INTERNATIONAL Poverty Centre
United Nations Development Programme
UNDP
Working Paper number 30 October, 2006

OPERATIONAL POVERTY TARGETING IN PERU – PROXY MEANS TESTING WITH NON-INCOME INDICATORS

Julia Johannsen
Department of Agricultural Economics and Rural Development, University of Goettingen, Germany

Working Paper

United Nations Development Programme
International Poverty Centre

International Poverty Centre
SBS – Ed. BNDES,10º andar
70076 900 Brasilia DF
Brazil

povertycentre@undp-povertycentre.org
www.undp.org/povertycentre
Telephone +55 61 2105 5000
Fax +55 61 2105 5001

Print ISSN: 1812-108X

OPERATIONAL POVERTY TARGETING IN PERU – PROXY MEANS TESTING WITH NON-INCOME INDICATORS*

Julia Johannsen **

ABSTRACT

The measurement of per capita daily expenditures relative to a monetary poverty line, also known as 'sophisticated means testing', is the most widely used approach to poverty assessment. However, it is reliant on the implementation of time- and cost-intensive household surveys. Hence, for operational purposes, it is not an effective method for targeting poor households with development services. This paper shows how to identify an alternative poverty assessment tool for Peru. The tool consists of a maximum of 15 indicators that are powerful predictors of per capita household expenditures. The indicators were selected out of a wide range of indicators used to gauge different poverty dimensions. The resultant poverty classification of households is based on the 'percent point function' of the predicted expenditures and validated by various accuracy measures and their confidence intervals. The results reveal that the 15 indicators correctly identify over 81 per cent of poor households when the national poverty line is employed as the benchmark. Thus, this tool might be considered, under certain conditions, as an alternative to the collection of detailed expenditure data. It offers an operational instrument for fairly accurate ex-ante poverty targeting and ex-post impact assessments.

Keywords: Poverty targeting, targeting accuracy, expenditure predictions, percent point function, Latin America, Peru.

---

* I gratefully acknowledge the conceptual and technical contributions of Manfred Zeller, Nanak Kakwani, Serguei Soares, Fabio Soares, and Walter Zucchini. My thanks are due to Stephan Klasen, Luis Castillo, and Pedro Llontop for valuable comments on earlier versions of the paper. Furthermore, this paper benefited from insights gained during a collaboration with the IRIS Center, University of Maryland, within the scope of the research project "Development of Poverty Assessment Tools," funded by the U.S. Agency for International Development (USAID) under the Accelerated Microenterprise Advancement Project (AMAP). The author was a resident scholar in the International Poverty Centre in 2005.

** Contact information: Julia Johannsen, Department of Agricultural Economics and Rural Development, University of Goettingen, Waldweg 26, D-37073 Goettingen, Germany.

Email: jjohann@gwdg.de, Fax: +49-551-393076, Phone: +49-551-392115, Cel.: +49-176-51373660.

1 INTRODUCTION

Both our understanding of poverty and the measurement approaches have considerably improved the targeting efforts of diverse intervention types during the last decades. The operational assessment of the money-metric dimension of poverty in the field, however, has still not reached a point where it could provide generally accepted blueprint solutions.[1] Instead, "(…) numerous approaches have been explored in the recent literature, and no 'best practice' approach has yet emerged" (Montgomery *et al.*, 2000: 155-156).[2] This situation reflects both the urgent need for operational poverty assessment and the difficulties involved.

The principles, (cost-related) potential benefits as well as ethical and political problems of targeting versus broad or universal services have been extensively discussed in the literature (e.g., van de Walle, 1998; Besley & Kanbur, 1993; Sen, 1995). There is also a growing amount of experience available respecting the suitability of different targeting forms (individual or household assessment, categorical or group targeting, self-selection, and their combinations) for different intervention types (cash transfers, social funds, food subsidies, etc.) in various contexts ((e.g., Skoufias *et al.* (2001) and Grosh (1994) for Latin America; Grootaert & Braithwaite (1998) for Eastern Europe and the former Soviet Union)).

It shows that (conditional) cash-transfer programs, in particular, and microfinance institutions, by nature, rely on targeting mechanisms at the individual or household levels, often preceded by some categorical targeting towards geographic or demographic sub-groups (cf., Coady *et al.*, 2003, 2004). Proxy means testing or "indicator targeting," which belongs to the approaches focusing on the household level, identifies potential beneficiary households by a single or, as in this paper, a set of a few weighted indicators that are highly correlated with (low) expenditures and, at the same time, can be obtained more easily and less costly than expenditures themselves (Besley & Kanbur, 1993).

The use of proxy indicators for poverty assessment is usually motivated by one of the following situations. Either no information on income or expenditures is available in an already existing data base as, for example, in demographic research based on DHS surveys (Filmer & Prichett, 1999, 2001; Sahn & Stifel, 2000; Montgomery *et al.*, 2000) or the existing income or expenditure data is not available in the form required for the intended evaluation and implies considerable methodological challenges (cf. Soares *et al.*, 2006). Or the purpose of the research requires the collection of new data anyway, but may not necessarily demand detailed expenditure surveys as, for example, in interdisciplinary research on the poverty impact of agricultural and ecological programs. Or finally, development practitioners want to conduct outreach or impact assessments in a situation where the scope or budget of the respective intervention does not permit an extensive collection of detailed income or expenditure data.

All of these situations call for operational assessment techniques of absolute poverty[3] that can be used not only in large-scale national programs but also in medium or even small-scale interventions, and that are motivated by the interest in "reducing our dependence on extremely expensive, time-consuming, and most likely, inaccurate consumption surveys. One or two questions are a good deal cheaper to ask than two or three hundred!" (Deaton, 2001: 144).

The old concern among donors, governments and practitioners about their success in reaching the poor has been re-enforced by the time-related urgency for effective action in the context of the Millennium Development Goals. And it has in some cases provoked consideration of targeting goals in national legislation, as shown by the Microenterprise for Self-Reliance Act passed by the US Congress in 2000 and amended in 2003, which requires that all microfinance institutions that receive funding from the U.S. Agency for International Development report the share of resources allocated to the very poor and the absolute number of very poor among their clients.

The objective of this paper is to contribute to the development of time and cost-saving instruments for poverty targeting and impact assessments that help to determine whether development interventions meet their poverty outreach and alleviation targets.

We propose a tool that makes it possible to estimate household expenditures as accurately as possible by means of a set of proxy indicators, an exercise that has been performed in the relevant literature using different approaches for the selection of indicators, determination of weights and classification of households (e.g., Ahmed & Bouis, 2002; Grosh & Baker, 1995; Wodon, 1997). The approach pursued here departs from existing traditions in two ways. The former refers to the prediction of household expenditures as such and the subsequent poverty classification of households in the sample, and the latter to the validation and robustness tests of the proposed tools with regard to the transferability of the accuracy performance to further applications in the country.

Firstly, although based on theoretical and logical considerations concerning expenditure predictions, we avoid an arbitrary indicator selection and the application of external weights, both common in most of the asset and housing indices currently used (e.g., Gibbons & DeWit, 1998; Deutsch & Silber, 2005). Furthermore, the poverty classification of households is then based on the percent point function of expenditures, an approach that considerably increases the achieved targeting accuracies and has not yet been used, to the best of our knowledge, for this purpose. Secondly, the tools are validated by diverse accuracy measures (including the conventional undercoverage and leakage measures as well as further combined criteria) and their respective confidence intervals, which are seldom reported, but provide important insights on the robustness of accuracy performance. Out-of-sample tests complete the validation of the indicator sets and confirm the promising results for accurate and still operational poverty assessments compared to conventional income and expenditure surveys.

2 IDENTIFYING AND VALIDATING POTENTIAL POVERTY INDICATORS

THE SELECTION OF PROXY INDICATORS OF POVERTY

Poverty assessment tools for a targeting or impact evaluation in a given country, as defined here, consist of suitable sets of five to 15 of such indicators characterized by a high predictive power of per capita daily expenditures, the relevant benchmark measure used in this study following the concept underlying the official Peruvian poverty statistics (Webb & Fernández 2003; Herrera, 2002).[4]

Depending on the multidimensionality of the indicator set, poverty assessments based on proxy variables have the additional advantages of facilitating the consideration of both various dimensions of capability deprivation (cf. Alkire, 2002) without further data collection as well as of avoiding unintended incentive effects, such as labor supply responses to targeting (cf. Kanbur *et al.*, 1995). Before making use of these advantages, of course, the suitable indicator set has to be identified and the respective survey implemented in practice. For tool identification, we use the most recent living standard measurement survey (LSMS) for Peru from the year 2000,[5] which, apart from the objective benchmark in terms of household expenditures, contains rich information for potential indicators belonging to the following "poverty dimensions":

- *demographics* (age, marital status, household size); ethnic and religious affiliation;
- *illness and disability*;
- *socioeconomic status* (education, occupation);
- *assets* (land, animals, farm assets, household durables);
- *housing* (ownership status, size, type of material, amenities);
- *access to communication* (internet, telecommunications);
- *credit and financial assets* (financial accounts);
- *selected single expenditure items* (clothing, remittances).

Tool identification is based on model (1), which regresses the logarithm of per capita daily expenditures (y_i) of household i on a set of variables (x_n) in order to identify the best five, 10 and 15 poverty indicators, i.e., the strongest correlates with the log of expenditures:

$$\ln(y_i) = \beta X + \mu_i \qquad (1)$$

where X describes a vector of household characteristics $\{x_1, x_2, ..., x_{142}\}$, and μ_i is the error term that describes the unobserved expenditure components (such as personal abilities) that affect the household's expenditure opportunities (idiosyncratic error), as well as the noise due to misspecifications of the empirical model (model error).

A detailed exploration of the survey data allows the construction of 142 potential predictors.[6] The initial regression model presented here is geared to the reduced form of the welfare model by Glewwe (1991). It leads us to the consideration of the different "poverty dimensions" listed above, to which the *selected single expenditure items* are added due to their assumed high potential for the prediction of total household expenditures (cf. Zeller *et al.*, 2006).

As a household's *human capital,* as well as productivity constraints caused by illness, are determined by the number and composition of its members, we introduce demographic variables, the most important of which are included in the control variables. Education itself is introduced in the form of a broad range of ordinal and binary variables for the different sex and age groups in the household. We consider the *specific human capital* in terms of personal knowledge and income-generating capacities by various occupation dummies, aware that they additionally reflect exogenous *labor market* responses to human capital. In these categories, indicators related to the head of the household are calculated separately from those of the remaining household members. Male and female household members are treated separately as well.

Variables on the ownership, number and value of the household durables, farm assets and/or housing characteristics represent the *physical capital* of a household and are widely used in Asian microfinance institutions' targeting instruments and as welfare proxies in various socioeconomic studies (cf., e.g., Sahn & Stifel, 2000; Filmer & Prichett, 2001; Deutsch & Silber, 2005; Gibbons & DeWit, 1998).

Given the large number of 142 initial regressors (x_n), we use the maximizing-R-squared-regression technique (MaxR) that identifies sets of consecutively increasing numbers of indicators while maximizing the explained variance R^2 in every step, rather than applying an arbitrary indicator selection within each dimension.[7]

The only restriction imposed is that, in all iterations of the MaxR procedure, we force nine control variables into the model (see Table 3). They ensure that the estimated coefficients are controlled for regional agro-ecological, cultural and socioeconomic differences, as well as for some demographic factors known as powerful factors, influencing household expenditures (cf., e.g., Ravallion, 1992; Glewwe, 1991). This way, three tools are obtained consisting of the best five, 10 and 15 indicators (not counting the control variables) and their objective weights that result from the regression coefficients.

THE POVERTY CLASSIFICATION OF HOUSEHOLDS

The identified indicator sets are tested for their accuracy in predicting the poverty status of the households. As the standard benchmark of reliable accuracy we chose i) the national poverty line of Peru z_n (hereafter referred to as identifying the 'poor') and ii) the corresponding expenditure cut-off of the bottom 50% of the population below this line, i.e., the median poverty line z_m, as an even stricter definition of poverty (hereafter referred to as identifying the 'very poor').

Due to geographic diversity, the national poverty line, as well as our alternative median poverty line, are disaggregated into seven regional ones. Both lines are listed in Table 1 with the corresponding poverty headcounts for each geographical domain.

TABLE 1

Comparison between the national poverty line A) and the median poverty line B) with the corresponding poverty headcounts for the seven regions in Peru

Expenditures May 2000	A) Daily national poverty line	Poverty headcount	B) Daily expenditures equivalent to 50% < national poverty line	Poverty headcount
Region	(Soles/ pers./ day)	(percent)*	(Soles/ pers./ day)	(percent)*
Metropolitan Lima	7.75	45.2%	5.48	22.6%
Urban Coast	6.41	53.1%	4.29	26.6%
Rural Coast	4.35	64.4%	2.78	32.2%
Urban Highland	5.51	44.3%	3.70	22.2%
Rural Highland	3.61	65.5%	2.18	32.8%
Urban Lowland	5.32	51.5%	3.51	25.8%
Rural Lowland	3.71	69.2%	2.39	34.6%
Total poor (national aggregate of headcounts)		54.1%		27.1%

Source: Author's calculations based on ENNIV data of 2000.

* The poverty headcount corresponds to the official figures based on ENNIV data of 2000, first published by Webb & Fernández (2000).

In order to test the resulting tools for their poverty accuracy, the predicted household expenditures are transferred into a binary variable that classifies each single household as either (very) poor or non-poor.

In contrast to previous work by Zeller et al. (2005), in which the poverty rates are calculated by comparing the predicted household expenditures $\hat{\beta} x_i$ in equation (1) directly with the poverty line, we opt for an approach that indirectly takes the unknown error term μ into account. By doing this, we consider that the residuals might contain additional information on immeasurable poverty determinants and avoid biased estimates of poverty rates (cf. also Hentschel et al., 2000; Ravallion, 1998). (In our case, the simple approach based on $\hat{\beta} x_i$ compared with the corresponding poverty line indeed results in a considerable underestimation of the predicted poverty rates, particularly when employing the stricter median poverty line.)

We derive the poor/non-poor classification from "percentile corrected" prediction values based on the empirical cumulative distribution or percent point function of the log of the observed, i.e., "true" daily household expenditures $\ln(y_i)$. This approach is geared to the procedures proposed by the poverty mapping literature, in particular, the approach by Hentschel *et al.* (2000) and Kakwani (2006), who use the cumulative standard normal distribution function of expenditures from which they derive the probability that a household is poor. In order to circumvent the problems related to transforming these probabilities into a binary poverty status of each household and to account for the non-normality of the expenditure distribution, we opt for the percent point function that makes it directly possible to establish a poor/non-poor classification based on the actual poverty headcount as a cut-off point. A similar idea is subject to the approach proposed by Ahmed & Bouis (2002), who, in contrast, use a flexible expenditure cut-off to force the poor/non-poor classification to minimize exclusion errors.

Let Fr be the empirical cumulative distribution function of the observed expenditures $\ln(y_i)$, and let Fp be the empirical cumulative distribution function of the predicted expenditures $\ln(\hat{y}_i)$. The "percentile corrected" predicted expenditures $\ln(\hat{y}_{ci})$ are defined as:

$$\ln(\hat{y}_{c_i}) = F_r^{-1}(F_p(\ln(\hat{y}_i))) \quad (2)$$

These corrected expenditures are compared to the corresponding poverty line z, below which a household is defined as (very) poor.

Alternatively, the poverty line itself can be "percentile corrected" in order to be directly applicable to the empirical cumulative distribution function of the predicted expenditures F_p. This provides the possibility of expenditure predictions using the poverty assessment tool in independent new samples without the need for information on observed expenditures, which constitutes the main purpose of an operational indicator-based poverty assessment tool. The percentile-corrected poverty line z* is defined as:

$$z^* = F_p^{-1}(F_r(z)) \quad (3)$$

It is derived from a comparison of the percent point function of the observed expenditures F_r to the true poverty headcount and is defined as the value of observed expenditures that corresponds to the household closest to the poverty headcount, as illustrated in Table 2.

TABLE 2

Illustration of calculating a "percentile corrected" poverty line based on the empirical cumulative distribution function of observed daily household expenditures (fictive data)

Household no.	Ranked observed expenditures per capita in Soles (y*)	Cumulative weight of household	Cumulative expenditure distribution (Y*)	Cumulative expenditure distribution (Y*)	Poverty classification
1	0.81	5,044	5,044/ 710,655	0.01	Poor
2	0.93	18,565	18,565/ 710,655	0.03	Poor
…	…	…	…	…	Poor
x	**0.98** = percentile-corrected poverty line **z***	191,877	191,877/ 710,655	**0.27** assuming that this matches observed poverty headcount	Non-poor
…	…	…	…	…	Non-poor
3,977	2.26	710,655	710,655/ 710,655	1	Non-poor

TESTING THE TOOLS FOR THEIR ACCURACY

The following accuracy measures and prediction errors are potentially relevant when validating the tools (for details and discussion, cf. IRIS, 2005; Hoddinott, 1999; Cornia & Stewart, 1995):

- *overall accuracy*: sum of correctly predicted poor and non-poor as a proportion of all;
- *poverty accuracy*: sum of correctly predicted poor as a proportion of the total poor;
- *undercoverage* (exclusion error): sum of actual poor wrongly classified as non-poor as a proportion of the total poor; and
- *leakage* (inclusion error): sum of actual non-poor wrongly classified as poor as a proportion of the total poor.

On the assumption that a budget-constrained policy maker is interested in both correctly targeting the (very) poor by identifying the households individually and in reaching a target population similar in size to the actual poverty headcount, the IRIS Center proposes an alternative accuracy criterion:

- *Balanced Poverty Accuracy Criterion (BPAC), defined as the poverty accuracy minus the absolute difference between undercoverage and leakage, all of them as given above (IRIS, 2005).*

We base our tool validation on BPAC as a summary accuracy measure. (Note, however, that the BPAC measure would still allow very high undercoverage and leakage figures without reducing the accuracy, provided that undercoverage and leakage errors are equal in size and cancel each other out. To correct for this, we also add a slightly different indicator called *Focused Poverty Accuracy Criterion* (FPAC) defined as poverty accuracy minus leakage. It directly deteriorates in case of any misclassification error, thus neglecting the policy objective of targeting a population similar in size to the "true" poor population share. Our results show that the ranking of tools with the FPAC measure is consistent with the tools found best for maximizing the BPAC measure.)

CONFIDENCE INTERVALS AND OUT-OF-SAMPLE TEST

Tool identification is undertaken with two-thirds of the original LSMS sample, i.e., 2,611 randomly drawn households out of 3,977. For each of the resulting accuracy criteria, a 95% confidence interval is calculated to test the reliability of the sample results and accuracy values. The confidence intervals are derived from a bootstrapping procedure based on 1,000 resampled datasets of the same size.

In order to test the robustness of the expenditure predictions achieved by the identified indicator sets, we conduct an independent out-of-sample test with the remaining one-third of the original sample. The test consists of the projection of expenditures by means of the corresponding indicators with their respective parameters, resulting from the in-sample regression analysis. The coefficients are introduced in the out-of-sample data, and all of the corresponding accuracy measures based on the percentile-corrected poverty line are calculated as usual.

3 FINDINGS

The methodology implies that the sets of best regressors are statistically determined by the search for the best model fit, as discussed in section two. The term 'best' indicator set should, therefore, not be misunderstood as being best in terms of any of the accuracy measures. We show the different tests for accuracy in the second part of this section after presenting the resulting tools consisting of five, 10 and 15 indicators.

THE RESULTING TOOLS

Table 3 shows the three poverty assessment tools with their respective parameter estimates, all of which are highly significant at $p < 0.001$. The goodness of fit of the three tools in the form of the adjusted R^2 value increases with the number of indicators and ranges from 0.722 in the first tool to 0.754 in the third one. As this study is not concerned with the causal determinants of expenditure poverty, we do not worry about endogeneity, nor comment on the magnitude of the regression coefficients as long as their direction of influence conforms to theory or logical considerations.

The main finding refers to the multidimensional character of all of the three tools. They consist of a balanced composition of variables representing the dimensions of selected expenditures, education, household durables and communication, in addition to housing characteristics in two of the three tools.

Not surprisingly, all of the three indicator sets include the variable of annual clothing expenditures, representing one of the sub-categories of total expenditures among the first indicators chosen. Furthermore, monetary asset indicators are preferred to the simple ownership or number of durables. Apart from the single asset indicator on video tapes, the summary indicator on the value of all household durables repeatedly appears among the most accurate proxies.

TABLE 3

The three best indicator sets and their parameter estimates on predicted daily household expenditures per capita (including nine control variables in each tool)

Best 5, Tool 1		Best 10, Tool 2		Best 15, Tool 3	
Adjusted R^2	0.722	Adjusted R^2	0.745	Adjusted R^2	0.754
Intercept	1.655***	Intercept	1.822***	Intercept	2.256***
Household size	-0.241***	Household size	-0.235***	Household size	-0.291***
Household size squared	0.010***	Household size squared	0.009***	Household size squared	0.011***
Age of head of hh	0.002***	Age of head of hh	0.002***	Age of head of hh	0.001**
Household lives in Urban coast	-0.097***	Household lives in Urban coast	-0.090***	Household lives in Urban coast	-0.092***
Household lives in Rural coast	-0.354***	Household lives in Rural coast	-0.322***	Household lives in Rural coast	-0.292***
Household lives in Urban highlands	-0.214***	Household lives in Urban highlands	-0.215***	Household lives in Urban highlands	-0.207***
Household lives in Rural highlands	-0.477***	Household lives in Rural highlands	-0.418***	Household lives in Rural highlands	-0.393***
Household lives in Urban lowlands	-0.179***	Household lives in Urban lowlands	-0.169***	Household lives in Urban lowlands	-0.157***
Household lives in Rural lowlands	-0.358***	Household lives in Rural lowlands	-0.365***	Household lives in Rural lowlands	-0.337***
Log of value of video	0.041***	Log of value of video	0.030***	Log of value of video	0.024***
Log of annual clothing exp p.c.	0.032***	Log of annual clothing exp p.c.	0.029***	Log of annual clothing exp p.c.	0.030***
Log of value of durables	0.123***	Log of value of durables	0.106***	Log of value of durables	0.094***
Household has fixed telephone	0.264***	Household has fixed telephone	0.234***	Household has fixed telephone	0.222***
Average years of education of all members	0.035***	Average years of education of all members	0.030***		
		Floor material: dirt/ other	-0.104***	Floor material: dirt/ other	-0.098***
		Log of remittances sent	0.024***	Log of remittances sent	0.024***
		Household owns cell phones	0.237***	Household owns cell phones	0.205***
		Number of members using internet	0.090***	Number of members using internet	0.085***
		Household uses no or inferior cooking fuel	-0.284***	Household uses no or inferior cooking fuel	-0.365***
				Household uses wood/ charcoal as cooking fuel	-0.122***
				Log of value of vacuum cleaners	0.027***
				Light source: candles	-0.162***
				Number of members that can read and write	0.042***
				Number of members with sup./ univ./post-grad. Educ.	0.050***
Level of statistical significance: ***$p < 0.001$.				Number of shovels/ rakes owned	0.014***

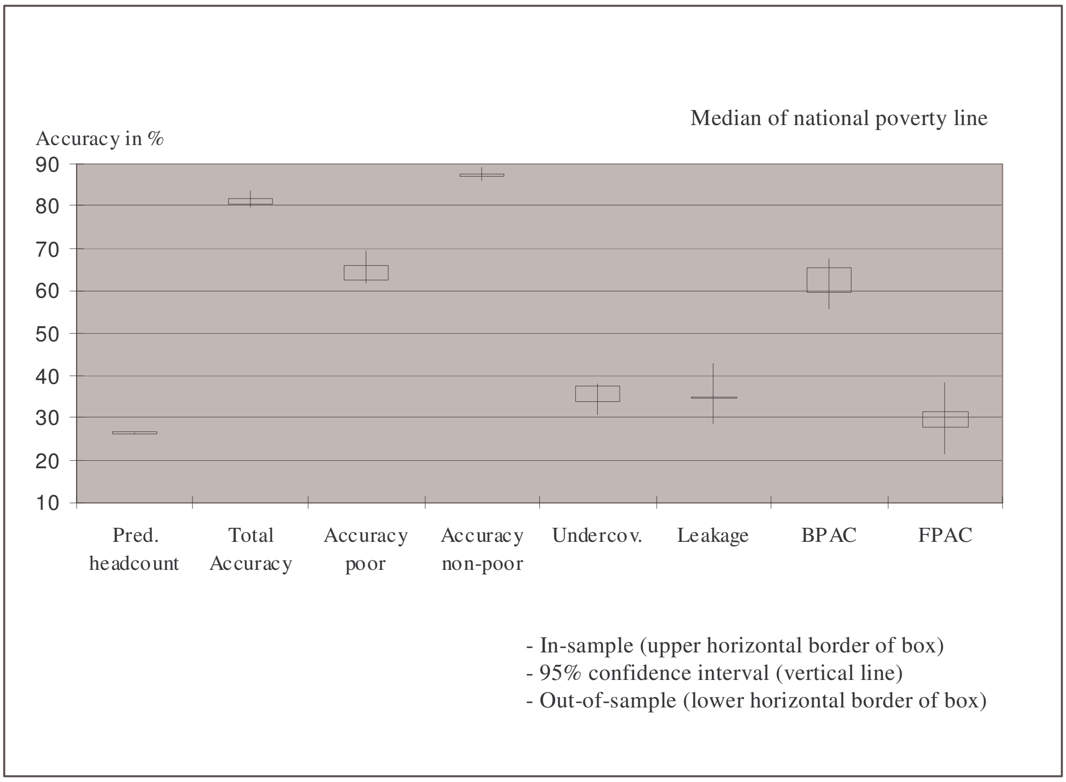
An operational poverty assessment tool should, however, not only be as accurate as possible in predicting the poverty status of the population, but also – in order to be suitable for implementation – be practical in addition, i.e., indicators should be easy to ask, to respond, and to verify under field conditions. From a viewpoint of practicability, it might be worth banning the variable on the value of all household durables from the indicator list because it demands an extended questionnaire section about the number and value of all household durables. In addition, this indicator (as well as all remaining monetary variables) is difficult to verify, a fact that is normally taken seriously in practical poverty assessment. Grosh & Baker (1995) show for the case of Jamaica, however, that the impact even of an exaggerated distortion caused, among others, by the underreporting of asset ownership of 25 percent of the surveyed households still produces almost identical targeting outcomes to those based on the supposedly undistorted information. We will discuss the implication of a practicability adjustment that excludes certain indicators in the following section when analyzing the accuracy performance of the tools.

ACCURACY PERFORMANCE OF THE DIFFERENT TOOLS

As expected, the best 15 set achieves the highest accuracy values and lowest misclassification errors among the three tools although the decrease in the accuracy of the smaller indicator sets is negligible, indicating that the best five set already achieves satisfactory poverty predictions. The obtained tool of 15 indicators is evaluated as depicted in Figure 1 and Figure 2 under the two scenarios determined by the two poverty lines (as in Table 1). For the detailed results of the best five to best 15 indicator sets, refer to Table 4.

FIGURE 1

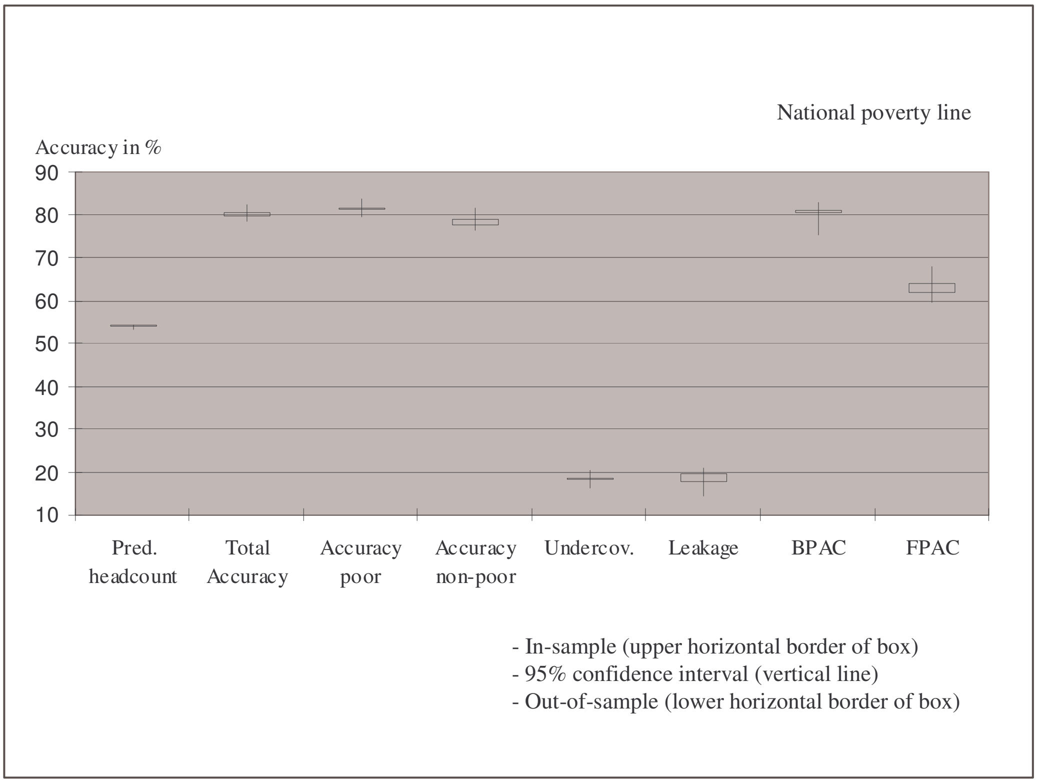
Evaluation of the best 15 indicators through different accuracy measures under the scenario of the median poverty line



The in- and out-of-sample accuracy values are depicted by the upper and lower horizontal border of a box or appear as a single horizontal line in the case of very similar values achieved out-of-sample. The in-sample BPAC value of 81.14% under the national poverty line is achieved with high accuracy by correctly identifying 81.76% of the poor, reduced by a small difference between the inclusion and exclusion errors, i.e., leakage and undercoverage. The corresponding confidence interval [75.44; 82.72], illustrated in the form of a vertical line, indicates that the in-sample value is closer to the upper bound of the interval, but the equally high out-of-sample BPAC value of 80.50% confirms the robustness of this in-sample estimate.

FIGURE 2

Evaluation of the best 15 indicators through different accuracy measures under the scenario of the national poverty line



In the case of the strict median poverty line, the accuracy among the poor is in particular notably lower, thus resulting in higher misclassification errors and a generally lower BPAC with a considerably wider confidence interval [55.85; 67.39] than when the national poverty line is employed. This observation is due to the fact that the tool's ability to correctly identify the poor increases with the percentage of poor in the sample, which is the case when using the higher poverty line that identifies more people as poor. However, particularly in the case of the stricter median line, the percent point function approach for the poverty classification of households proposed here performs much better in terms of a BPAC confidence interval of approximately 56 to 67% than a classification based on the direct expenditure calculation $\hat{\beta} x_i$ would (yielding a BPAC interval for all of the three tools with a much wider span and below 50% for the data used).

TABLE 4

Evaluation of the best five to best 15 indicators through different accuracy measures using two different poverty lines, including 95% confidence intervals and out-of-sample tests

Poverty line used	Robustness tests	Type	Actual poverty headcount	Predicted poverty headcount	Total accuracy	Poverty accuracy	Non-poverty accuracy	Under coverage	Leakage	Balanced poverty accuracy BPAC	Focused poverty accuracy FPAC
Median of national	In-sample	Best five	26.24	26.27	81.02	63.88	87.12	36.12	36.21	63.79	27.67
	95% confidence interval			26.14 26.60	79.03 82.88	60.17 68.13	85.58 88.59	31.87 39.83	29.88 44.36	54.28 65.83	18.75 35.85
	Out-of-sample		27.08	26.27	80.15	61.85	86.95	38.15	35.14	58.85	26.72
Median of national	In-sample	Best 10	26.24	26.45	81.56	65.26	87.36	34.74	35.53	64.47	29.73
	95% confidence interval			26.16 26.61	79.66 83.53	61.31 69.08	85.85 89.09	30.92 38.69	28.66 43.35	55.70 67.04	20.74 38.33
	Out-of-sample		27.08	26.32	81.05	63.62	87.53	36.38	33.58	60.82	30.04
Median of national	In-sample	Best 15	26.24	26.44	82.02	66.11	87.67	33.89	34.64	65.36	31.47
	95% confidence interval			26.13 26.60	79.91 83.72	61.76 69.40	85.98 89.13	30.60 38.24	28.44 42.89	55.85 67.39	21.42 38.41
	Out-of-sample		27.08	26.36	80.38	62.43	87.04	37.57	34.89	59.75	27.55
National	In-sample	Best five	54.27	53.90	79.58	80.84	78.07	19.16	18.48	80.16	62.36
	95% confidence interval			53.42 54.29	77.70 81.82	78.69 83.15	75.67 81.03	16.85 21.32	15.02 22.12	74.86 82.28	58.44 66.69
	Out-of-sample		53.76	54.11	79.65	81.40	77.62	18.60	19.25	80.75	62.15
National	In-sample	Best 10	54.27	53.83	80.09	81.25	78.71	18.75	17.95	80.45	63.30
	95% confidence interval			53.41 54.28	77.87 81.93	78.81 83.37	75.88 81.13	16.63 21.20	14.96 21.65	74.80 82.25	58.87 67.01
	Out-of-sample		53.76	54.11	79.05	80.85	76.97	19.15	19.81	80.19	61.04
National	In-sample	Best 15	54.27	53.93	80.54	81.76	79.10	18.24	17.62	81.14	64.14
	95% confidence interval			53.43 54.30	78.42 82.28	79.53 83.80	76.37 81.71	16.20 20.47	14.38 21.16	75.44 82.72	59.62 67.78
	Out-of-sample		53.76	54.21	79.49	81.34	77.34	18.66	19.50	80.50	61.85

In the preceding section, the possible banning of summary expenditure indicators has been considered. Of course, an exclusion of such powerful monetary proxies reduces the accuracy of a tool. In our case, the replacement of all of the monetary variables (including the summary value of all household durables) by other (next-) best indicators reduces the BPAC value in the in-sample from 63.79 to 58.46 for the best five and from 65.36 to 62.53 for the best 15 set, both under the scenario of the median poverty line. This implies the logical trade-off between the predictive power of a tool and its practicability.

Regardless of these considerations, the achieved accuracy levels (especially in the case of the national poverty line) are high and do not differ much between the tools, for which any of the indicator sets could be proposed to a policy-maker depending on the budgetary constraints and accuracy requirements of the intended assessment.

4 CONCLUSIONS

The paper presents a methodology for identifying an operational poverty assessment tool for Peru and shows how to make concrete statements on its performance based on different accuracy measures. The innovative issues include:

Firstly, with respect to the prediction of household expenditures, an arbitrary indicator selection and the application of external weights are avoided, both common in many of the asset and housing indices currently used. Secondly, the subsequent poverty classification of households is grounded on a transformation of the poverty line based on the percent point function of expenditures, an approach that considerably increases the achieved targeting accuracies and has not yet been used for this purpose to our knowledge. Thirdly, regarding the validation and robustness tests, the resulting tools are validated by diverse accuracy measures (including the conventional undercoverage and leakage measures as well as further combined criteria) and their respective confidence intervals, which provide important insights on the robustness of the accuracy performance. Out-of-sample tests complete the validation of the indicator sets and confirm the promising results for accurate and still operational poverty assessments.

In the case of employing the national poverty line as the relevant benchmark, the proposed tool consisting of 15 indicators achieves an accuracy of correctly predicting the poverty status of 80 to 84% of the poor and a Balanced Poverty Accuracy of 75 to 83% (according to the 95% confidence interval of the poverty accuracy and BPAC, respectively). In the case of the median poverty line, this tool still correctly identifies 62 to 69% of the poor and a BPAC of 56 to 67%, (according to the 95% confidence interval of the poverty accuracy and BPAC, respectively). It makes it possible to both i) identify *ex ante* those households that lie below a certain pre-defined minimum threshold and should, therefore, be offered to participate in a development project or program and ii) assess *ex post* the impact of such interventions on the households' current expenditure level. In order to be employed practically, the indicators should be transformed into a short, focused poverty questionnaire. Compared to a detailed expenditure questionnaire, a tool of five to 15 indicators represents a short and low-cost option for poverty assessments as compared to conventional income and expenditure surveys.

REFERENCES

Ahmed, A. U., & Bouis, H. E. (2002): Weighing what's practical: proxy means tests for targeting food subsidies in Egypt. *Food Policy* 27, 519-540.

Alkire, S. (2002): *Valuing Freedoms. Sen's Capability Approach and Poverty Reduction*. Oxford University Press. New York.

Besley, T., & Kanbur, R. (1993): The Principles of Targeting, pp. 67-90. In M. Lipton, and J. van der Gaag (Eds): *Including the poor - Proceedings of a symposium organized by the World Bank and the International Food Policy Research Institute*, World Bank, Washington, DC.

Coady, D., Grosh, M. E., & Hoddinott, J. (2003): Targeted anti-poverty interventions: A selected annotated bibliography, World Bank, Washington, DC.

Coady, D., Grosh, M. E., & Hoddinott, J. (2004): Targeting outcomes redux. *World Bank Research Observer* 19, 61-85.

Cornia, G. A., & Stewart, F. (1995): Two errors of targeting, pp. 350-386. In D. van de Walle, & K. Nead (Eds): *Public spending and the poor - Theory and evidence*, World Bank, Washington, DC.

Deaton, A. (1997): *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*. Johns Hopkins University Press. Baltimore, United States.

Deaton, A. (2001): Counting the World's Poor: Problems and Possible Solutions. *World Bank Research Observer* 16, 125-148.

Deutsch, J., & Silber, J. (2005): The order of acquisition of durable goods and the multidimensional measurement of poverty. Paper presented at the International Conference "The many dimensions of poverty", organized by the Department for International Development (DFID), Inter-American Development Bank (IDB), International Poverty Centre (IPC), United Nations Development Programme (UNDP), and Institute for Applied Economic Research (IPEA), August 29-31, 2005, Brasilia.

Filmer, D., & Prichett, L. (1999): The Effect of Household Wealth on Educational Attainment: Evidence from 35 Countries. *Population and Development Review* 25, 85-120.

Filmer, D., & Prichett, L. (2001): Estimating wealth effects without expenditure data – or tears: An application to educational enrollments in states of India. *Demography* 38, 115-132.

Gibbons, D., & De Wit, J. (1998): Targeting the poorest and covering costs, Countdown 2005 Newsletter of Microcredit Summit Campaign 1(3), Microcredit Summit.

Glewwe, P. (1991): Investigating the determinants of household welfare in Côte d'Ivoire. Journal of *Development Economics* 35, 307-337.

Glewwe, P., & van der Gaag, J. (1990): Identifying the poor in developing countries: Do different definitions matter? *World Development* 18, 803-814.

Grootaert, C., & Braithwaite, J. (1998): *Poverty correlates and indicator-based targeting in eastern Europe and the Former Soviet Union*. World Bank. Washington, DC.

Grosh, M. E. (1994): *Administering targeted social programs in Latin America - From platitude to practice*. World Bank. Washington, DC.

Grosh, M. E., & Baker, J. L. (1995): Proxy means tests for targeting social programs, LSMS working paper No. 118, World Bank, Washington, DC.

Hentschel, J., Lanjouw, J. O., Lanjouw, P., & Poggi, J. (2000): Combining census and survey data to trace the spatial dimensions of poverty: A case study of Ecuador. *World Bank Economic Review* 14, 147-165.

Herrera, J. (2002): La pobreza en el Perú en 2001. Una visión departamental, Instituto Naional de Estatística e Información (INEI), Institut de recherche pour le développment (IRD), Lima, Peru.

Hoddinott, J. (1999): Targeting: Principles and Practice, International Food Policy Research Institute, Washington, DC.

IRIS (2005): Note on Assessment and Improvement of Tool Accuracy, IRIS Center, University of Maryland, College Park.

Kakwani, N. (2006): Poverty Analysis for Policy Use: Poverty Profiles and Mapping. In United Nations Statistics Division (Ed): *United Nations Handbook on Poverty Statistics: Concepts, Methods and Policy Use* (Chapter 7). United Nations, New York, forthcoming.

Kanbur, R., Keen, M., & Tuomala, M. (1995): Labor supply and targeting in poverty-alleviating programs, pp. 91-113. In D. van de Walle, and K. Nead (Eds): *Public spending and the poor - Theory and evidence*, World Bank, Washington, DC.

Medeiros, M. (2006): Poverty, inequality and redistribution: A methodology to define the rich, Working Paper No. 18, International Poverty Centre of the United Nations Development Programme and the Institute for Applied Economic Research, Brasilia.

Montgomery, M., Gragnolati, M., Burke, K. A., & Paredes, E. (2000): Measuring Living Standards with Proxy Variables. *Demography* 37, 155-174.

Nussbaum, M. C. (1995): Human capabilities, female human beings, pp. 61-104. In M. C. Nussbaum, & J. Glover (Eds): *Women, culture and development*, Clarendon Press, Oxford.

Nussbaum, M. C. (2000): *Women and human development: the capabilities approach*. Cambridge University Press. Cambridge.

Ravallion, M. (1992): Poverty comparisons: a guide to concepts and methods, LSMS Working Paper No. 88, World Bank, Washington, D.C.

Ravallion, M. (1998): Expected Poverty under Risk-Induced Welfare Variability. *Economic Journal* 98, 1171-82.

Robeyns, I. (2005): The capability approach: A theoretical survey. *Journal of Human Development* 6, 93-114.

Sahn, D. E., & Stifel, D. C. (2000): Poverty comparison over time and across countries in Africa. *World Development* 28, 2123-2155.

Sen, A. (1995): The political economy of targeting, pp. 11-24. In D. van de Walle, & K. Nead (Eds): *Public spending and the poor - Theory and evidence*, World Bank, Washington, DC.

Sen, A. K. (1985): *Commodities and capabilities*. North-Holland. Amsterdam.

Sen, A. K. (2000): *Development as freedom*. Anchor books. New York.

Skoufias, E., Davis, B., & de la Vega, S. (2001): Targeting the poor in Mexico: An evaluation of the selection of households into PROGRESA. *World Development* 29, 1769-1784.

Soares, F. V., Soares, S., Medeiros, M., & Osório, R. G. (2006): Cash Transfer Programmes in Brazil: Impacts on inequality and poverty, Working Paper No. 21, International Poverty Centre of the United Nations Development Programme and the (IPEA) Institute for Applied Economic Research, Brasilia.

van de Walle, D. (1998): Targeting revisited. *World Bank Research Observer* 13, 231-248.

Webb, R., & Fernández , G. (2003): *Anuario Estadístico - Perú en Números 2003*. Instituto Cuánto. Lima.

Wodon, Q. T. (1997): Targeting the poor using ROC curves. *World Development* 25, 2083-2092.

Zeller, M., Johannsen, J., & Alcaraz, G. (2005): Developing and testing poverty assessment tools: Results from accuracy tests in Peru, IRIS Center, University of Maryland, College Park, Maryland.

Zeller, M., Sharma, M., Henry, C., & Lapenu, C. (2006): An operational method for assessing the poverty outreach performance of development policies and projects: Results of case studies in Africa, Asia, and Latin America. *World Development* 34, 446-464.

NOTES

1. It should be mentioned that the concepts regarding the money-metric dimension of poverty and its alleviation are not limited to the identification of the poor alone. An interesting approach by Medeiros (2006) starts from the question of defining the rich by means of a money-metric affluence line and then analyzes their role and potential contributions to reducing poverty.

2. The development in poverty research implies that since the capability concept of poverty (cf, e.g., Sen, 1985; Nussbaum, 1995, 2000) at the latest, we can no longer measure monetary income or expenditures alone and seriously claim that we are assessing well-being in a comprehensive way. However, in view of the great challenges involved in transferring holistic poverty concepts to practical poverty assessment (Alkire, 2002), money-metric approaches continue to play a vital role in political decision-making on public spending, in general, and targeting, and impact evaluations, in particular. In order to take this dilemma into consideration, we explicitly refer to *expenditure poverty* in this paper, i.e., the partial capability deprivation that is caused by the lack of monetary *means* or, in a slightly different interpretation, the deprivation at the household level respecting food and non-food goods and services used as instrumental economic proxies for selected achieved *functionings*. We critically note that we, as a consequence, neglect to establish whether every individual in the household has and takes advantage of the opportunity to make choices and whether the observed *functionings* are the desired outcomes of these choices (Sen, 2000; Robeyns, 2005).

3. For an established operational assessment tool concerning *relative* poverty (based on expenditure rankings of households within a certain locality), refer to Zeller *et al.* (2006).

4. Apart from the justification based on the sound theoretical basis of expenditure-based welfare measures (Ravallion, 1992; Deaton, 1997) and the simple fact that they constitute the preferred measures by the Peruvian official institutions themselves, Glewwe & van der Gaag (1990) provide empirical support for the suitability of (unadjusted) per capita consumption as a meaningful welfare indicator preferable to many other economic poverty definitions.

5. The survey is called "Encuesta Nacional de Hogares sobre Medición de Niveles de Vida (ENNIV)" and was conducted by the "Instituto Cuánto" (Lima, Peru) in 2000.

6. In fact, an exhaustive exploration of the data allows the calculation of nearly 400 potential predictors, many sub-groups of which measure the same phenomenon. By retaining only the most powerful ones from each sub-group (we call this in-dimension pre-selection by MaxR) and by excluding variables with measurement error and those that are too closely correlated in terms of variance inflation factors above 10 or bivariate correlations above 0.65, the initial number of indicators are reduced to 142. Information on the detailed derivation and pre-selection of indicators as well as the summary statistics of all variables are available on request.

7. Formally, OLS is not the most appropriate technique for predicting poverty as it is meant to minimize the sum of squared residuals with respect to a continuous expenditure variable expressing welfare (not poverty). This is a different analytical problem than that of a logistic poverty model with a binary dependent variable (poor/ non-poor). For a more detailed discussion, refer to Grosh & Baker (1995). As Grootaert & Braithwaite (1998) show for the specific economic context of transition countries, there might indeed be quantitative (rather than qualitative) differences between the indicator weights resulting from a welfare and a poverty regression. Zeller *et al.* (2006) compare the approach based on the MaxR OLS regression to other regression approaches for Peru, including logistic regression. Their findings suggest, however, that the probit technique performs worse than OLS and quantile regressions in LSMS-based models for the purpose of accurately identifying poor households. We, therefore, opt for the MaxR OLS approach, which, besides, proves to be convenient for the large number of discrete and continuous variables at this initial stage in the process of developing poverty assessment tools.

International Poverty Centre
SBS – Ed. BNDES,10º andar
70076 900 Brasilia DF
Brazil

povertycentre@undp-povertycentre.org
www.undp.org/povertycentre
Telephone +55 61 2105 5000
Fax +55 61 2105 5001