000
Reported income	3.443	0.083	P > t = 0.000	P > t = 0.000
PROGRESA-style logit	3.364	0.084	P > t = 0.000	P > t = 0.000
Asset index	3.237	0.084	P > t = 0.000	P > t = 0.000
Food share	2.879	0.076	P > t = 0.000	P > t = 0.000

Comparing the remaining indicators with reported consumption, we find that the hypothesis that there are no welfare losses associated with using alternative targeting indicators is consistently rejected.[19] This finding suggests that using indicators based on income or food share or assets is likely to result in significantly higher welfare losses.[20] Nevertheless, the better performing indicators, such as reported income and the probability of being poor, exhibited welfare losses that were significantly lower than the rest of the indicators.

---

[19] Chaudhuri and Ravallion (1994) also find that consumption is not necessarily the best indicator of their dynamic concept of "persistent" or "chronic" poverty.

[20] We have also experimented with the share of cereals in total consumption and found that it also performs as poorly at the share of food in total consumption.

Generally similar results hold when focusing on $\lambda(2)$. One notable difference is that using $\lambda(2)$ as a measure of welfare yields that the welfare losses associated with using expenditures instead of reported consumption are now significantly higher. Based on this evidence we conclude that the degree of aversion to inequality has an impact on the welfare losses associated with specific targeting indicators.

5. SUMMARY AND CONCLUSIONS

In this paper we make an effort at quantifying and comparing the size of the welfare losses from using alternative "imperfect" welfare indicators as substitutes for the conventionally preferred consumption indicator. We find that the size of the welfare losses associated with different indicators varies considerably. Our preferred welfare index implies that the losses from the two best targeting indicators (reported expenditures and reported income) are very low if not trivial (less than 5 percent). Moreover, the welfare losses suggested by our preferred welfare index are always lower than those suggested by the poverty, undercoverage, and leakage welfare indices. In contrast, the welfare losses suggested by undercoverage and leakage indices are substantially higher, while those based on poverty indices are relatively high for the worst performing indicator (food share). In the case of our preferred welfare index, this reflects the fact that most of the targeting errors (exclusion and inclusion) are highly concentrated around the poverty line; thus, the differences in welfare weights between those receiving and not

receiving the transfers are insufficient to make much of a difference to the overall welfare impact.

Recognizing that welfare indices are subject to sampling error leads us to conclude also that there are significant welfare losses associated with different targeting indicators. An asset-based index and the share of food as targeting indicators were found to have the highest welfare losses relative to all other targeting indicators examined in this paper. Although there may be room for improvement in our construction of a "gold standard" consumption indicator, it is not obvious that improvements in this direction would overturn the conclusions drawn here. We also find, based on our preferred welfare indicator, that whether we reject the hypothesis that there are no welfare losses associated with using the better performing alternative indicators, e.g., reported expenditures or reported income, depends on the extent of aversion to inequality. It may also be that the profile of consumption and other household characteristics are so different across countries that these results are country-specific. In future work we hope to test the robustness of our conclusions across countries.

APPENDIX 1

Simulating the Impact of the Various Targeting and Transfer Schemes

THE STRUCTURE OF TRANSFERS

In this appendix, we describe briefly the main steps we have taken for conducting our simulations. We have used the amounts, and age and gender structure of the benefits to be similar to that of the PROGRESA program. Specifically, the level of primary school benefits received by each potentially participating household is determined as follows:

Primary School benefits at the household level =

(number of boys and girls of 8 yrs of age) * 60P/month +

(number of boys and girls of 9 yrs of age) * 70P/month +

(number of boys and girls of 10 yrs of age) * 90P/month +

(number of boys and girls of 11 yrs of age) * 120P/month.

The level of secondary school benefits received by each potentially participating household is determined as follows:

Secondary school benefits at the household level =

(number of boys 12-14 yrs of age) * 175P/month +

(number of girls 12-14 yrs of age) * 185P/month +

(number of boys 15-16 yrs of age) * 185P/month +

(number of girls 15-16 yrs of age) * 205P/month +

(number of boys 17-18 yrs of age) * 195P/month +

(number of girls 17-18 yrs of age) * 225P/month.

We then summed the benefits from having all children enrolled either in primary and secondary school grades with the fixed allowance of 115P/month given to PROGRESA beneficiaries. For households that the total exceeded the maximum of 695P/month as allowed by PROGRESA, we replaced the benefit that could be received by the amount of 695P.

Finally, we added to the total cash transfer the allowances given to households for school utilities of the children, these been equal to the number of children in primary school*(135P/12) + number of children enrolled in secondary school*(170/12). The school allowances were divided by 12 since these are given on an annual, not monthly, frequency. The cash transfer of each household was discounted to July 1994 prices in order to make them comparable to the consumption expenditures.

BUDGET SIZE

We used the 50th percentile of the reported consumption per adult equivalent as the cutoff point below which a household is classified as a beneficiary. Given a cutoff line, we then derive the value of the budget used in all of our simulations by giving each beneficiary the benefits of the PROGRESA program. The average transfer was 210 pesos per adult equivalent in June 1994 prices while the total budget amounted to 52.13 percent

of the poverty gap. The poverty gap is defined as the sum (across all poor households) of the difference between the poverty line and the per-adult equivalent consumption expenditure of the household both multiplied by the number of adult-equivalent units in the household.

CLASSIFYING HOUSEHOLDS AS POOR WITH ALTERNATIVE TARGETING INDICATORS

Next we employed the alternative indicators used to classify households (discussed in the main paper). Each of these indicators allows us to sort or rank all the households in the sample from the lowest to the highest value. Cash transfers are given to households beginning with households having the lowest value of the indicator to progressively higher values until the budget is exhausted. The households that end up receiving program benefits are the poor, while those left out because of the limited budget are the nonpoor.

NONPARAMETRIC GRAPHS

In deriving the nonparametric graphs, we follow the methods used by Subramanian and Deaton (1996). The procedure works as follows. At any given point x, we calculated a weighted mean of the variable identifying whether a household is misclassified or not on the log of reported per-adult equivalent household consumption. The weights are chosen to be largest for sample points close to x and to diminish with distance from x; they are also set so that, as the sample size increases, the weight given to

the immediate neighborhood of x is increased so that, in the limit, only x is represented. In our case, we chose an evenly spaced grid of 50 points in the distribution of the predicted log per adult equivalent consumption and for each grid x, observation i gets the (quartic kernel) weight

$$w_i(x) = \frac{15}{16}\left[1 - \left(\frac{x - x_i}{h}\right)^2\right]^2$$

if $-h \leq x - x_i \leq h$ and zero otherwise. The quantity h is a bandwidth that is set so as to trade off bias and variance, and that tends to zero with the sample size. We have set the bandwidth to the value of 0.5.

APPENDIX 2

Welfare and Poverty Indices Under Various Targeting/Transfer Schemes

Sample: All households in URBAN areas
Poverty line: 25th percentile of reported consumption in URBAN areas

(percent difference from value of index using reported consumption)

	Lambda(5)	Lambda(2)	Lambda(1)	P(0) Headcount Index	P(1) Poverty Gap	P(2) Severity Index	U in percent	L in percent
No transfer (no anti-poverty program)				0.2500	0.0718	0.0284		
Targeting based on reported (total) consumption	16.509	2.548	1.541	0.179	0.039	0.012	0.00	0.00
Targeting based on reported expenditures	16.294 -(1.31)	2.453 -(3.74)	1.488 -(3.41)	0.213 (18.66)	0.042 (8.19)	0.012 (3.58)	5.65	8.18
Targeting based on reported income	58.294 -(7.37)	2.297 -(9.85)	1.423 -(7.66)	0.210 (17.42)	0.046 (17.90)	0.014 (18.23)	9.03	10.55
Targeting based on the probability of being poor	14.764 -(10.57)	2.2284 -(12.56)	1.3940 -(9.53)	0.2131 (18.91)	0.0480 (23.94)	0.0151 (28.19)	11.20	10.37
Targeting based on asset index	13.223 -(19.91)	2.004 -(21.38)	1.292 -(16.14)	0.212 (18.42)	0.049 (26.89)	0.016 (34.24)	12.55	21.23
Targeting based on the share of food in total consumption	10.071 -(39.00)	1.653 -(35.13)	1.147 -(25.54)	0.217 (21.21)	0.054 (39.33)	0.019 (59.63)	17.57	27.56

REFERENCES

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

Atkinson, A. (1995). On Targeting Social Security: Theory and Western Experience with Family Benefits. In *Public Spending and the Poor: Theory and Evidence*, ed. van de Walle and Nead. Baltimore, Md., U.S.A.: Johns Hopkins University Press.

Besley, T., and R. Kanbur (1993). The Principles of Targeting. In *Including the Poor*, ed. M. Lipton and J. van der Gaag. Washington, DC: World Bank.

Chaudhuri, S., and M. Ravallion (1994). How Well Do Static Indicators Identify the Chronically Poor? *Journal of Public Economics* 53: 367–394.

Coady, D., and J. Drèze (2000). Commodity Taxation and Social Welfare: The Generalized Ramsey Rule. International Food Policy Research Institute, Washington, DC. Photocopy.

Coady, D., and E. Skoufias (2001). *On the Targeting and Redistributive Efficiencies of Alternative Transfer Instruments*. Food Consumption and Nutrition Division Discussion Paper 100. Washington, DC: International Food Policy Research Institute.

Cornia, G., and F. Stewart (1995). Two Errors of Targeting. In *Public Spending and the Poor: Theory and Evidence*, ed. van de Walle and Nead. Baltimore, Md., U.S.A.: Johns Hopkins University Press.

Deaton, A. (1997). *The Analysis of Household Surveys: Microeconometric Analysis for Development*. Baltimore, Md., U.S.A.: Johns Hopkins University Press.

Deaton, A., and S. Zaidi (1999). *Guidelines for Constructing Consumption Aggregates for Welfare Analysis.* Washington, DC: World Bank.

Dr**P**ze, 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.

Filmer, D., and L. Pritchett (1998). *Estimating Wealth Effects Without Expenditure Data or Tears*.Policy Research Working Paper No. 1994. Washington, DC: World Bank.

Foster, J., J. Greer, and E. Thorbecke (1984). A Class of Decomposable Poverty Measures. *Econometrica* 52: 761–765.

Glewwe, P., and J. van der Gaag (1990). Identifying the Poor in Developing Countries: Do Different Definitions Matter? *World Development* 18 (6): 803–814.

Grosh, M. (1994). *From Platitudes to Practice: Targeting Social Programs in Latin America*. Washington, DC: World Bank.

Hentschel, J., and P. Lanjouw (1996). *Constructing an Indicator of Consumption for the Analysis of Poverty.* Living Standards Measurement Study Working Paper No. 124. Washington, DC: World Bank.

Ravallion, M. (1989). Land-Contingent Poverty Alleviation Schemes. *World Development* 17: 1223–1233.

Ravallion, M. (1992). *Poverty Comparisons: A Guide to Concepts and Methods*. Washington, DC: World Bank.

Schady, N. (1999). Picking the Poor: Indicators for Geographic Targeting in Peru. Photocopy.

Sen, A. (1992). *Inequality Reexamined*. Oxford: Clarendon Press.

Sharma, S. (1994). *Applied Multivariate Techniques*. John Wiley.

Skoufias, E., B. Davis, and J. Behrman (1999). An Evaluation of the Selection of Beneficiary Households in the Education, Health, and Nutrition Program (PROGRESA) of Mexico. Final Report. International Food Policy Research Institute, Washington, DC.

Subramanian, S., and A. Deaton (1996). The Demand for Food and Calories. *Journal of Political Economy* 104: 133-162.

van de Walle, D., and K. Nead, eds. (1995). *Public Spending and the Poor: Theory and Evidence*. Baltimore, Md., U.S.A.: Johns Hopkins University Press.

FCND DISCUSSION PAPERS

FCND DISCUSSION PAPERS

FCND DISCUSSION PAPERS

FCND DISCUSSION PAPERS

69   *Supply Response of West African Agricultural Households: Implications of Intrahousehold Preference Heterogeneity*, Lisa C. Smith and Jean-Paul Chavas, July 1999

70   *Child Health Care Demand in a Developing Country: Unconditional Estimates from the Philippines*, Kelly Hallman, August 1999

71   *Social Capital and Income Generation in South Africa, 1993-98*, John Maluccio, Lawrence Haddad, and Julian May, September 1999

72   *Validity of Rapid Estimates of Household Wealth and Income for Health Surveys in Rural Africa*, Saul S. Morris, Calogero Carletto, John Hoddinott, and Luc J. M. Christiaensen, October 1999

73   *Social Roles, Human Capital, and the Intrahousehold Division of Labor: Evidence from Pakistan*, Marcel Fafchamps and Agnes R. Quisumbing, October 1999

74   *Can Cash Transfer Programs Work in Resource-Poor Countries? The Experience in Mozambique*, Jan W. Low, James L. Garrett, and Vitória Ginja, October 1999

75   *Determinants of Poverty in Egypt, 1997*, Gaurav *Datt* and Dean Jolliffe, October 1999

76   *Raising Primary School Enrolment in Developing Countries: The Relative Importance of Supply and Demand*, Sudhanshu Handa, November 1999

77   *The Political Economy of Food Subsidy Reform in Egypt*, Tammi Gutner, November 1999.

78   *Determinants of Poverty in Mozambique: 1996-97*, Gaurav Datt, Kenneth Simler, Sanjukta Mukherjee, and Gabriel Dava, January 2000

79   *Adult Health in the Time of Drought*, John Hoddinott and Bill Kinsey, January 2000

80   *Nontraditional Crops and Land Accumulation Among Guatemalan Smallholders: Is the Impact Sustainable?* Calogero Carletto, February 2000

81   *The Constraints to Good Child Care Practices in Accra: Implications for Programs*, Margaret Armar-Klemesu, Marie T. Ruel, Daniel G. Maxwell, Carol E. Levin, and Saul S. Morris, February 2000

82   *Pathways of Rural Development in Madagascar: An Empirical Investigation of the Critical Triangle of Environmental Sustainability, Economic Growth, and Poverty Alleviation*, Manfred Zeller, Cécile Lapenu, Bart Minten, Eliane Ralison, Désiré Randrianaivo, and Claude Randrianarisoa, March 2000

83   *Quality or Quantity? The Supply-Side Determinants of Primary Schooling in Rural Mozambique*, Sudhanshu Handa and Kenneth R. Simler, March 2000

84   *Intrahousehold Allocation and Gender Relations: New Empirical Evidence from Four Developing Countries*, Agnes R. Quisumbing and John A. Maluccio, April 2000

85   *Intrahousehold Impact of Transfer of Modern Agricultural Technology: A Gender Perspective*, Ruchira Tabassum Naved, April 2000

86   *Women's Assets and Intrahousehold Allocation in Rural Bangladesh: Testing Measures of Bargaining Power*, Agnes R. Quisumbing and Bénédicte de la Brière, April 2000

87   *Changes in Intrahousehold Labor Allocation to Environmental Goods Collection: A Case Study from Rural Nepal*, Priscilla A. Cooke, May 2000

88   *The Determinants of Employment Status in Egypt*, Ragui Assaad, Fatma El-Hamidi, and Akhter U. Ahmed, June 2000

89   *The Role of the State in Promoting Microfinance Institutions*, Cécile Lapenu, June 2000

90   *Empirical Measurements of Households' Access to Credit and Credit Constraints in Developing Countries: Methodological Issues and Evidence*, Aliou Diagne, Manfred Zeller, and Manohar Sharma, July 2000

91   *Comparing Village Characteristics Derived From Rapid Appraisals and Household Surveys: A Tale From Northern Mali*, Luc Christiaensen, John Hoddinott, and Gilles Bergeron, July 2000

FCND DISCUSSION PAPERS

FCND DISCUSSION PAPERS