A Simple Poverty Scorecard for Brazil

Mark Schreiner

17 March 2010

This document and related tools are at `http://www.microfinance.com/#Brazil`.

Abstract

This study uses Brazil's 2008 National Household Survey to construct an easy-to-use scorecard that estimates the likelihood that a household has income below a given poverty line. The scorecard uses ten simple indicators that field workers can quickly collect and verify. Poverty scores can be computed on paper in the field in five to ten minutes. The scorecard's accuracy and precision are reported for a range of poverty lines. The poverty scorecard is a practical way for pro-poor programs in Brazil to monitor poverty rates, track changes in poverty rates over time, and target services.

Acknowledgements

This paper was funded by Grameen Foundation with a grant from Ford Foundation. Data are from the *Instituto Brasileiro de Geografia e Estatística*. Thanks go to Nigel Biggar, Clara Galeazzi, and Jeff Toohig. The simple poverty scorecard is the same as what Grameen Foundation calls the Progress out of Poverty Index™. The PPI™ is a performance-management tool that Grameen Foundation promotes to help institutions achieve their social objectives more effectively.

Author

Mark Schreiner is the Director of Microfinance Risk Management, L.L.C., 2441 Tracy Avenue, Kansas City, MO 64108–2935, U.S.A., `mark@microfinance.com`. He is also Senior Scholar, Center for Social Development, Washington University in Saint Louis, One Brookings Drive, Saint Louis, MO 63130–4899.

Figure 1: A simple poverty scorecard for Brazil

Entity	Name	ID	Date (DD/MM/YY)	
Participant:			Joined:	
Field agent:			Today:	
Service point:			Household size:	

Indicator	Value	Points	Score
1. How many members does the household have?	A. Five or more	0	
	B. Four	6	
	C. Three	11	
	D. Two	17	
	E. One	20	
2. Do any household members ages 5 to 18 go to private school or private pre-school?	A. No	0	
	B. Yes	5	
	C. No members ages 5 to 18	7	
3. How many years of schooling has the female head/spouse completed?	A. Three or less	0	
	B. Four to eleven	2	
	C. Twelve or more	8	
	D. No female head/spouse	8	
4. How many household members work as employees with a written contract, as civil servants for the government, or in the military?	A. None	0	
	B. One	4	
	C. Two or more	13	
5. In their main occupation, how many household members are managers, administrators, professionals in the arts and sciences, mid-level technicians, or clerks?	A. None	0	
	B. One or more	8	
6. How many rooms does the residence have?	A. One to four	0	
	B. Five	2	
	C. Six	5	
	D. Seven	7	
	E. Eight or more	11	
7. How does the household dispose of sewage?	A. Ditch, other, or no bathroom	0	
	B. Simple hole, or directly into river, lake, or ocean	2	
	C. Septic tank not connected to public sewage/rainwater system	3	
	D. Septic tank connected to public sewage/rainwater system	4	
	E. Direct connection to public sewage/rainwater system	5	
8. Does the household have a refrigerator?	A. No	0	
	B. Yes, with one door	5	
	C. Yes, with two doors	10	
9. Does the household have a washing machine?	A. No	0	
	B. Yes	7	
10. Does the household have a cellular or land-line telephone?	A. None	0	
	B. Cellular but not land-line	5	
	C. Land-line but not cellular	6	
	D. Both	11	

Microfinance Risk Management, L.L.C., `http://www.microfinance.com` **Total score:**

Figure 1: A simple poverty scorecard for Brazil (no points)

Entity	Name	ID	Date (DD/MM/YY)
Participant:			Joined:
Field agent:			Today:
Service point:			Household size:

Indicator	Value
1. How many members does the household have?	A. Five or more B. Four C. Three D. Two E. One
2. Do any household members ages 5 to 18 go to private school or private pre-school?	A. No B. Yes C. No members ages 5 to 18
3. How many years of schooling has the female head/spouse completed?	A. Three or less B. Four to eleven C. Twelve or more D. No female head/spouse
4. How many household members work as employees with a written contract, as civil servants for the government, or in the military?	A. None B. One C. Two or more
5. In their main occupation, how many household members are managers, administrators, professionals in the arts and sciences, mid-level technicians, or clerks?	A. None B. One or more
6. How many rooms does the residence have?	A. One to four B. Five C. Six D. Seven E. Eight or more
7. How does the household dispose of sewage?	A. Ditch, other, or no bathroom B. Simple hole, or directly into river, lake, or ocean C. Septic tank not connected to public sewage/rainwater system D. Septic tank connected to public sewage/rainwater system E. Direct connection to public sewage/rainwater system
8. Does the household have a refrigerator?	A. No B. Yes, with one door C. Yes, with two doors
9. Does the household have a washing machine?	A. No B. Yes
10. Does the household have a cellular or land-line telephone?	A. None B. Cellular but not land-line C. Land-line but not cellular D. Both

A Simple Poverty Scorecard for Brazil

1. Introduction

This paper presents an easy-to-use poverty scorecard that pro-poor programs in Brazil can use to estimate the likelihood that a household has income below a given poverty line. This poverty likelihood can then be used to monitor groups' poverty rates at a point in time, to track changes in groups' poverty rates between two points in time, and to target services to households.

The direct approach to poverty measurement via surveys is difficult and costly. As a case in point, Brazil's 2008 National Household Survey (*Pesquisa Nacional por Amostra de Domicílios*, PNAD) runs more than 25 pages and covers about 250 questions.

In contrast, the indirect approach via poverty scoring is simple, quick, and inexpensive. It uses ten verifiable indicators (such as "How many rooms does the residence have?" and "Does the household have a washing machine?") to get a score that is highly correlated with poverty status as measured by income from the exhaustive survey.

The poverty scorecard here differs from "proxy means tests" (Coady, Grosh, and Hoddinott, 2002) in that it is tailored to the capabilities and purposes not of national governments but rather of local, pro-poor organizations. The feasible poverty-measurement options for local organizations are typically subjective and relative (such as participatory wealth ranking by skilled field workers) or blunt (such as rules based on land-ownership or housing quality). These approaches may be costly, their results are not comparable across organizations or across countries, and their accuracy and precision are unknown.

The scorecard here can be used by organizations who want to know what share of their participants are below a poverty line, perhaps because they want to relate their poverty status to the Millennium Development Goals' $1.25/day poverty line at 2005 purchase-power parity (PPP). It can also be used by USAID microenterprise partners who want to report how many of their participants are among the poorest half of people below the national poverty line. Or it can be used by organizations that want to measure movement across a poverty line (for example, Daley-Harris, 2009). The simple poverty scorecard is an income-based, objective tool with known accuracy that can serve for monitoring, management, and/or targeting. While income surveys are difficult and costly even for governments, a simple, inexpensive scorecard can be feasible for many local, pro-poor organizations.

The statistical approach here aims to be understood by non-specialists. After all, if managers are to adopt poverty scoring on their own and apply it to inform their decisions, they must first trust that it works. Transparency and simplicity build trust. Getting "buy-in" matters; proxy means tests and regressions on the "determinants of poverty" have been around for three decades, but they are rarely used to inform decisions by local pro-poor organizations. This is not because these tools do not work, but because they are presented (when they are presented at all) as tables of regression coefficients incomprehensible to non-specialists (with indicator names such as "`LGHHSZ_2`", negative points, and points with many decimal places). Thanks to the predictive-modeling phenomenon known as the "flat maximum", simple scorecards are about as accurate as complex ones.

The technical approach here is innovative in how it associates scores with poverty likelihoods, in the extent of its accuracy tests, and in how it derives formulas for standard errors. Although these accuracy tests are simple and commonplace in statistical practice and in the for-profit field of credit-risk scoring, they have rarely been applied to poverty scorecards.

The scorecard (Figure 1) is based on data from the 2008 PNAD from the *Instituto Brasileiro de Geografia e Estatística*). Indicators are selected to be:

- Inexpensive to collect, easy to answer quickly, and simple to verify
- Strongly correlated with poverty
- Liable to change over time as poverty status changes

All points in the scorecard are zeroes or positive integers, and total scores range from 0 (most likely below a poverty line) to 100 (least likely below a poverty line). Non-specialists can collect data and tally scores on paper in the field in five to ten minutes.

Poverty scoring can be used to estimate three basic quantities. First, it can estimate a particular household's "poverty likelihood", that is, the probability that the household has per-capita income below a given poverty line.

Second, poverty scoring can estimate the poverty rate of a group of households at a point in time. This is the average poverty likelihood of households in the group.

Third, poverty scoring can estimate changes in the poverty rate for a given group of households (or for two independent samples, both of which are representative of the same group) between two points in time. This estimate is simply the change in the average poverty likelihood of the group(s) of households over time.

Poverty scoring can also be used for targeting services to poorer households. To help managers choose a targeting cut-off, this paper reports several measures of targeting accuracy for a range of possible cut-offs.

This paper presents a single scorecard whose indicators and points are derived from Brazil's one-minimum-wage poverty line and data on household income. Scores from this scorecard are calibrated to poverty likelihoods for nine poverty lines.

The scorecard is constructed and calibrated using a sub-sample from the 2008 PNAD. Its accuracy is then validated on a different sub-sample from the 2008 PNAD. While all three scoring estimators are unbiased when applied to the population from which they were derived (that is, they match the true value on average in repeated samples from the same population from which the scorecard is built), they are—like all predictive models—biased to some extent when applied to a different population.[1]

[1] Examples of "different populations" include nationally representative samples at another point in time or non-representative sub-groups (Tarozzi and Deaton, 2007).

Thus, while the indirect scoring approach is less costly than the direct survey approach, it is also biased. (The direct survey approach is unbiased by definition.) There is bias because scoring must assume that the relationships between indicators and poverty will be the same in the future as they are in the data used to build the scorecard. Scoring must also assume that these relationships will be the same in all sub-groups as in the population as a whole. Of course, these assumptions—ubiquitous and inevitable in predictive modeling—hold only partly.

When applied to the validation sample for Brazil with the half-minimum-wage poverty line and $n = 16{,}384$, the average difference between scorecard estimates of groups' poverty rates and true rates at a point in time is +0.5 percentage points. Across all nine lines, the average absolute difference is 0.3 percentage points, and the maximum absolute difference is 0.5 percentage points.

Because the validation sample is representative of the same population as the data that is used to construct the scorecard and because all the data come from the same time frame, the scorecard estimators are unbiased and these observed differences are due to sampling variation; the average difference would be zero if the 2008 PNAD were to be repeatedly redrawn and then divided into sub-samples before repeating the entire scorecard-building and accuracy-testing process.

For $n = 16{,}384$, the 90-percent confidence intervals for these estimates are +/–0.5 percentage points or less. For $n = 1{,}024$, these intervals are +/–2.0 percentage points or less.

Section 2 below documents data, poverty rates, and poverty lines for Brazil. Sections 3 and 4 describe scorecard construction and offer practical guidelines for use. Sections 5 and 6 detail the estimation of households' poverty likelihoods and of groups' poverty rates at a point in time. Section 7 discusses estimating changes in poverty rates, and Section 8 covers targeting. Section 9 places the new scorecard here in the context of similar existing exercises for Brazil. The final section is a summary.

2. Data and poverty lines

This section discusses the data used to construct and validate the poverty scorecard. It also documents the poverty lines to which scores are calibrated.

2.1 Data

The scorecard is based on data from the 109,640 households in the 2008 PNAD with valid income data. This is the most recent national income survey available for Brazil.[2] Households are randomly divided into three sub-samples (Figure 2):

- *Construction* for selecting indicators and points
- *Calibration* for associating scores with poverty likelihoods
- *Validation* for measuring accuracy on data not used in construction or calibration

2.2 Poverty rates and poverty lines

2.2.1 Rates

As a general definition, the *poverty rate* is the share of people in a given group who live in households whose total household income (divided by the number of members) is below a given poverty line.

Beyond this general definition, there two special cases, *household-level poverty rates* and *person-level poverty rates*. With household-level rates, each household is counted as if it had only one person, regardless of true household size, so all households

---

[2] Expenditure in the *Pesquisa de Orcamentos Familiares* (POF) would be preferred to income in the PNAD, but the most recent POF is from 2002/3.

are counted equally. With person-level rates (the "head-count index"), each household is weighted by the number of people in it, so larger households have greater weight.

For example, consider a group of two households, the first with one member and the second with two members. Suppose further that the first household has per-capita income above a poverty line (it is "non-poor") and that the second household has per-capita income below a poverty line (it is "poor"). The household-level rate counts both households as if they had only one member and so gives a poverty rate for the group of $1 \div (1 + 1) = 50$ percent. In contrast, the person-level rate weighs each household by the number of people in it and so gives a poverty rate for the group of $2 \div (1 + 2) = 67$ percent.

Whether the household-level rate or the person-level rate is most relevant depends on the situation. If an organization's "participants" include all the people in a household, then the person-level rate is relevant. Governments, for example, are concerned with the well-being of their people, regardless of how those people are arranged in households, so governments typically report person-level poverty rates.

If an organization has only one "participant" per household, however, then the household-level rate is relevant. For example, if a microlender has only one borrower per household, then it might want to report household-level poverty rates.

The poverty scorecard here is constructed using Brazil's 2008 PNAD and household-level lines. Scores are calibrated to household-level poverty likelihoods, and accuracy is measured for household-level rates. This household-level focus reflects the belief that it is the most relevant for most pro-poor organizations.

Organizations can estimate person-level poverty rates by taking a household-size-weighted average of the household-level poverty likelihoods. It is also possible to construct a scorecard based on person-level lines, to calibrate scores to person-level likelihoods, and to measure accuracy for person-level rates, but it is not done here.

2.2.2 Poverty lines

Figure 2 reports poverty lines and household- and person-level poverty rates for Brazil, based on the 2008 PNAD.

Brazil has no official poverty lines (Ferreira Loureiro and Cirilo Suliano, 2009). According to World Bank (2007, p. ii):

> The most commonly used set of poverty lines for policy are the "administrative poverty lines" that are typically set at arbitrary low levels of income such as fractions of the minimum wage (*e.g.*, one-half or one-fourth). . . . These cut-offs have been widely used for determining eligibility for social programs. In fact, most social-transfer programs use these cut-offs, including the *Bolsa Familia* and its predecessors (*Bolsa Escola*, *Bolsa Alimentação*, *Cartão Alimentação* under *Fome Zero*, and *Auxilio Gas*), state and municipal safety-net programs, and other constitutional social-assistance programs such as the BPC-LOAS programs for poor elderly and disabled. These cut-offs are also widely used in the government's Multi-Year Plan.

These "administrative poverty lines" have two limitations. First, they are applied country-wide, ignoring regional differences in cost-of-living. According to World Bank (2007, pp. 11, 31), Brazil has "no satisfactory spatial cost-of-living index" even though "price variations across this continent-sized nation are substantial." Second, the "administrative poverty lines" are based on income, not expenditure, and furthermore the PNAD has "incomplete measurement of income (particularly for income from transfers, housing, in-kind benefits, self-employment, and agricultural production for own-consumption). . . . [This] may result in inaccurate measures of poverty for two important groups: self-employed informal sector workers and cultivating households" (World Bank, 2007, p. 1).

Nevertheless, the best option with the PNAD for Brazil is a nationwide income-based poverty line. Thus, the scorecard here is built using the one-minimum-wage poverty line, and the examples focus on the half-minimum-wage line, as this is the line the most relevant for policy.

Because local pro-poor organizations in Brazil may want to use different or various poverty lines, this paper calibrates scores from its single scorecard to poverty likelihoods for nine lines:

- Half-minimum-wage
- Quarter-minimum-wage
- One-minimum-wage
- Two-minimum-wage
- USAID "extreme"
- $1.25/day 2005 PPP
- $2.50/day 2005 PPP
- $3.75/day 2005 PPP
- $5.00/day 2005 PPP

The lines based on multiples of the minimum wage are self-explanatory.

The USAID "extreme" line is defined as the median aggregate household per-capita income of people (not households) below the national line (U.S. Congress, 2002), here taken as the half-minimum-wage line.

The \$1.25/day 2005 PPP line is derived from:

- 2005 PPP exchange rate for "individual consumption expenditure by households" (World Bank, 2008): BRL1.57 per \$1.00
- Price deflators for Brazil overall: 119.2853 in September 2008, and 107.592 for 2005 on average[3]

Using the formula in Sillers (2006), the \$1.25/day 2005 PPP line for Brazil as a whole in September 2008 is:

$$(\text{2005 PPP exchange rate}) \cdot \$1.25 \cdot \frac{\text{CPI}_{\text{Sept. 2008}}}{\text{CPI}_{\text{Ave. 2005}}} =$$
$$\left(\frac{\text{BRL1.57}}{\$1.00}\right) \cdot \$1.25 \cdot \frac{119.2853}{107.592} = \text{BRL2.18.}$$

The \$2.50/day, \$3.75/day, and \$5.00/day 2005 PPP lines are multiples of the \$1.25/day 2005 PPP line.

[3] Derived from `http://www.gwu.edu/~ibi/Statistics%20PDF%20Files/IPCA%20Price%20Index.pdf`, retrieved 2 February 2010.

3. Scorecard construction

For the Brazil scorecard, about 85 potential indicators are initially prepared in the areas of:

- Family composition (such as household size)
- Education (such as attendance by children at private schools)
- Employment (such as the number of household members who work as employees with a written contract, as civil servants for the government, or in the military)
- Housing (such as the number of rooms in the residence)
- Ownership of durable goods (such as washing machines or refrigerators)

Figure 3 lists all the candidate indicators, ranked by the entropy-based "uncertainty coefficient" that is a measure of how well an indicator predicts poverty on its own (Goodman and Kruskal, 1979). For a given indicator, responses are ordered starting with those associated with higher poverty likelihoods.

The scorecard also aims to measure *changes* in poverty through time. This means that, when selecting indicators and holding other considerations constant, preference is given to more sensitive indicators. For example, ownership of a washing machine is probably more likely to change in response to changes in poverty than is the age of the male head/spouse.

The scorecard itself is built using the one-minimum-wage poverty line and Logit regression on the construction sub-sample. Indicator selection uses both judgment and statistics (forward stepwise, based on "c"). The first step is to use Logit to build one scorecard for each candidate indicator. Each scorecard's accuracy is taken as "c", a measure of ability to rank by poverty status (SAS Institute Inc., 2004).

One of these one-indicator scorecards is then selected based on several factors (Schreiner *et al.*, 2004; Zeller, 2004), including improvement in accuracy, likelihood of acceptance by users (determined by simplicity, cost of collection, and "face validity" in terms of experience, theory, and common sense), sensitivity to changes in poverty status, variety among indicators, and verifiability.

A series of two-indicator scorecards are then built, each based on the one-indicator scorecard selected from the first step, with a second candidate indicator added. The best two-indicator scorecard is then selected, again based on "c" and judgment. These steps are repeated until the scorecard has 10 indicators.

This algorithm is the Logit analogue to the familiar R^2-based stepwise with least-squares regression. It differs from naïve stepwise in that the criteria for selecting indicators include not only statistical accuracy but also judgment and non-statistical factors. The use of non-statistical criteria can improve robustness through time and helps ensure that indicators are simple and make sense to users.

The final step is to transform the Logit coefficients into non-negative integers such that total scores range from 0 (most likely below a poverty line) to 100 (least likely below a poverty line).

The single poverty scorecard here applies to all of Brazil. Tests for Mexico and India (Schreiner, 2006a and 2006b), Sri Lanka (Narayan and Yoshida, 2005), and Jamaica (Grosh and Baker, 1995) suggest that segmenting scorecards by urban/rural does not improve targeting much, although such segmentation may improve the accuracy of estimated poverty rates (Tarozzi and Deaton, 2007).

4. Practical guidelines for scorecard use

The main challenge of scorecard design is not to squeeze out the last drops of accuracy but rather to improve the chances that scoring is actually used (Schreiner, 2005). When scoring projects fail, the reason is not usually technical inaccuracy but rather the failure of an organization to decide to do what is needed to integrate scoring in its processes and to learn to use it properly (Schreiner, 2002). After all, most reasonable scorecards predict tolerably well, thanks to the empirical phenomenon known as the "flat maximum" (Falkenstein, 2008; Hand, 2006; Baesens *et al.*, 2003; Lovie and Lovie, 1986; Kolesar and Showers, 1985; Stillwell, Barron, and Edwards, 1983; Dawes, 1979; Wainer, 1976; Myers and Forgy, 1963). The bottleneck is less technical and more human, not statistics but organizational-change management. Accuracy is easier to achieve than adoption.

The scorecard here is designed to encourage understanding and trust so that users will want to adopt it and use it properly. Of course, accuracy is important, but so are simplicity, ease-of-use, and "face validity". Programs are more likely to collect data, compute scores, and pay attention to the results if, in their view, scoring does not imply a lot of additional work and if the whole process generally seems to make sense.

To this end, the poverty scorecard fits on a single page. The construction process, indicators, and points are simple and transparent. Additional work is minimized; non-specialists can compute scores by hand in the field because the scorecard has:

- Only ten indicators
- Only categorical indicators
- Simple weights (non-negative integers, and no arithmetic beyond addition)

The scorecard in Figure 1 is ready to be photocopied and can be used with a simple spreadsheet database (Microfinance Risk Management, L.L.C., 2010) that records identifying information, indicator values, scores, and poverty likelihoods.

A field worker using the paper scorecard would:

- Record participant identifiers
- Read each question verbatim from the scorecard
- Circle each response and its points
- Write the points in the far-right column
- Add up the points to get the total score
- Implement targeting policy (if any)
- Deliver the paper scorecard to a central office for data entry and filing

4.1 Quality control

Of course, field workers must be trained. High-quality outputs require high-quality inputs. If organizations or field workers gather their own data and if they believe that they have an incentive to exaggerate poverty rates (for example, if funders reward them for higher poverty rates), then it is wise to do on-going quality control via data review and random audits (Matul and Kline, 2003).[4] IRIS Center (2007a) and Toohig (2008) are useful nuts-and-bolts guides for planning, budgeting, training field workers and supervisors, logistics, sampling, interviewing, piloting, recording data, and controlling quality.

In particular, while collecting indicators for the poverty scorecard is relatively easier than most alternatives, it is still absolutely difficult. Training and explicit definitions of the terms and concepts in the scorecard is essential.[5] For example, one study in Nigeria finds distressingly low inter-rater and test-retest correlations for indicators as seemingly simple and obvious as whether the household owns an automobile (Onwujekwe, Hanson, and Fox-Rushby, 2006).

For the example of a Mexican social program that uses self-reported indicators in the first stage of scorecard-based targeting, Martinelli and Parker (2007) find that "underreporting [of asset ownership] is widespread but not overwhelming, except for a few goods . . . [and] overreporting is common for a few goods, which implies that self-reporting may lead to the exclusion of deserving households" (pp. 24–25). Still, as done in the second stage of the Mexican program, field agents using poverty scoring can verify responses with a home visit and correct any false reports.

---

[4] If an organization does not want field workers to know the points associated with indicators, then they can use the version of Figure 1 without points and apply the points later at the central office.

[5] Appendix A is a guide for interpreting indicators in Brazil's poverty scorecard.

4.2 Implementation and sampling

In terms of implementation and sample design, an organization must make choices about:

- Who will do the scoring
- How scores will be recorded
- What participants will be scored
- How many participants will be scored
- How frequently participants will be scored
- Whether scoring will be applied at more than one point in time
- Whether the same participants will be scored at more than one point in time

The non-specialists who apply the scorecard with participants in the field can be:

- Employees of the organization
- Third-party contractors

Responses, scores, and poverty likelihoods can be recorded:

- On paper in the field and then filed at an office
- On paper in the field and then keyed into a database or spreadsheet at an office
- On portable electronic devices in the field and then downloaded to a database

The subjects to be scored can be:

- All participants
- A representative sample of all participants
- All participants in a representative sample of branches
- A representative sample of all participants in a representative sample of branches
- A representative sample of a sub-group that is relevant for a particular question

If not determined by other factors, the number of participants to be scored can be derived from sample-size formulas (presented later) for a desired confidence level and a desired confidence interval.

Frequency of application can be:

- At in-take of new clients only (precluding measuring changes in poverty rates)
- As a once-off project for current participants (precluding measuring changes)
- Once a year or at some other fixed time interval (allowing measuring changes)
- Each time a field worker visits a participant at home (allowing measuring changes)

When the scorecard is applied more than once in order to measure changes in poverty rates, it can be applied with:

- Different sets of participants, with each set representative of a given group
- A single set of participants

An example collection of implementation and design choices is provided by BRAC and ASA, two microlenders in Bangladesh (each with more than 7 million participants) who are applying a poverty scorecard (Chen and Schreiner, 2009a). Their design is that loan officers in a random sample of branches score all their clients each time they visit a homestead (about once a year) as part of their standard due diligence prior to loan disbursement. Responses in the field are recorded on paper before being sent to a central office to be entered into a spreadsheet database. The sampling plans of ASA and BRAC cover 50,000–100,000 participants each, which is far more than would be required to inform most decisions at a typical pro-poor organization.

5. Estimates of household poverty likelihoods

The sum of scorecard points for a household is called the *score*. For Brazil, scores range from 0 (most likely below a poverty line) to 100 (least likely below a poverty line). While higher scores indicate less likelihood of being below a poverty line, the scores themselves have only relative units. For example, doubling the score does not double the likelihood of being above a poverty line.

To get absolute units, scores must be converted to *poverty likelihoods*, that is, probabilities of being below a poverty line. This is done via simple look-up tables. For the example of the half-minimum-wage line with the 2008 PNAD, scores of 25–29 correspond to a poverty likelihood of 54.2 percent, and scores of 30–34 correspond to a poverty likelihood of 41.1 percent (Figure 4).

The poverty likelihood associated with a score varies by poverty line. For example, scores of 25–29 are associated with a poverty likelihood of 54.2 percent for the half-minimum-wage line but 16.1 percent for the quarter-minimum-wage line.[6]

---

[6] Starting with Figure 4, many figures have nine versions, one for each of the nine poverty lines. Single tables that pertain to all poverty lines are placed with the tables for the half-minimum-wage line.

5.1 Calibrating scores with poverty likelihoods

A given score is non-parametrically associated ("calibrated") with a poverty likelihood by defining the poverty likelihood as the share of households in the calibration sub-sample who have the score and who are below a given poverty line.

For the example of the half-minimum-wage line (Figure 5), there are 6,725 (normalized) households in the calibration sub-sample with a score of 25–29, of whom 3,646 (normalized) are below the poverty line. The estimated poverty likelihood associated with a score of 25–29 is then 54.2 percent, because 3,646 ÷ 6,725 = 0.542.

As another illustration, consider the half-minimum-wage line and a score of 30–34. Now there are 7,701 (normalized) households in the calibration sample, of whom 3,164 (normalized) are below the line (Figure 5). Thus, the poverty likelihood for this score is 3,164 ÷ 7,701 = 0.411, or 41.1 percent.

The same method is used to calibrate scores with estimated poverty likelihoods for all nine poverty lines.

Figure 6 shows, for all scores, the likelihood that income falls in a range demarcated by two adjacent poverty lines. For example, the daily income of someone with a score of 25–29 falls in the following ranges with probability:

- 7.1 percent   less than $1.25/day 2005 PPP
- 9.0 percent   between $1.25/day 2005 PPP and quarter-minimum-wage
- 7.9 percent   between quarter-minimum-wage and $2.50/day 2005 PPP
- 23.6 percent   between $2.50/day 2005 PPP and $3.75/day 2005 PPP
- 6.7 percent   between $3.75/day 2005 PPP and half-minimum-wage
- 13.3 percent   between half-minimum-wage and $5.00/day 2005 PPP
- 24.7 percent   between $5.00/day 2005 PPP and one-minimum-wage
- 7.2 percent   between one-minimum-wage and two-minimum-wage
- 0.6 percent   more than two-minimum-wage

Even though the scorecard is constructed partly based on judgment, this calibration process produces poverty likelihoods that are objective, that is, derived from survey data on income and quantitative poverty lines. The poverty likelihoods would be objective even if indicators and/or points were selected without any data at all. In fact, objective scorecards of proven accuracy are often based only on judgment (Fuller, 2006; Caire, 2004; Schreiner *et al.*, 2004). Of course, the scorecard here is constructed with both data and judgment. The fact that this paper acknowledges that some choices in scorecard construction—as in any statistical analysis—are informed by judgment in no way impugns the objectivity of the poverty likelihoods, as this depends on using data in score calibration, not on using data (and nothing else) in scorecard construction.

Although the points in Brazil's poverty scorecard are transformed coefficients from a Logit regression, scores are not converted to poverty likelihoods via the Logit formula of $2.718281828^{\text{score}}$ x $(1+ 2.718281828^{\text{score}})^{-1}$. This is because the Logit formula is esoteric and difficult to compute by hand. It is more intuitive to define the poverty likelihood as the share of households with a given score in the calibration sample who are below a poverty line. Converting scores to poverty likelihoods requires no arithmetic at all, just a look-up table. This non-parametric calibration can also improve accuracy, especially with large calibration samples.

5.2 Accuracy of estimates of households' poverty likelihoods

As long as the relationships between indicators and poverty do not change and as long as the scorecard is applied to households who are representative of the same population from which the scorecard is constructed, this calibration process produces unbiased estimates of poverty likelihoods. *Unbiased* means that in repeated samples from the same population, the average estimate matches the true poverty likelihood. The scorecard also produces unbiased estimates of poverty rates at a point in time, as well as unbiased estimates of changes in poverty rates between two points in time.[7]

But the relationships between indicators and poverty do change with time, and they also change across sub-groups in Brazil's population. Thus, the scorecard will generally be biased when applied after the end date of fieldwork for the 2008 PNAD (as it must be applied in practice) or when applied with non-nationally representative groups (as it probably will be applied by local, pro-poor organizations).

[7] This follows because these estimates of groups' poverty rates are linear functions of the unbiased estimates of households' poverty likelihoods.

How accurate are these estimates of households' poverty likelihoods, given the assumption of representativeness? To check, the scorecard is applied to 1,000 bootstrap samples of size n = 16,384 from the validation sub-sample. Bootstrapping entails (Efron and Tibshirani, 1993):

- Score each household in the validation sample
- Draw a new bootstrap sample *with replacement* from the validation sample
- For each score, compute the true poverty likelihood in the bootstrap sample, that is, the share of households with the score who have income below a poverty line
- For each score, record the difference between the estimated poverty likelihood (Figure 4) and the true poverty likelihood in the bootstrap sample
- Repeat the previous three steps 1,000 times
- For each score, report the average difference between estimated and true poverty likelihoods across the 1,000 bootstrap samples
- For each score, report the two-sided interval containing the central 900, 950, or 990 differences between estimated and true poverty likelihoods

For each score range and for n = 16,384, Figure 7 shows the average difference between estimated and true poverty likelihoods as well as confidence intervals for the differences.

For the half-minimum-wage line in the validation sample, the average poverty likelihood across bootstrap samples for scores of 25–29 is too high by 1.8 percentage points. For scores of 30–34, the estimate is too high by 3.0 percentage points.[8]

The 90-percent confidence interval for the differences for scores of 25–29 is +/– 2.7 percentage points (Figure 7). This means that in 900 of 1,000 bootstraps, the

[8] These differences are not zero, despite the estimator's unbiasedness, because the scorecard comes from a single sample. The average difference by score would be zero if samples were repeatedly drawn from the population and split into sub-samples before repeating the entire construction and calibration process.

difference between the estimate and the true value is between –0.9 and +4.5 percentage points (because +1.8 – 2.7 = –0.9, and +1.8 + 2.7 = +4.5). In 950 of 1,000 bootstraps (95 percent), the difference is +1.8 +/–3.2 percentage points, and in 990 of 1,000 bootstraps (99 percent), the difference is +1.8 +/–4.0 percentage points.

For all scores, Figure 7 shows differences—usually small—between estimated poverty likelihoods and true values. The differences are not all zero because the validation sub-sample is a single sample that—thanks to sampling variation—differs in distribution from the construction/calibration sub-samples and from Brazil's population. Also, some score ranges have few households in them, increasing the importance of sampling variation.

For targeting, what matters is less the differences across all score ranges and more the differences in score ranges just above and just below the targeting cut-off. This mitigates the effects of bias and sampling variation on targeting (Friedman, 1997). Section 8 below looks at targeting accuracy in detail.

Of course, if estimates of groups' poverty rates are to be usefully accurate, then errors for individual households must largely balance out. As discussed in the next section, this is always the case.

Another possible source of bias is overfitting. By construction, the scorecard here is unbiased, but it may still be *overfit* when applied after the end of field work for the 2008 PNAD. That is, the scorecard may fit the 2008 data so closely that it captures not only some real patterns but also some false patterns that, due to sampling variation, show up only in the 2008 data. Or the scorecard may be overfit in the sense that it is not robust to changes in the relationships between indicators and poverty over time. Finally, the scorecard could also be overfit when it is applied to samples from non-nationally representative sub-groups.

Overfitting can be mitigated by simplifying the scorecard and by not relying only on data but rather also considering experience, judgment, and theory. Of course, the scorecard here does this. Bootstrapping scorecard construction—which is not done here—can also mitigate overfitting by reducing (but not eliminating) dependence on a single sampling instance. Combining scorecards can also help, at the cost of complexity. Simplifying the scorecard can also reduce overfitting (at the cost of decreased precision), although the poverty scorecard here is already parsimonious with limited scope for simplification. Often the best option is simply to update the scorecard as soon as new data is available.

In any case, errors in individual households' likelihoods largely balance out in the estimates of groups' poverty rates (see the next section). Furthermore, much of the differences between scorecard estimates and true values may come from non-scorecard sources. These factors can be addressed only by improving data quantity and quality, which is beyond the scope of the scorecard.

6. Estimates of a group's poverty rate at a point in time

A group's estimated poverty rate at a point in time is the average of the estimated poverty likelihoods of the individual households in the group.

To illustrate, suppose a program samples three households on Jan. 1, 2010 and that they have scores of 20, 30, and 40, corresponding to poverty likelihoods of 68.7, 41.1, and 17.4 percent (half-minimum-wage line, Figure 4). The group's estimated poverty rate is the households' average poverty likelihood of (68.7 + 41.1 + 17.4) ÷ 3 = 42.4 percent.[9]

6.1 Accuracy of estimated poverty rates at a point in time

How accurate is this estimate? For a range of sample sizes, Figure 9 reports average differences between estimated and true poverty rates as well as precision (confidence intervals for the differences) for the scorecard applied to 1,000 bootstrap samples from the validation sample.

Summarizing Figure 9 across poverty lines and years for n = 16,384, Figure 8 shows that the absolute differences between estimated poverty rates and true rates for the scorecard applied to the validation sample are 0.5 percentage points or less. The

[9] The group's poverty rate is *not* the poverty likelihood associated with the average score. Here, the average score is (20 + 30 + 40) ÷ 3 = 30, and the poverty likelihood associated with the average score is 41.1 percent. This is not the 42.4 percent found as the average of the three poverty likelihoods associated with each of the three scores.

average absolute difference across the nine poverty lines for the validation sample is 0.3 percentage points.

In terms of precision, the 90-percent confidence interval for a group's estimated poverty rate at a point in time with n = 16,384 is +/–0.5 percentage points or less (Figure 8). This means that in 900 of 1,000 bootstraps of this size, the absolute difference between the estimate and the average estimate is 0.5 percentage points or less.

In the specific case of the half-minimum-wage line and the validation sample, 90 percent of all samples of n = 16,384 produce estimates that differ from the true value in the range of +0.5 – 0.4 = +0.1 to +0.5 + 0.4 = +0.9 percentage points. This is because +0.5 is the average difference and +/–0.4 is its 90-percent confidence interval. The average difference is +0.5 because the average scorecard estimate is too high by 0.5 percentage points; the scorecard tends to estimate a poverty rate of 24.3 percent for the validation sample, but the true value is 23.8 percent (Figure 2).

Part of these differences is due to sampling variation in the division of the 2008 PNAD into three sub-samples. Of course, estimates of poverty rates at a point in time from now on will be most accurate for periods that resemble September 2008, the reference period for data in the 2008 PNAD.

6.2 Standard-error formula for estimates of poverty rates at a point in time

How precise are the point-in-time estimates? Because they are averages, the estimates have a Normal distribution and can be characterized by their average difference vis-à-vis true values, along with the standard error of the average difference.

To derive a formula for the standard errors of estimated poverty rates at a point in time for indirect measurement via poverty scorecards (Schreiner, 2008a), note that the textbook formula (Cochran, 1977) that relates confidence intervals with standard errors in the case of direct measurement of poverty rates is $c = +/- z \cdot \sigma$, where:

c is a confidence interval as a proportion (*e.g.*, 0.02 for +/–2 percentage points),

z is from the Normal distribution and is $\begin{cases} 1.64 \text{ for confidence levels of 90 percent} \\ 1.96 \text{ for confidence levels of 95 percent} \\ 2.58 \text{ for confidence levels of 99 percent} \end{cases}$,

σ is the standard error of the estimated poverty rate, that is, $\sqrt{\dfrac{p \cdot (1-p)}{n}}$,

p is the proportion of households below the poverty line in the sample, and

n is the sample size.

For example, with a sample n = 16,384, 90-percent confidence (z = 1.64), and a poverty rate p of 23.8 percent (the true rate in the validation sample for the half-minimum-wage line in Figure 2), the confidence interval c is

$$+/- z \cdot \sqrt{\frac{p \cdot (1-p)}{n}} = +/- 1.64 \cdot \sqrt{\frac{0.238 \cdot (1-0.238)}{16{,}384}} = +/–0.546 \text{ percentage points.}$$

Poverty scorecards, however, do not measure poverty directly, so this formula is not applicable. To derive a formula for the Brazil scorecard, consider Figure 9, which reports empirical confidence intervals c for the differences for the scorecard applied to 1,000 bootstrap samples of various sample sizes from the validation sample. For n = 16,384, the half-minimum-wage line, and the validation sub-sample, the 90-percent confidence interval is +/–0.420 percentage points.[10] Thus, the ratio of confidence intervals for poverty scoring versus direct measurement is 0.420 ÷ 0.546 = 0.77.

Now consider the same case, but with n = 8,192. The confidence interval under direct measurement is $+/-1.64 \cdot \sqrt{\frac{0.238 \cdot (1-0.238)}{8{,}192}}$ = +/–0.772 percentage points. The empirical confidence interval with the Brazil scorecard for the half-minimum-wage line (Figure 9) is +/–0.630 percentage points. Thus for n = 8,192, the ratio for poverty scoring to direct measurement is 0.630 ÷ 0.772 = 0.82.

This ratio of 0.82 for n = 8,192 is not far from the ratio of 0.77 for n = 16,384. Indeed, across all sample sizes of 256 or more in Figure 9, the average ratio turns out to be 0.79, implying that confidence intervals for indirect estimates of poverty rates via the Brazil scorecard and this poverty line are about 21 percent narrower than those for direct estimates. This 0.79 appears in Figure 8 as the "α factor" because if α = 0.79, then the formula relating confidence intervals c and standard errors σ for the Brazil

[10] Due to rounding, Figure 9 displays 0.4, not 0.420.

scorecard is $c = +/- z \cdot \alpha \cdot \sigma$. The standard error σ for point-in-time estimates of poverty rates via scoring is $\alpha \cdot \sqrt{\frac{p \cdot (1-p)}{n}}$.

In general, α could be more or less than 1.00. When α is less than 1.00, it means that the scorecard is more precise than direct measurement. This occurs for all nine lines for the validation sample in Figure 8.

The formula relating confidence intervals to standard errors for poverty scoring can be rearranged to give a formula for determining sample size n before measurement.[11] If $\hat{p}$ is the expected poverty rate before measurement, then the formula for n based on the desired confidence level that corresponds to z and the desired confidence interval $+/-c$ under poverty scoring is $n = \left(\frac{\alpha \cdot z}{c}\right)^2 \cdot \hat{p} \cdot (1 - \hat{p})$.

To illustrate how to use this, suppose $c = 0.03445$ and $z = 1.64$ (90-percent confidence), and $\hat{p} = 0.2385$ (the average poverty rate for the half-minimum-wage line in the construction and calibration sub-samples, Figure 2). Then the formula gives $n = \left(\frac{0.79 \cdot 1.64}{0.03445}\right)^2 \cdot 0.2385 \cdot (1 - 0.2385) = 257$, which is almost the same as the sample size of 256 observed for these parameters in Figure 9.

[11] IRIS Center (2007a and 2007b) says that a sample size of $n = 300$ is sufficient for reporting estimated poverty rates to USAID. If a scorecard is as precise as direct measurement, if the expected (before measurement) poverty rate is 50 percent, and if the confidence level is 90 percent, then $n = 300$ implies a confidence interval of +/–2.2 percentage points. In fact, USAID has not specified confidence levels or intervals. Furthermore, the expected poverty rate may not be 50 percent, and the scorecard could be more or less precise than direct measurement.

Of course, the α factors in Figure 8 are specific to Brazil, its poverty lines, its poverty rates, and this scorecard. The method for deriving the formulas, however, is valid for any poverty scorecard following the basic approach in this paper.

In practice after the end of the 2008 PNAD field work in September 2008, an organization would select a poverty line (say, the half-minimum-wage line), select a desired confidence level (say, 90 percent, or $z = 1.64$), select a desired confidence interval (say, +/–2.0 percentage points, or $c = 0.02$), make an assumption about $\hat{p}$ (perhaps based on a previous measurement such as the 23.8 percent average for the half-minimum-wage line in Figure 2), look up α (here, 0.79), assume that the scorecard will work the same in the future and/or for non-nationally representative sub-groups,[12] and then compute the required sample size. In this illustration,

$$n = \left(\frac{0.79 \cdot 1.64}{0.02}\right)^2 \cdot 0.238 \cdot (1 - 0.238) = 762.$$

[12] This paper reports accuracy for the scorecard applied to the validation sample, but it cannot test accuracy for later years or for other groups. Performance will deteriorate with time to the extent that the relationships between indicators and poverty change.

7. Estimates of changes in group poverty rates over time

The change in a group's poverty rate between two points in time is estimated as the change in the average poverty likelihood of the households in the group. With data for 2008 only, this paper cannot estimate changes over time, nor can it present sample-size formula. Nevertheless, the relevant concepts are presented here because, in practice, pro-poor organizations can apply the scorecard to measure change over time.

7.1 Warning: Change is not impact

Scoring can estimate change. Of course, change could be for the better or for the worse, and scoring does not indicate what caused change. This point is often forgotten, confused, or ignored, so it bears repeating: poverty scoring simply estimates change, and it does not, in and of itself, indicate the reason for the change. In particular, estimating the impact of program participation on poverty status requires knowing what would have happened to participants if they had not been participants. Knowing this requires either strong assumptions or a control group that resembles participants in all ways except participation. To belabor the point, poverty scoring can help estimate program impact only if there is some way to know what would have happened in the absence of the program. And that information must come from somewhere beyond poverty scoring.

7.2 Calculating estimated changes in poverty rates over time

Consider the illustration begun in the previous section. On Jan. 1, 2010, a program samples three households who score 20, 30, and 40 and so have poverty likelihoods of 68.7, 41.1, and 17.4 percent (half-minimum-wage line, Figure 4). The group's baseline estimated poverty rate is the households' average poverty likelihood of (68.7 + 41.1 + 17.4) ÷ 3 = 42.4 percent.

After baseline, two sampling approaches are possible for the follow-up round:

- Score a new, independent sample, measuring change by cohort across samples
- Score the same sample at follow-up as at baseline

By way of illustration, suppose that a year later on Jan. 1, 2011, the program samples three additional households who are in the same cohort as the three households originally sampled (or suppose that the program scores the same three original households a second time) and finds that their scores are now 25, 35, and 45 (poverty likelihoods of 54.2, 26.1, and 12.4 percent, half-minimum-wage line, Figure 4). Their average poverty likelihood at follow-up is (54.2 + 26.1 + 12.4) ÷ 3 = 30.9 percent, an improvement of 42.4 – 30.9 = 11.5 percentage points.[13]

This suggests that about one of eight participants moved above the poverty line in 2010. (This is a net figure; some people start above the line and end below it, and vice versa.) Among those who started below the line, about one in four (11.5 ÷ 42.4 =

[13] Of course, such a huge reduction in poverty is unlikely in a year's time, but this is just an example to show how poverty scoring can be used to estimate change.

27.1 percent) ended up above the line. Of course, poverty scoring does not reveal the reasons for this change.

7.3 Accuracy for estimated change in two independent samples

For two equal-sized independent samples, the same logic as in the previous section can be used to derive a formula relating the confidence interval c with the standard error σ of a poverty scorecard's estimate of the change in poverty rates over time:

$$c = +/- z \cdot \sigma = +/- z \cdot \alpha \cdot \sqrt{\frac{2 \cdot p \cdot (1-p)}{n}}.$$

z, c, and p are defined as before, n is the sample size at both baseline and follow-up,[14] and α is the average (across a range of bootstrapped sample sizes) of the ratio of the observed confidence intervals from a poverty scorecard and the theoretical confidence intervals from the textbook formula for direct measurement for two equal-sized independent samples.

[14] This means that, for a given precision and with direct measurement, estimating the change in a poverty rate over time requires four times as many measurements (not twice as many) as does estimating a poverty rate at a point in time.

As before, the formula for standard errors can be rearranged to give a formula for sample sizes before indirect measurement via a poverty scorecard, where $\hat{p}$ is based on previous measurements and is assumed equal at both baseline and follow-up:

$$n = 2 \cdot \left(\frac{\alpha \cdot z}{c} \right)^2 \cdot \hat{p} \cdot (1 - \hat{p}) .$$

For countries for which this α has been measured (Schreiner, 2009a, 2009b, 2009c, 2009d, 2009e, and 2008b and Chen and Schreiner, 2009a and 2009b), the simple average of α across poverty lines, years, and countries is 1.11. This is as reasonable a figure as any to use for Brazil.

To illustrate the use of the formula above to determine sample size for estimating changes in poverty rates across two independent samples, suppose the desired confidence level is 90 percent ($z = 1.64$), the desired confidence interval is 2 percentage points ($c = 0.02$), the poverty line is the half-minimum-wage line, $\alpha = 1.11$, and $\hat{p} = 0.238$ (from Figure 2). Then the baseline sample size is $n = 2 \cdot \left(\frac{1.11 \cdot 1.64}{0.02} \right)^2 \cdot 0.238 \cdot (1 - 0.238) = 3{,}005$, and the follow-up sample size is also 3,005.

7.4 Accuracy for estimated change for one sample, scored twice

The general formula relating the confidence interval c to the standard error σ when using scoring to estimate change for a single group of households, all of whom are scored at two points in time, is:[15]

$$c = +/- z \cdot \sigma = +/- z \cdot \alpha \cdot \sqrt{\frac{p_{12} \cdot (1 - p_{12}) + p_{21} \cdot (1 - p_{21}) + 2 \cdot p_{12} \cdot p_{21}}{n}}.$$

z, c, and α are defined as before, p_{12} is the share of all sampled households that move from below the poverty line to above it, and p_{21} is the share of all sampled households that move from above the line to below it.

As usual, the formula for σ can be rearranged to give a formula for sample size n before measurement. This requires an estimate (based on information available before measurement) of the expected shares of all households who cross the poverty line $\hat{p}_{12}$ and $\hat{p}_{21}$. Before measurement, it is reasonable to assume that the overall change in the poverty rate will be zero, which implies $\hat{p}_{12} = \hat{p}_{21} = \hat{p}_*$, giving:

$$n = 2 \cdot \left(\frac{\alpha \cdot z}{c}\right)^2 \cdot \hat{p}_*.$$

[15] See McNemar (1947) and Johnson (2007). John Pezzullo helped find this formula.

$\hat{p}_*$ could be anything between 0 and 0.5, so more information is needed before applying this formula. Suppose that the observed relationship between $\hat{p}_*$, the number of years y between baseline and follow-up, and $p_{\text{baseline}} \cdot (1 - p_{\text{baseline}})$ is—as in Peru (Schreiner, 2009a)—close to:

$$\hat{p}_* = -0.02 + 0.016 \cdot y + 0.47 \cdot [p_{\text{baseline}} \cdot (1 - p_{\text{baseline}})].$$

Given this, a sample-size formula for a group of households to whom the Brazil poverty scorecard is applied twice (once after the end of field work for the 2008 PNAD and then again later) is:

$$n = 2 \cdot \left(\frac{\alpha \cdot z}{c}\right)^2 \cdot \left\{-0.02 + 0.016 \cdot y + 0.47 \cdot \left[p_{\text{baseline}} \cdot \left(1 - p_{\text{baseline}}\right)\right]\right\}.$$

In Peru (the only other country for which there is a data-based estimate, Schreiner 2009a), the average α across years and poverty lines is about 1.30.

To illustrate the use of this formula, suppose the desired confidence level is 90 percent ($z = 1.64$), the desired confidence interval is 2.0 percentage points ($c = 0.02$), the poverty line is the half-minimum-wage line, and the sample will be scored first in 2010 and then again in 2013 ($y = 3$). The before-baseline poverty rate is 23.8 percent ($p_{2008} = 0.238$, Figure 2), and suppose $\alpha = 1.30$. Then the baseline sample size is $n = 2 \cdot \left(\frac{1.3 \cdot 1.64}{0.02}\right)^2 \cdot \left\{-0.02 + 0.016 \cdot 3 + 0.47 \cdot \left[0.238 \cdot \left(1 - 0.238\right)\right]\right\} = 2{,}574$. The same group of 2,574 households is scored at follow-up as well.

8. Targeting

When a program uses poverty scoring for targeting, households with scores at or below a cut-off are labeled *targeted* and treated—for program purposes—as if they are below a given poverty line. Households with scores above a cut-off are labeled *non-targeted* and treated—for program purposes—as if they are above a given poverty line.

There is a distinction between *targeting status* (scoring at or below a targeting cut-off) and *poverty status* (having income below a poverty line). Poverty status is a fact that depends on whether income is below a poverty line as directly measured by a survey. In contrast, targeting status is a program's policy choice that depends on a cut-off and on an indirect estimate from a scorecard.

Targeting is successful when households truly below a poverty line are targeted (*inclusion*) and when households truly above a poverty line are not targeted (*exclusion*). Of course, no scorecard is perfect, and targeting is unsuccessful when households truly below a poverty line are not targeted (*undercoverage*) or when households truly above a poverty line are targeted (*leakage*).

Figure 10 depicts these four possible targeting outcomes. Targeting accuracy varies by cut-off; a higher cut-off has better inclusion (but greater leakage), while a lower cut-off has better exclusion (but higher undercoverage).

A program should weigh these trade-offs when setting a cut-off. A formal way to do this is to assign net benefits—based on a program's values and mission—to each of the four possible targeting outcomes and then to choose the cut-off that maximizes total net benefits (Adams and Hand, 2000; Hoadley and Oliver, 1998).

Figure 11 shows the distribution of households by targeting outcome. For an example cut-off of 29 or less and the scorecard applied to the validation sample, outcomes for the half-minimum-wage line are:

- Inclusion: 13.7 percent are below the line and correctly targeted
- Undercoverage: 10.1 percent are below the line and mistakenly not targeted
- Leakage: 6.0 percent are above the line and mistakenly targeted
- Exclusion: 70.1 percent are above the line and correctly not targeted

Increasing the cut-off to 34 or less improves inclusion and undercoverage but worsens leakage and exclusion:

- Inclusion: 16.8 percent are below the line and correctly targeted
- Undercoverage: 7.0 percent are below the line and mistakenly not targeted
- Leakage: 10.7 percent are above the line and mistakenly targeted
- Exclusion: 65.5 percent are above the line and correctly not targeted

Which cut-off is preferred depends on total net benefit. If each targeting outcome has a per-household benefit or cost, then total net benefit for a given cut-off is:

Benefit per household correctly included	x	Households correctly included	–
Cost per household mistakenly not covered	x	Households mistakenly not covered	–
Cost per household mistakenly leaked	x	Households mistakenly leaked	+
Benefit per household correctly excluded	x	Households correctly excluded.	

To set an optimal cut-off, a program would:

- Assign benefits and costs to possible outcomes, based on its values and mission
- Tally total net benefits for each cut-off using Figure 11 for a given poverty line
- Select the cut-off with the highest total net benefit

The most difficult step is assigning benefits and costs to targeting outcomes. Any program that uses targeting—with or without scoring—should thoughtfully consider how it values successful inclusion or exclusion versus errors of undercoverage or leakage. It is healthy to go through a process of thinking explicitly and intentionally about how possible targeting outcomes are valued.

A common choice of benefits and costs is "Total Accuracy" (IRIS Center, 2005; Grootaert and Braithwaite, 1998). With "Total Accuracy", total net benefit is the number of households successfully included or successfully excluded:

Total Accuracy =	1	x	Households correctly included	–
	0	x	Households mistakenly undercovered	–
	0	x	Households mistakenly leaked	+
	1	x	Households correctly excluded.	

Figure 11 shows "Total Accuracy" for all cut-offs for Brazil's scorecard. For the half-minimum-wage line in the validation sample, total net benefit is greatest (83.9) for a cut-off of 29 or less, with about five in six households in Brazil correctly classified.

"Total Accuracy" weighs successful inclusion of households below the line the same as successful exclusion of households above the line. If a program values inclusion more (say, twice as much) than exclusion, it could reflect this by setting the benefit for inclusion to 2 and the benefit for exclusion to 1. Then the chosen cut-off would maximize (2 x Households correctly included) + (1 x Households correctly excluded).[16]

[16] Figure 11 also reports "BPAC", the Balanced Poverty Accuracy Criteria adopted by USAID as its criterion for certifying poverty scorecards. IRIS Center (2005) says that BPAC considers accuracy both in terms of the estimated poverty rate and in terms of targeting inclusion. After normalizing by the number of people below a poverty line, BPAC = (Inclusion – |Undercoverage – Leakage|) x [100 ÷ (Inclusion+Undercoverage)].

As an alternative to assigning benefits and costs to targeting outcomes and then choosing a cut-off to maximize total net benefits, a program could set a cut-off to achieve a desired poverty rate among targeted households. The third column of Figure 12 ("% targeted who are poor") shows the expected poverty rate among Brazil households who score at or below a given cut-off. For the example of the half-minimum-wage line and the validation sample, targeting households who score 29 or less would target 19.8 percent of all households (second column) and produce a poverty rate among those targeted of 69.5 percent (third column).

Figure 12 also reports two other measures of targeting accuracy. The first is a version of inclusion ("% of poor who are targeted"). For the example of the half-minimum-wage line and the validation sample with a cut-off of 29 or less, 57.7 percent of all poor households are covered.

The final targeting measure in Figure 12 is the number of successfully targeted poor households for each non-poor household mistakenly targeted (right-most column). For the half-minimum-wage line, the validation sample, and a cut-off of 29 or less, covering 2.3 poor households means leaking to one non-poor household.

9. The context of poverty scorecards for Brazil

This section discusses seven existing scorecards for Brazil in terms of their goals, methods, poverty lines, poverty definitions, indicators, cost, accuracy, and precision. The advantages of the new scorecard here are its use of the latest nationally representative data, its focus on feasibility for local, pro-poor organizations, its testing of accuracy and precision out-of-sample, and its reporting of formulas for standard errors.

9.1 Gwatkin *et al.*

Gwatkin *et al.* (2007) apply to Brazil an approach used in 56 countries with Demographic and Health Surveys (Rutstein and Johnson, 2004). They use Principal Components Analysis to make an asset index from simple, low-cost indicators available for the 13,283 households in Brazil's 1996 DHS. The PCA index is like the poverty scorecard here except that, because the DHS does not collect data on income or expenditure, it is based on a different conception of poverty, its accuracy vis-à-vis income-based poverty is unknown, and it can only be assumed to be a proxy for long-term wealth/economic status.[17] Well-known examples of the PCA asset-index approach

[17] Still, because the indicators are similar and because the "flat maximum" is important, carefully built PCA indices and expenditure-based poverty scorecards may pick up the same underlying construct (perhaps "permanent income", see Bollen, Glanville, and Stecklov, 2007), and they rank households much the same. Tests of how well rankings

include Ferguson *et al.* (2003), Sahn and Stifel (2000 and 2003), and Filmer and Pritchett (2001). Some applications to Brazil are discussed later in this section.

The 13 indicators in Gwatkin *et al.* are similar to those in the new scorecard here in terms of their simplicity, low cost, and verifiability:

- Characteristics of the residence:
    - Presence of electricity
    - Type of floors
    - Type of walls
    - Type of roof
    - Source of drinking water
    - Type of toilet arrangement
    - Number of people per sleeping room
- Ownership of consumer durables:
    - Radio
    - Television
    - Refrigerator
    - Car
- Presence of a domestic worker not related to the head
- Whether any household members work their own or family's agricultural land

Gwatkin *et al.* discuss three basic uses for their index:

- Segmenting households by quintiles to see how health, population, and nutrition vary with socio-economic status
- Monitoring (via exit surveys) how well local health-service posts reach the poor
- Measuring coverage of health services via local, small-scale surveys

The first goal is akin to targeting, and the last two goals resemble the monitoring goals here, so the uses of the index are similar to those of the scorecard here.

Still, the Gwatkin *et al.* index is more difficult and costly: it has 13 indicators, one indicator required computing a ratio, it does not fit on a single page, and it cannot

---

by PCA indices correspond with rankings by expenditure-based scorecards include Lindelow (2006), Wagstaff and Watanabe (2003), and Montgomery *et al.* (2000).

be computed by hand in the field. Finally, it has 86 point values, half of them negative, and all with five decimal places.

Unlike the PCA index, the scorecard here is linked directly to an absolute, income-based poverty line. Thus, while both approaches can rank households, only the poverty scorecard can estimate income-based poverty status.

In essence, Gwatkin *et al.*—like all PCA asset indices—define poverty in terms of the indicators in their index. Thus, the index can be seen not as a proxy standing in for something else (such as income) but rather as a direct measure of a non-income-based definition of poverty. There is nothing wrong—and a lot right—about defining poverty in this way, but it is not as common as an income-based definition.

9.2 Vyas and Kumaranayake

Vyas and Kumaranayake (2006) is billed as a "how-to" primer on PCA indices. As a running example, they use urban and rural indices based on Brazil's 1996 DHS (the same data as Gwatkin *et al.*). Vyas and Kumaranayake's indicators resemble those here and in Gwatkin *et al.* (2007) in that they are few, simple, and verifiable:

- Characteristics of the residence:
  - — Presence of electricity
  - — Type of floor
  - — Source of drinking water
  - — Type of toilet facility
  - — Number of rooms for sleeping
- Asset ownership:
  - — Radio
  - — Television
  - — Refrigerator
  - — Telephone

— Bicycle
— Car

As usual, Vyas and Kumaranayake can only assume that the indices represent economic status. Indeed, they do not relate their index to anything (Gwatkin *et al.* relate their index to a series of quantitative measures of health). Vyas and Kumaranayake do not present ready-to-use indices. Moreover, the issues that they are concerned about as possible limitations of PCA-based asset indices—that indicators are selected *ad hoc* and rely on "face validity", that many households are clumped in a few score ranges, that scores for many households may be truncated at the lowest or highest scores, that categorical indicators are not appropriate with PCA, that many indicators are highly correlated with each other, that urban and rural areas are too different to be covered together with a single index—are in fact non-issues. They are made moot by the flat maximum, reflect reality, or do not make a material difference in practice.

9.3 Barros and Victora

Barros and Victora (2005) use a 1-percent sample (418,348 households) from the 2000 Brazilian Demographic Census to build a user-friendly PCA index for urban areas to enable comparisons of the poverty of small groups with larger reference groups. Their strategy is similar to that in this paper: "The number of [indicators] to compose the scorecard is kept manageable for small-scale surveys and epidemiological studies. The 13 [indicators] that compose the scorecard are straightforward to collect and code" (p. 6). PCA coefficients are transformed into points that are non-negative integers. Finally, Barros and Victora present reference distributions by score decile for Brazil as a whole, its 27 states, and its five regions. In short, they take care to make a tool that is useful to non-specialists.

The 13 indicators in Barros and Victora are simple, inexpensive, and verifiable:

- Characteristics of the residence:
  — Number of bedrooms
  — Number of bathrooms
  — Presence of air conditioning
- Ownership of consumer durables:
  — Presence of a radio
  — Number of televisions
  — Presence of a videocassette
  — Presence of a refrigerator or freezer
  — Presence of a washing machine
  — Presence of a microwave oven
  — Presence of a land-line telephone
  — Presence of a microcomputer
  — Number of cars
- Schooling of the head

Scores range from 20 (most poor) to 1,086 (least poor). Barros and Victora compare the distribution of their index to the distribution of income, and find a Pearson correlation coefficient of 0.67.

To illustrate the use of the index, Barros and Victora apply it to 3,000 households in the Family Health Program in the city of Porto Alegre, the capital of one of Brazil's best-off states, Rio Grande do Sul. The index shows that participant families are disproportionately more poor than other households in Porto Alegre, but they are disproportionately less poor than other households in Brazil overall.

On the whole, Barros and Victora is an excellent example of a practical PCA-based asset index, designed and presented to be useful and feasible for non-specialists. Its only technical limitations are applying only to rural areas and constraining the values of some variables (schooling of head, number of bedrooms, bathrooms, televisions, and cars) to have a linear relationship with the score (a mistake highlighted by Vyas and Kumaranayake), when the scorecard here shows that they may be non-linear.

9.4 Filmer and Scott

Filmer and Scott (2008) test how well different approaches to constructing asset indices produce ranks that correlate with ranks from other asset indices, with income as directly measured by a survey, and with income as predicted by a regression. They run tests on 11 countries, one of which is Brazil.

Filmer and Scott find that different approaches to constructing asset indices generally lead to similar rankings vis-à-vis the benchmarks of directly measured income and regression-predicted income. Furthermore, this result is most robust in countries where regression works well for predicting income and in less-poor countries where most income is not spent on food.

For Brazil, Filmer and Scott use data on the 4,940 households in the 1996/7 *Pesquisa sobre Padrões de Vida* (PPV, Living Standards Survey) to select 28 indicators that are simple, low-cost, and verifiable:

- Characteristics of the residence:
  - — Type of structure
  - — Type of floors
  - — Type of walls
  - — Type of roof
  - — Type of toilet arrangement
  - — Rooms per person
- Ownership of consumer durables:
  - — Radio
  - — Tape recorder
  - — Audio system
  - — Television
  - — Video player
  - — Personal computer
  - — Fan
  - — Iron
  - — Washing machine
  - — Clothes dryer
  - — Dishwasher
  - — Vacuum cleaner
  - — Sewing machine
  - — Floor-waxing machine
  - — Air conditioner
  - — Bicycle
  - — Motorbike
  - — Car
  - — Stove
  - — Blender
  - — Microwave oven
  - — Freezer

As Filmer and Scott's goal is to establish general properties of approaches to constructing asset indices (rather than provide asset indices that local, pro-poor organizations can use), they do not report scorecard points.

9.5 Cosenza Faria, Britz do Nascimiento Silva, and Aparecida Feijó

Cosenza Faria, Britz do Nascimiento Silva, and Aparecida Feijó ("CFBNSAF", 2007) use Logit, the 2003 PNAD, and the half-minimum-wage poverty line to construct a scorecard that estimates poverty likelihoods. They focus on targeting, and they propose their scorecard as an alternative to the use of self-reported (and unverified) income to select beneficiaries for the *Cadastro Único* program.

CFBNSAF use 14 indicators that are mostly simple, low-cost, and verifiable:

- Household demographics:
    - Family structure
    - Number of children 14-years-old or younger
    - Age of head
    - Dependency ratio
- Years of schooling of the head
- Characteristics of the residence:
    - Presence of electricity
    - Use of resilient construction materials
    - Source of drinking water
    - Type of sewerage arrangement
    - Type of disposal of garbage
    - Persons per room
- Ownership of land-line telephone
- Place of residence:
    - Region
    - Urban/rural

Conspicuous in its absence is the ownership of consumer durables other than a land-line telephone. Also, the indicators for the dependency ratio and persons per room require computing ratios. Although CFBNSAF report Logit coefficients, non-specialists in local, pro-poor organizations may puzzle over how to use them with the Logit formula to estimate poverty likelihoods.

Targeting accuracy for CFBNSAF can be compared with that of the new scorecard here. The comparison favors CFBNSAF because their tests are *in-sample*, that is, they check accuracy with the same data that is used to construct the scorecard in the first place. In-sample tests overstate accuracy. In contrast, this paper reports only *out-of-sample* tests with data that is not used to construct the scorecard. Johanssen (2006, for BPAC) and Copestake *et al.* (2005, for a variety of measures) find that accuracy measures can deteriorate 8 to 17 percent going from in-sample to out-of-sample. Out-of-sample is also more relevant because, in practice, scorecards are applied out-of-sample to data on households that were not used to construct the scorecard.

CFBNSAF use the 2003 PNAD with a poverty line that leads to a household-level poverty rate of 22 percent. For the new scorecard here with the 2008 PNAD, the $3.75/day 2005 PPP line produces a similar household-level poverty rate of 20.1 percent (Figure 2). Thus, this is the relevant line for the comparison.

For a targeting cut-off that leads to 33 percent of households being targeted, CFBNSAF have inclusion of 17.2 percent and exclusion of 61.6 percent. For a cut-off of 39 or less with the $3.75/day 2005 PPP line (Figure 11), the new scorecard here has inclusion of 17.3 percent (matching CFBNSAF) and exclusion of 77.6 percent (sounding beating CFBNSAF), even though it uses out-of-sample tests.

For a targeting cut-off that leads to 52 percent of households being targeted, CFBNSAF obtain inclusion of 19.8 percent and exclusion of 46.8 percent. For a cut-off of 59 or less with the $3.75/day 2005 PPP line (Figure 11), the new scorecard here has inclusion of 19.9 percent (matching CFBNSAF) and exclusion of 45.4 percent. Given sampling variation and the disadvantage of out-of-sample testing, this suggests that the two scorecards have similar targeting accuracy at this level of inclusion.

9.6 Elbers *et al.*

Elbers *et al.* (2004) construct 10 regional poverty scorecards using direct measures of expenditure from about 5,000 households in the 1996/7 PPV, a poverty line of BRL2.14 per person per day in units of São Paulo prices in 1996, and regressions on the logarithm of per-capita expenditure. The regressions use only indicators available in both the PPV and in the PNAD or other databases of community-level indicators. These scorecards are then applied to the 111,000 households in the Northeast and Southeast regions from the 1996 and 1997 PNAD to construct a "poverty map" (Elbers, Lanjouw, and Lanjouw, 2003) that features consumption-based estimates of poverty rates and measures of inequality at more disaggregated levels and with greater precision than would be possible with the PPV alone. Previous poverty mapping in Brazil has had a record of widespread influence on public opinion that has nudged social policy to be more explicit, transparent, and pro-poor (Snel and Henninger, 2002).

Poverty mapping in Elbers *et al.* and poverty scoring in this paper are similar in that they both:

- Build scorecards with nationally representative survey data and then apply them to other data on sub-groups that may not be nationally representative
- Use simple, verifiable indicators that are quick and inexpensive to collect
- Provide unbiased estimates when their assumptions hold
- Are used to estimate poverty rates for groups
- Seek to be useful in practice and so aim to be understood by non-specialists

Strengths of poverty mapping include that it:

- Has formally established theoretical properties
- Can be applied straightforwardly to measures of well-being beyond poverty rates
- Requires less data for scorecard construction and calibration
- Includes community-level indicators
- Uses only indicators that appear in a census

Strengths of poverty scoring include that it:

- Is simpler in terms of both construction and application
- Tests accuracy empirically
- Associates poverty likelihoods with scores non-parametrically
- Uses judgment and theory in scorecard construction to reduce overfitting
- Estimates poverty likelihoods for individual households
- Reports simple formulas for standard errors

The basic difference between the two approaches is that poverty mapping seeks to help governments design pro-poor policies, while poverty scoring seeks to help small, local pro-poor organizations to manage their outreach when implementing policies.[18] For

[18] Another apparent difference is that the developers of the poverty-mapping approach (Elbers, Lanjouw, and Lanjouw, 2003; Demombynes *et al.*, 2002) say that it is too inaccurate to be used for targeting particular households, while Schreiner (2008c) supports such targeting as a legitimate, potentially useful application of poverty scoring. Recently, the developers of poverty mapping seem to have taken some small steps away from their original position (Elbers *et al.*, 2007).

example, Elbers *et al.* have as one of their two explicit purposes the demonstration of their method, and they do not report scorecard indicators or points at all.

Elbers *et al.* compare direct estimates of poverty rates from the PPV with indirect estimates from their poverty scorecards applied to the PNAD, finding a close correspondence. Because the PNAD does not measure expenditure directly, however, out-of-sample measures of bias are not possible. Likewise, while Elbers *et al.* report standard errors for estimated poverty rates, they do not report sample sizes nor standard-error formula, so a comparison with the precision of the new poverty scorecard here is not possible.

9.7 Tarozzi and Deaton, and Elbers, Lanjouw, and Leite

A general debate on the accuracy of the poverty-mapping approach (and by extension, the poverty-scoring approach here) has played out against the background of Brazil in Tarozzi and Deaton (2007) and Elbers, Lanjouw, and Leite (2008).

9.7.1 Tarozzi and Deaton

The starting point of the debate is Tarozzi and Deaton ("TD", 2007). They point out that sub-groups in a population (such as a given municipality, or participants of a given pro-poor organization) may differ from the population as a whole in ways that are both linked with poverty and not fully captured by a poverty map or scorecard. These differences cause estimates based on poverty mapping (or poverty scoring) to differ from true values. For this reason, TD say that reports of accuracy should include not only standard errors but also differences from true values. Differences in true values—always reported in the poverty-scoring papers by the present author—have not been reported for poverty mapping, simply because the true values are unknown.

TD use Monte Carlo tests to demonstrate that sub-group differences can matter. To show that their concern is not merely theoretical, they use the 2000 Mexico census to create synthetic household surveys of rural households in Chiapas, Oaxaca, and Veracruz. They apply poverty mapping to these surveys, generate estimates of poverty rates, and compare the estimates out-of-sample to census data on households not used in scorecard construction. At the time, this was the first such an accuracy test for poverty mapping, although, as in this paper, such test have always been standard in the poverty-scoring approach.[19]

9.7.2 Elbers, Lanjouw, and Leite

Elbers, Lanjouw, and Leite ("ELL", 2008) defend the poverty-mapping approach against TD's critique. They use data from one state in Brazil's 2000 Demographic Census, which collected "fairly detailed" (p. 4) income data from one of eight households (606,000 households in the 853 municipalities of the state of Minas Gerais). ELL draw 20 synthetic surveys from the census data that replicate the sampling design and the enumeration areas of the PNAD (averaging about 12,000 households each), 20 synthetic

[19] Tarozzi (2008) further shows that the poverty-mapping approach, when applied to literacy rates in the 2000 Mexico census, leads to inaccuracies for sub-groups, suggesting that there probably are also inaccuracies when applying the approach to income or poverty rates.

surveys that replicate the sampling design and the enumeration areas of the POF (averaging about 2,800 households each), and 1 synthetic survey with new "enumeration areas" following the sampling design of the POF.

ELL use the standard poverty-mapping approach—stepwise regression on the logarithm of per-capita household income—with a poverty line of BRL3.29 per person per day to construct one poverty scorecard for each of the 41 pseudo-surveys. They then apply the scorecards out-of-sample to census data, comparing estimates to true values for poverty rates and for measures of inequality.

The 41 models have 17 to 45 indicators. These indicators include both household-level indicators (like the poverty scorecard here) as well as higher-level aggregate indicators derived from the census or ancillary sources (unlike the poverty scorecard here). ELL say that the higher-level aggregate indicators are critical to keep bias low enough (and precision high enough) to mitigate TD's critique.

ELL report 26 indicators and regression coefficients for a single example scorecard based on a synthetic PNAD:

- Demographics:
    - — Race/ethnicity of head
    - — Whether the household head is female
    - — Age of head
    - — Number of household members 66-years-old or older
    - — Family composition
- Education:
    - — Years of schooling of head
    - — School attendance of children
- Characteristics of the residence:
    - — Location on a paved street
    - — Tenancy status

— Type of structure
— Number of rooms
— Source of drinking water
— Type of toilet arrangement
— Type of sewage disposal
- Ownership of consumer durables:
  — Washing machine
  — Refrigerator
  — Microwave oven
- Municipal characteristics:
  — Percentage of households with curb-side garbage collection
  — Share of population who are migrants
- District characteristics:
  — Average years of schooling in the population
  — Share of people who are out of the labor force
  — Share of people who are self-employed
  — Average income per-capita
- Enumeration-area characteristics:
  — Average income of household heads
  — Average years of schooling of household heads
  — Average number of members per household

After applying this and the other scorecards out-of-sample to census data and comparing estimates to true values, ELL acknowledge that Deaton and Tarozzi have a point, but they conclude that the poverty-mapping approach "performs reasonably well" and "is able to produce estimates of welfare that line up quite closely to their true values . . . [and] confidence intervals for the poverty estimates also appear to be appropriate. However, this latter conclusion holds only after carefully controlling for community-level factors that are correlated with household-level welfare."

Similar conclusions come out of another paper with a similar test with data from Mexico (Demombynes, Elbers, and Lanjouw, 2006). It also concedes that accuracy is reduced when a sub-group is not representative of the population from which the scorecard is built, and is also contends that the use of community-level indicators mitigates such inaccuracies. After all, if a sub-group is different, then group-level indicators should help control for these differences. Demombynes, Elbers, and Lanjouw (2006) conclude that "bias is low" (p. 18) and that the use of community-level indicators "can go a long way" (p. 19) toward mitigating sub-group differences.

Demombynes, Elbers, and Lanjouw (2006) report that the average difference between estimated poverty rates and true values across 20 "small areas" with an average sample size of about 1,010 is +0.7 percentage points, and the average 90-percent confidence interval for this difference is +/–0.7 percentage points.

Even though they share two authors with Demombynes, Elbers, and Lanjouw (2006), ELL never report these two simple summary measures of bias and precision across their synthetic surveys and applications to census data for Brazil. Instead, they present a graph of estimated poverty rates versus true rates at the municipal level for two of the 41 synthetic surveys, saying that "the figures demonstrate that poverty estimates are randomly assigned around the main diagonal" and that the correlation between estimates and true values ranges from 75 to 90 percent, depending on the survey and the poverty measure.

But the question remains: if the poverty-mapping estimate of poverty rates is unbiased (or if bias is low), then why don't ELL come right out and say that? Why do they fail to report an average figure for bias, relying instead on graphs in which bias is impossible to discern? After all, reporting bias is precisely TD's main recommendation.

In terms of precision, ELL report that about 90 percent of 95-percent confidence intervals at the municipal level contain the true value. This confirms the TD critique that confidence intervals are too wide in poverty mapping, but ELL argue that the confidence intervals are close enough and "do not appear unreasonable—particularly for [estimates of the poverty rate]."

In essence, ELL argue that poverty mapping's confidence intervals may be too big, but not so big that the TD critique matters much in practice. Fair enough. But by what standard are the confidence intervals "close enough" (McCloskey, 1998)? The standard for ELL is that in practice "a hypothetical policy maker, presented with [a poverty map] and its accompanying standard errors, would not come away with a wildly unrealistic picture of the spatial distribution of poverty". For targeting social programs, this seems like a reasonable standard.

9.7.3 What this means for poverty scoring

In some senses, poverty scoring is a simpler version of poverty mapping, designed to be accurate enough to be useful, inexpensive enough to be used by local pro-poor organizations, and straightforward enough for non-specialists to understand and accept. Given this, does the critique of poverty mapping by Deaton—a possible future Noble Prize winner in economics—mean that poverty scoring should be abandoned?

In their concluding remarks, TD say (pp. 24–25):

> Overall, we believe that efforts to calculate welfare estimates for small areas . . . are certainly worthwhile, but we also believe that the current literature has not emphasized enough the limitations of the current methodologies and the very strong assumptions that they require in order

> to allow for meaningful inference. Such limitations must be stressed, and the precision of the estimates should be judged accordingly . . . [Users] should be aware that such maps may be subject to much more uncertainty and error than previously thought.

In essence, Tarozzi and Deaton ask that the authors of poverty maps document not only standard errors but also differences between estimates and true values (that is, bias), as well as the broader limitations of the approach. This is reasonable; users of any tool need to know what the tool can and cannot do, and in what contexts.

This type of reporting has been standard for poverty scoring since the beginning of 2008. In particular, reporting includes both bias and standard errors and explicitly points out that reported accuracy holds only for sub-groups that are representative of a given country's population at a particular point in time. Thus, poverty scoring is still potentially useful, even though poverty mapping's limitations were not fully reported.

But is poverty mapping/scoring accurate enough? The answer depends on the purpose. Consider, for example, for-profit lenders with billions of dollars at risk in loans underwritten largely via credit-risk scorecards. Not only are these credit-risk scorecards much less accurate for their purposes than poverty maps/scorecards, but they are also subject to the same sub-group critiques. But even though credit-risk scorecards have limitations, they are nevertheless more useful than alternatives from a benefit/cost perspective.

The next question is then, What is the benefit of improved decisions versus the cost of improved decision-making? If national governments are targeting funds at the state-level, then the cost of poverty mapping/scoring is probably not worth the benefit; after all, governments already can rank states by poverty without any additional information. If, however, federal governments are targeting funds at lower levels, then they may not know what the poorest entities are (although governments at lower levels should know). In any case, poverty mapping provides an objectivity that will likely favor poorer entities in the budget process, raising awareness among the polity and allowing politicians to deflect accusations of political bias by referring to the poverty map.

In the case of local, pro-poor organizations, no alternative for targeting households compares well with poverty scoring's combination of inexpensiveness, accuracy, and objectivity. Other targeting tools may be more accurate, but they cost more, are less objective, and their accuracy is unquantified. For raising organizational awareness about performance in terms of poverty outreach, scoring's measures of poverty rates are also valuable, showing managers which branches and field agents serve poorer people, and whether the pro-poor organization as a whole is indeed pro-poor.

Of course, no tool is a silver bullet, and poverty mapping and poverty scoring are accurate enough for some uses and not accurate enough for others. A central strength of both approaches is their ability to report quantitative measures of accuracy in terms of bias and precision.

In most cases, the errors in poverty scoring are likely to be small, relative to the benefit/cost of additional accuracy. That is, given the alternatives for their purposes, poverty scorecards are usually “good enough for government work”.

TD apply their critique only to estimates of poverty rates. Their critique may not to apply as strongly to estimates of changes in poverty rates or to rankings used for targeting. For example, Schreiner (2006a) finds little degradation for targeting when a single all-Mexico scorecard is applied to urban/rural sub-groups. Of course, on the continuum of sub-groups between urban/rural down to a single household, at some point rankings may very well become too inaccurate for a given purpose. Still, even relatively inaccurate credit-risk scorecards have proven useful for targeting individual households, and, depending on the context and alternatives, poverty scorecards may likewise turn out to be a good choice for targeting as well as other uses.

Given that ELL emphasize that higher-level aggregate indicators improve accuracy, a final question is whether poverty scorecards should also include them. Local pro-poor organizations, however, do not seek estimates of poverty rates for a community; they want measures for their clients (a sub-group within a community or across a group of communities). It would also be difficult to provide all local organizations with the values of the higher-level aggregate indicators.

10. Conclusion

This paper presents a simple poverty scorecard for Brazil that can be used to estimate the likelihood that a household has income below a given poverty line, to estimate the poverty rate of a group of households at a point in time, and to estimate changes in the poverty rate of a group of households between two points in time. The scorecard can also be used for targeting.

The scorecard is inexpensive to use and can be understood by non-specialists. It is designed to be practical for local pro-poor organizations who want to improve how they monitor and manage their social performance.

The scorecard is built with a sub-sample of data from the 2008 PNAD, calibrated to nine poverty lines, and tested on a different sub-sample from the 2008 PNAD.

Accuracy is reported for estimates of households' poverty likelihoods, groups' poverty rates at a point in time, and changes in groups' poverty rates over time. Of course, the scorecard's estimates of changes in poverty rates are not the same as estimates of program impact. Targeting accuracy and formula for standard errors are also reported.

When the scorecard is applied to the validation sample with $n = 16{,}384$, the absolute difference between estimates and true poverty rates at a point in time is 0.5 percentage points or less and averages (across the nine poverty lines) 0.3 percentage points. With 90-percent confidence, the precision of these differences for all lines is +/– 0.5 percentage points or less.

For targeting, programs can use the results reported here to select a cut-off that fits their mission and values.

Although the statistical technique is innovative, and although technical accuracy is important, the design of the poverty scorecard focuses on transparency and ease-of-use. After all, a perfectly accurate scorecard is worthless if programs feel so daunted by its complexity or its cost that they do not even try to use it. For this reason, the poverty scorecard is kept simple, using ten indicators that are inexpensive to collect and that are straightforward to verify. Points are all zeros or positive integers, and scores range from 0 to 100. Scores are related to poverty likelihoods via simple look-up tables, and targeting cut-offs are likewise simple to apply. The design attempts to facilitate adoption by helping managers understand and trust scoring and by allowing non-specialists to generate scores quickly in the field.

In sum, the simple poverty scorecard is a practical, objective way for pro-poor programs in Brazil to monitor poverty rates, track changes in poverty rates over time, and target services, provided that it is applied during a period similar to that of September 2008, the reference period for the data used to construct the scorecard. The same approach can be applied to any country with similar data from a national income or expenditure survey.

References

Adams, Niall M.; and David J. Hand. (2000) "Improving the Practice of Classifier Performance Assessment", *Neural Computation*, Vol. 12, pp. 305–311.

Baesens, Bart; Van Gestel, Tony; Viaene, Stijn; Stepanova, Maria; Suykens, Johan A. K.; and Jan Vanthienen. (2003) "Benchmarking State-of-the-Art Classification Algorithms for Credit Scoring", *Journal of the Operational Research Society*, Vol. 54, pp. 627–635.

Barros, Aluísio J.D.; and Cesar G. Victora. (2005) "A Nationwide Wealth Score Based on the 2000 Brazilian Demographic Census", *Revista de Saudé Pública*, Vol. 39, No. 4, pp. 1–6.

Bollen, Kenneth A.; Glanville, Jennifer L.; and Guy Stecklov. (2007) "Socio-Economic Status, Permanent Income, and Fertility: A Latent-Variable Approach", *Population Studies*, Vol. 61, No. 1, pp. 15–34.

Caire, Dean. (2004) "Building Credit Scorecards for Small Business Lending in Developing Markets", Bannock Consulting, `http://www.microfinance.com/English/Papers/Scoring_SMEs_Hybrid.pdf`, retrieved 1 February 2010.

Chen, Shiyuan; and Mark Schreiner. (2009a) "A Simple Poverty Scorecard for Bangladesh", `http://www.microfinance.com/English/Papers/Scoring_Poverty_Bangladesh.pdf`, retrieved 1 February 2010.

______. (2009b) "A Simple Poverty Scorecard for Vietnam", `http://www.microfinance.com/English/Papers/Scoring_Poverty_Vietnam_EN.pdf`, retrieved 1 February 2010.

Coady, David; Grosh, Margaret; and John Hoddinott. (2002) "The Targeting of Transfers in Developing Countries: Review of Experience and Lessons", `http://info.worldbank.org/etools/docs/library/79646/Dc%202003/courses/dc2003/readings/targeting.pdf`, retrieved 1 February 2010.

Cochran, William G. (1977) *Sampling Techniques, Third Edition*, New York: Wiley.

Copestake, James G.; Dawson, Peter; Fanning, John-Paul; McKay, Andrew; and Katie Wright-Revolledo. (2005) "Monitoring the Diversity of the Poverty Outreach and Impact of Microfinance: A Comparison of Methods Using Data from Peru", *Development Policy Review*, Vol. 23, No. 6, pp. 703–723.

Cosenza Faria, Ana Lucia; Britz do Nascimiento Silva; and Carmem Aparecida Feijó. (2007) "Aplicação do Teste de Elegibilidade Multidimensional na Definição do Público-Alvo Beneficiário de Políticas Públicas", *Revista de Economia Contemporânea*, Vol. 11, No. 2, pp. 253–278, http://www.ie.ufrj.br/revista/pdfs/v11n2-Aplicacao_do_Teste_de_Elegibilidade.pdf, retrieved 3 February 2010.

Daley-Harris, Sam. (2009) *State of the Microcredit Summit Campaign Report 2009*, Washington, D.C., http://www.microcreditsummit.org/state_of_the_campaign_report/, retrieved 1 February 2010.

Dawes, Robyn M. (1979) "The Robust Beauty of Improper Linear Models in Decision Making", *American Psychologist*, Vol. 34, No. 7, pp. 571–582.

Demombynes, Gabriel; Elbers, Chris; and Peter Lanjouw. (2007) "How Good a Map? Putting Small-Area Estimation to the Test", World Bank Policy Research Working Paper No. 4155, Washington, D.C., http://www-wds.worldbank.org/servlet/WDSContentServer/WDSP/IB/2007/03/26/000016406_20070326150728/Rendered/PDF/wps4155.pdf, retrieved 4 February 2010.

_____; Elbers, Chris; Lanjouw, Jenny; Lanjouw, Peter; Mistiaen, Johan; and Berk Özler. (2002) "Producing an Improved Geographic Profile of Poverty: Methodology and Evidence from Three Developing Countries", World Institute for Development Economics Research Discussion Paper No. 2002/39, http://go.worldbank.org/UMQCZ1BW00, retrieved 4 February 2010.

Efron, Bradley; and Robert J. Tibshirani. (1993) *An Introduction to the Bootstrap*, New York: Chapman and Hall.

Elbers, Chris; Fujii, Tomoki; Lanjouw, Peter; Özler, Berk; and Wesley Yin. (2007) "Poverty Alleviation through Geographic Targeting: How Much Does Disaggregation Help?", *Journal of Development Economics*, Vol. 83, pp. 198–213.

_____; Lanjouw, Peter; and Phillippe George Leite. (2008) "Brazil within Brazil: Testing the Poverty Map Methodology in Minas Gerais", Policy Research Working Paper No. 4513, Washington, D.C.: World Bank, http://www-wds.worldbank.org/servlet/WDSContentServer/WDSP/IB/2008/02/26/000158349_20080226134003/Rendered/PDF/wps4513.pdf, retrieved 3 February 2010.

_____; Lanjouw, Jean Olson; and Peter Lanjouw. (2003) "Micro-Level Estimation of Poverty and Inequality", *Econometrica*, Vol. 71, No. 1, pp. 355–364, http://siteresources.worldbank.org/DEC/Resources/micestpovineq.pdf, retrieved 1 February 2010.

_____; Lanjouw, Jean Olson; Lanjouw, Peter; and Phillippe George Leite. (2004) "Poverty and Inequality in Brazil: New Estimates from Combined PPV-PNAD Data", http://unstats.un.org/unsd/methods/poverty/RioWS-poverty-in-Brazil.pdf, retrieved 3 February 2010.

Falkenstein, Eric. (2008) "DefProb™: A Corporate Probability of Default Model", http://www.defprob.com/publications/DefProb.pdf, retrieved 2 December 2009.

Ferguson, Brodie D.; Tandon, Ajay; Gakidou, Emmanuela; and Christopher J.L. Murray. (2003) "Estimating Permanent Income Using Indicator Variables", pp. 747–760 in Christopher J.L. Murray and David B. Evans (eds) *Health Systems Performance Assessment: Debates, Methods, and Empiricism*, Geneva: World Health Organization.

Ferreira Loureiro, André Oliveira; and Daniel Cirilo Suliano. (2009) "As Principais Linhas de Pobreza Utilizadas no Brasil", Nota Técnica No. 38, Fortaleza: Instituto de Pesquisa e Estratégia Econômica do Ceará, http://www.ipece.ce.gov.br/publicacoes/notas_tecnicas/NT_38.pdf, retrieved 4 February 2010.

Filmer, Deon; and Lant Pritchett. (2001) "Estimating Wealth Effects without Expenditure Data—or Tears: An Application to Educational Enrollments in States of India", *Demography*, Vol. 38, No. 1, pp. 115–132.

Filmer, Deon; and Kinnon Scott. (2008) "Assessing Asset Indices", Policy Research Working Paper No. 4605, Washington, D.C.: World Bank, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1149108, retrieved 1 February 2010.

Friedman, Jerome H. (1997) “On Bias, Variance, 0–1 Loss, and the Curse-of-Dimensionality”, *Data Mining and Knowledge Discovery*, Vol. 1, pp. 55–77.

Fuller, Rob. (2006) “Measuring the Poverty of Microfinance Clients in Haiti”, `http://www.microfinance.com/English/Papers/Scoring_Poverty_Haiti_Fuller.pdf`, retrieved 1 February 2010.

Goodman, Leo A.; and Kruskal, William H. (1979) *Measures of Association for Cross Classification*, New York, NY: Springer-Verlag.

Grootaert, Christiaan; and Jeanine Braithwaite. (1998) “Poverty Correlates and Indicator-Based Targeting in Eastern Europe and the Former Soviet Union”, Policy Research Working Paper No. 1942, Washington, D.C.: World Bank, `http://go.worldbank.org/VPMWVLU8E0`, retrieved 1 February 2010.

Grosh, Margaret; and Judy L. Baker. (1995) “Proxy Means Tests for Targeting Social Programs: Simulations and Speculation”, LSMS Working Paper No. 118, Washington, D.C.: World Bank, `http://go.worldbank.org/W9OWN57PD0`, retrieved 1 February 2010.

Gwatkin, Davidson R.; Rutstein, Shea; Johnson, Kiersten; Suliman, Eldaw; Wagstaff, Adam; and Agbessi Amouzou. (2007) “Socio-Economic Differences in Health, Nutrition, and Population: Brazil”, Country Reports on HNP and Poverty, Washington, D.C.: World Bank, `http://go.worldbank.org/T6LCN5A340`, retrieved 3 February 2010.

Hand, David J. (2006) “Classifier Technology and the Illusion of Progress”, *Statistical Science*, Vol. 22, No. 1, pp. 1–15.

Hoadley, Bruce; and Robert M. Oliver. (1998) “Business Measures of Scorecard Benefit”, *IMA Journal of Mathematics Applied in Business and Industry*, Vol. 9, pp. 55–64.

IRIS Center. (2007a) “Manual for the Implementation of USAID Poverty Assessment Tools”, `http://www.povertytools.org/training_documents/Manuals/USAID_PAT_Manual_Eng.pdf`, retrieved 1 February 2010.

_____. (2007b) “Introduction to Sampling for the Implementation of PATs”, `http://www.povertytools.org/training_documents/Sampling/Introduction_Sampling.ppt`, retrieved 1 February 2010.

_____. (2005) “Notes on Assessment and Improvement of Tool Accuracy”, `http://www.povertytools.org/other_documents/AssessingImproving_Accuracy.pdf`, retrieved 1 February 2010.

Johannsen, Julia. (2006) “Operational Poverty Targeting in Peru—Proxy Means Testing with Non-Income Indicators”, International Poverty Centre Working Paper No. 30, `http://www.undp-povertycentre.org/pub/IPCWorkingPaper30.pdf`, retrieved 4 February 2010.

Johnson, Glenn. (2007) “Lesson 3: Two-Way Tables—Dependent Samples”, `http://www.stat.psu.edu/online/development/stat504/03_2way/53_2way_compare.htm`, retrieved 1 February 2010.

Kolesar, Peter; and Janet L. Showers. (1985) “A Robust Credit-Screening Model Using Categorical Data”, *Management Science*, Vol. 31, No. 2, pp. 124–133.

Lindelow, Magnus. (2006) “Sometimes More Equal Than Others: How Health Inequalities Depend on the Choice of Welfare Indicator”, *Health Economics*, Vol. 15, pp. 263–279.

Lovie, Alexander D.; and Patricia Lovie. (1986) “The Flat-Maximum Effect and Linear Scoring Models for Prediction”, *Journal of Forecasting*, Vol. 5, pp. 159–168.

Martinelli, César; and Susan W. Parker. (2007) “Deception and Misreporting in a Social Program”, Centro de Investigación Económica and Instituto Tecnológico Autónomo de México, `http://ciep.itam.mx/~martinel/lies4.pdf`, retrieved 1 February 2010.

Matul, Michal; and Sean Kline. (2003) “Scoring Change: Prizma’s Approach to Assessing Poverty”, MFC Spotlight Note No. 4, Warsaw, Poland: Microfinance Centre for Central and Eastern Europe and the New Independent States, `http://www.mfc.org.pl/doc/Research/ImpAct/SN/MFC_SN04_eng.pdf`, retrieved 1 February 2010.

McCloskey, Deirdre N. (1998) *The Rhetoric of Economics, Second Edition*, Madison: University of Wisconsin Press.

McNemar, Quinn. (1947) “Note on the Sampling Error of the Difference between Correlated Proportions or Percentages”, *Psychometrika*, Vol. 17, pp. 153–157.

Microfinance Risk Management, L.L.C. (2010) "Data-Entry Software for a Simple Poverty Scorecard for Brazil", http://www.microfinance.com/#Brazil, retrieved 1 February 2010.

Montgomery, Mark; Gragnolati, Michele; Burke, Kathleen A.; and Edmundo Paredes. (2000) "Measuring Living Standards with Proxy Variables", *Demography*, Vol. 37, No. 2, pp. 155–174.

Myers, James H.; and Edward W. Forgy. (1963) "The Development of Numerical Credit Evaluation Systems", *Journal of the American Statistical Association*, Vol. 58, No. 303, pp. 779–806.

Narayan, Ambar; and Nobuo Yoshida. (2005) "Proxy Means Tests for Targeting Welfare Benefits in Sri Lanka", Report No. SASPR–7, Washington, D.C.: World Bank, http://siteresources.worldbank.org/EXTSAREGTOPPOVRED/Resources/493440-1102216396155/572861-1102221461685/Proxy+Means+Test+for+Targeting+Welfare+Benefits.pdf, retrieved 1 February 2010.

Onwujekwe, Obinna; Hanson, Kara; and Julia Fox-Rushby. (2006) "Some Indicators of Socio-Economic Status May Not Be Reliable and Use of Indices with These Data Could Worsen Equity", *Health Economics*, Vol. 15, pp. 639–644.

Rutstein, Shea Oscar; and Kiersten Johnson. (2004) "The DHS Wealth Index", DHS Comparative Reports No. 6, Calverton, MD: ORC Macro, http://www.measuredhs.com/pubs/pdf/CR6/CR6.pdf, retrieved 1 February 2010.

Sahn, David E.; and David Stifel. (2003) "Exploring Alternative Measures of Welfare in the Absence of Expenditure Data", *Review of Income and Wealth*, Series 49, No. 4, pp. 463–489.

_____. (2000) "Poverty Comparisons over Time and across Countries in Africa", *World Development*, Vol. 28, No. 12, pp. 2123–2155.

SAS Institute Inc. (2004) "The LOGISTIC Procedure: Rank Correlation of Observed Responses and Predicted Probabilities", in *SAS/STAT User's Guide, Version 9*, Cary, NC, http://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/statug_logistic_sect035.htm, retrieved 1 February 2010.

Schreiner, Mark. (2009a) "A Simple Poverty Scorecard for Peru", http://www.microfinance.com/English/Papers/Scoring_Poverty_Peru.pdf, retrieved 1 February 2010.

_____. (2009b) "A Simple Poverty Scorecard for the Philippines", http://www.microfinance.com/English/Papers/Scoring_Poverty_Philippines.pdf, retrieved 1 February 2010.

_____. (2009c) "A Simple Poverty Scorecard for Pakistan", http://www.microfinance.com/English/Papers/Scoring_Poverty_Pakistan_2005.pdf, retrieved 1 February 2010.

_____. (2009d) "A Simple Poverty Scorecard for Bolivia", http://www.microfinance.com/English/Papers/Scoring_Poverty_Bolivia_EN_2007.pdf, retrieved 1 February 2010.

_____. (2009e) "A Simple Poverty Scorecard for Mexico", http://www.microfinance.com/English/Papers/Scoring_Poverty_Mexico_2008_EN.pdf, retrieved 1 February 2010.

_____. (2008a) "A Simple Poverty Scorecard for Peru", http://www.microfinance.com/English/Papers/Scoring_Poverty_Peru_May_2008.pdf, retrieved 1 February 2010.

_____. (2008b) "A Simple Poverty Scorecard for India", http://www.microfinance.com/English/Papers/Scoring_Poverty_India.pdf, retrieved 1 February 2010.

_____. (2008c) "A Simple Poverty Scorecard for Ecuador", http://www.microfinance.com/English/Papers/Scoring_Poverty_Ecuador.pdf, retrieved 3 February 2010.

_____. (2006a) "Un índice de pobreza para México", memo for Grameen Foundation, http://www.microfinance.com/Castellano/Documentos/Scoring_Pobreza_Mexico_2002.pdf, retrieved 1 February 2010.

_____. (2006b) "Is One Simple Poverty Scorecard Enough for India?", memo for Grameen Foundation, http://www.microfinance.com/English/Papers/Scoring_Poverty_India_Segments.pdf, retrieved 1 February 2010.

_____. (2005) "IRIS Questions on Poverty Scorecards", memo for Grameen Foundation, http://www.microfinance.com/English/Papers/Scoring_Poverty_Response_to_IRIS.pdf, retrieved 1 February 2010.

_____. (2002) *Scoring: The Next Breakthrough in Microfinance?* Occasional Paper No. 7, Consultative Group to Assist the Poor, Washington, D.C., http://pdf.usaid.gov/pdf_docs/PNACQ633.pdf, retrieved 1 February 2010.

_____; Matul, Michal; Pawlak, Ewa; and Sean Kline. (2004) "Poverty Scorecards: Lessons from a Microlender in Bosnia-Herzegovina", Microfinance Risk Management, L.L.C., http://www.microfinance.com/English/Papers/Scoring_Poverty_in_BiH_Short.pdf, retrieved 1 February 2010.

Sillers, Don. (2006) "National and International Poverty Lines: An Overview", Washington, D.C.: United States Agency for International Development, http://www.povertytools.org/other_documents/siller-povertylines.doc, retrieved 1 February 2010.

Snel, Mathilde; and Norbert Henninger. (2002) "Where are the Poor? Experiences with the Use and Development of Poverty Maps", Washington, D.C.: World Resources Institute, http://pdf.wri.org/wherepoor.pdf, retrieved 4 February 2009.

Stillwell, William G.; Barron, F. Hutton; and Ward Edwards. (1983) "Evaluating Credit Applications: A Validation of Multi-Attribute Utility-Weight Elicitation Techniques", *Organizational Behavior and Human Performance*, Vol. 32, pp. 87–108.

Tarozzi, Alesandro. (2008) "Can Census Data Alone Signal Heterogeneity in the Estimation of Poverty Maps?", http://www.econ.duke.edu/~taroz/TarozziHet2008.pdf, retrieved 4 February 2010.

_____; and Angus Deaton. (2007) "Using Census and Survey Data to Estimate Poverty and Inequality for Small Areas", http://www.princeton.edu/~deaton/downloads/20080301SmallAreas_FINAL.pdf, retrieved 1 February 2010.

Toohig, Jeff. (2008) "PPI Pilot Training Guide", Grameen Foundation, http://www.progressoutofpoverty.org/toolkit, retrieved 1 February 2010.

United States Congress. (2002) "Amendments to the Microenterprise for Self-Reliance Act of 2000 (Public Law 106–309)", October 8, http://www.microlinks.org/file_download.php/AmendMicroenterpriseAct2000.pdf?URL_ID=7744&filename=11205460851AmendMicroenterpriseAct2000.pdf&filetype=application%2Fpdf&filesize=95834&name=AmendMicroenterpriseAct2000.pdf&location=user-S/, retrieved 1 February 2010.

Wagstaff, Adam; and Naoko Watanabe. (2003) "What Difference Does the Choice of SES Make in Health-Inequality Measurement?", *Health Economics*, Vol. 12, No. 10, pp. 885–890.

Wainer, Howard. (1976) "Estimating Coefficients in Linear Models: It Don't Make No Nevermind", *Psychological Bulletin*, Vol. 83, pp. 223–227.

World Bank. (2007) *Brazil: Measuring Poverty Using Household Consumption*, Report No. 36358-BR, Washington, D.C., http://go.worldbank.org/U1KR0U6L80, retrieved 4 February 2010.

______. (2008) "International Comparison Project: Tables of Results", Washington, D.C., http://siteresources.worldbank.org/ICPINT/Resources/icp-final-tables.pdf, retrieved 1 February 2010.

Vyas, Seema; and Lilani Kumaranayake. (2006) "Constructing Socio-Economic Status Indices: How to Use Principal Components Analysis", *Health Policy Planning*, Vol. 21, No. 6, pp. 459–468.

Zeller, Manfred. (2004) "Review of Poverty Assessment Tools", Accelerated Microenterprise Advancement Project, http://www.povertytools.org/other_documents/Review%20of%20PAT%20Tools.pdf, retrieved 1 February 2010.

Figure 2: Sample sizes and poverty rates, by sub-sample and poverty line

			% with expenditure below a poverty line								
			National (minimum wages)				USAID	International 2005 PPP			
Sub-sample	**Item**	**Households**	**1/2**	**1/4**	**1**	**2**	**'Extreme'**	**$1.25/day**	**$2.50/day**	**$3.75/day**	**$5.00/day**
All Brazil	Poverty line (BRL/person/day)		6.90	3.43	13.83	27.67	4.27	2.18	4.35	6.53	8.70
	Poverty rate (household level)	109,640	23.8	8.5	51.6	77.7	11.3	4.3	11.6	20.1	30.3
	Poverty rate (person level)	109,640	30.3	11.2	58.5	82.2	15.2	5.2	15.6	26.7	38.0
Construction											
Selecting indicators and weights	Poverty rate (household level)	36,439	23.9	8.5	51.5	77.7	11.3	4.3	11.5	20.1	30.2
Calibration											
Associating scores with likelihoods	Poverty rate (household level)	36,409	23.8	8.5	51.4	77.7	11.3	4.2	11.6	20.2	30.2
Validation											
Measuring accuracy	Poverty rate (household level)	36,792	23.8	8.5	51.8	77.6	11.3	4.3	11.6	20.1	30.4
Change in household-level poverty rate (percentage points)											
From construction/calibration to validation			+0.0	–0.0	–0.4	+0.1	–0.1	–0.1	–0.1	+0.1	–0.3

Source: 2008 PNAD

Figure 3: Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (responses are ordered starting with those associated with higher poverty likelihoods)
1536	What is the most years of schooling that a member of the household has completed? (Five to seven; One to four; Eight to ten; None or less than one year; Eleven; Twelve to fourteen; Fifteen or more)
1368	Does the household have a cellular or land-line telephone? (None; Cellular but not land-line; Land-line but not cellular; Both)
1297	Does the household have a personal computer? (No; Yes, but no internet connection; Yes, and has an internet connection)
1204	Does the household have a washing machine? (No; Yes)
1189	How many household members are 17-years-old or younger? (Three or more; Two; One; None)
1186	How many household members are 18-years-old or younger? (Three or more; Two; One; None)
1170	How many household members are 16-years-old or younger? (Three or more; Two; One; None)
1150	Do any household members ages 5 to 18 go to private school or private pre-school? (No; Yes; No members ages 5 to 18)
1139	Do any household members ages 5 to 17 go to private school or private pre-school? (No; Yes; No members ages 5 to 17)
1120	Do any household members ages 5 to 16 go to private school or private pre-school? (No; Yes; No members ages 5 to 16)
1118	In their main occupation, how many household members are managers, administrators, professionals in the arts and sciences, mid-level technicians, or clerks? (None; One or more)
1079	Do any household members ages 5 to 15 go to private school or private pre-school? (No; Yes; No members ages 5 to 15)
1059	How many years of schooling has the male head/spouse completed? (One or less; Two to three; Five to six; Four or seven; No male head/spouse; Eight to ten; Eleven to twelve; Thirteen or more)
1044	How many household members are 15-years-old or younger? (Two or more; One; None)

Figure 3 (cont.): Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (responses are ordered starting with those associated with higher poverty likelihoods)
1033	Do any household members ages 5 to 14 go to private school or private pre-school? (No; Yes; No members ages 5 to 14)
1022	Is there a bathroom or water closet in the residence or on the property? (No, or yes, but shared with other households; Yes, one, not shared; Yes, two, not shared)
1004	How many household members are 14-years-old or younger? (Two or more; One; None)
971	Do any household members ages 5 to 13 go to private school or private pre-school? (No; Yes; No members ages 5 to 13)
957	How many household members are 13-years-old or younger? (Two or more; One; None)
912	How many household members are 12-years-old or younger? (Two or more; One; None)
911	What is the occupation of the female head/spouse in her main line of work? (Farmer, rancher, or undefined; Does not work; Service worker; Sales person or shopkeeper; Skilled worker in production, maintenance, and repair; No female head/spouse; Manager, administrator, professional in the arts and sciences, mid-level technician, or clerk)
909	Do any household members ages 5 to 12 go to private school or private pre-school? (No; Yes; No members ages 5 to 12)
884	Does the household have a refrigerator? (No; Yes, with one door; Yes, with two doors)
854	How many household members are 11-years-old or younger? (Two or more; One; None)
850	How many years of schooling has the female head/spouse completed? (Three or less; Four to eleven; Twelve or more; No female head/spouse)
842	Do any household members ages 5 to 11 go to private school or private pre-school? (No; Yes; No members ages 5 to 11)

Figure 3 (cont.): Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (responses are ordered starting with those associated with higher poverty likelihoods)
815	What is the occupation of the male head/spouse in his main line of work? (Farmer, rancher, or undefined; Does not work; Service worker; No male head/spouse; Skilled worker in production, maintenance, and repair; Sales person or shopkeeper; Clerk; Mid-level technician; Manager, administrator, or professional in the arts and sciences)
764	How does the household dispose of sewage? (Ditch, other, or no bathroom; Simple hole, or directly into river, lake, or ocean; Septic tank not connected to public sewage/rainwater system; Septic tank connected to public sewage/rainwater system; Direct connection to public sewage/rainwater system)
718	In what sector is the main line of work of the female head/spouse? (Agriculture, or undefined categories; Does not work; Domestic service; Construction, food service or lodging, or other social/collective services; Manufacturing; No female head/spouse; Sales or repairs; Transportation, logistics, communications, public administration, education, health care, social services, or other industries or activities)
685	What type of employment does the female head/spouse have? (Subsistence farmer, working on own house, or unremunerated worker; Domestic servant without a written contract; Does not work; Employee without a written contract; Self-employed without employees; Domestic servant with a written contract; No female head/spouse; Employee with a written contract, civil servant or military, or self-employed with employees)
672	Do all household members ages 5 to 17 go to school or pre-school? (No; Yes; No members ages 5 to 17)
672	Do all household members ages 5 to 16 go to school or pre-school? (No; Yes; No members ages 5 to 16)
661	Do all household members ages 5 to 18 go to school or pre-school? (No; Yes; No members ages 5 to 18)
655	Do all household members ages 5 to 15 go to school or pre-school? (No; Yes; No members ages 5 to 15)
633	Do all household members ages 5 to 14 go to school or pre-school? (No; Yes; No members ages 5 to 14)

Figure 3 (cont.): Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (responses are ordered starting with those associated with higher poverty likelihoods)
605	How many rooms does the residence have? (One to four; Five; Six; Seven; Eight or more)
600	Do all household members ages 5 to 13 go to school or pre-school? (No; Yes; No members ages 5 to 13)
564	Do all household members ages 5 to 12 go to school or pre-school? (No; Yes; No members ages 5 to 12)
554	How many members does the household have? (Five or more; Four; Three; Two; One)
540	What is the race/ethnicity of the female head/spouse? (Mixed, indigenous, or not declared; Black; White or Asian; No female head/spouse)
531	How many household members work as employees with a written contract, as civil servants for the government, or in the military? (None; One; Two or more)
525	Do all household members ages 5 to 11 go to school or pre-school? (No; Yes; No members ages 5 to 11)
471	In what sector is the main line of work of the male head/spouse? (Agriculture; Construction, domestic service, or undefined categories; Does not work; No male head/spouse; Food service or lodging, sales or repairs, or other social/collective services; Manufacturing; Transportation, logistics, communications; Public administration, education, health care, social services, or other industries or activities)
433	What is the source of water for the household? (Not piped into the house, or piped into the house from some source other than a well, spring, or public water network; Piped from a well or spring into the house; Piped from the public water network into the house)
425	How many household members are 5-years-old or younger? (One or more; None)
414	What type of residence does the household have? (Detached house, or room; Apartment)
410	What type of employment does the male head/spouse have? (Subsistence farmer, working on own house, unremunerated worker, or employee without a written contract; Domestic servant, or self-employed without employees; Does not work; No male head/spouse; Employee with a written contract; Civil servant, military, or self-employed with employees)

Figure 3 (cont.): Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (responses are ordered starting with those associated with higher poverty likelihoods)
396	How does the household dispose of garbage? (Burned or buried on the property, thrown on the ground or public area, thrown in a river, lake, or ocean, or other; Picked up by a garbage truck from a community collection point; Picked up by a garbage truck from the residence)
391	How many household members attend a private school? (None; One or more)
368	In their main occupation, how many household members are farmers or ranchers? (Two or more; One; None)
357	Can the male head/spouse read and write? (No; No male head/spouse; Yes)
356	How many household members work in their main job in agriculture? (Two or more; One; None)
355	In their main job, how many household members work in the agricultural sector? (Two or more; One; None)
339	Can the female head/spouse read and write? (No; Yes; No female head/spouse)
333	What is the race/ethnicity of the male head/spouse? (Mixed, indigenous, or not declared; Black; White or Asian; No male head/spouse)
296	Does the household have a freezer? (No; Yes)
292	How many household members work as employees with a written contract? (None; One; Two or more)
266	What is the main construction material of the roof of the residence? (Other; Pre-fabricated concrete)
261	What type of fuel is used for cooking? (Other; LPG or piped gas)
257	How many household members work as civil servants for the government or in the military? (None; One or more)
234	How many household members in their main job work in sales or repair, food service or lodging, transportation, logistics, and communications, public administration, education, health, and social services, domestic service, or other services? (None; One; Two or more)
188	Does the household have a radio? (No; Yes)
166	How many household members work as subsistence farmers? (One or more; None)
162	What is the tenancy status of the household in its residence? (Provided for free, or other; Owned free-and-clear; Rented; Owned with a mortgage outstanding)

Figure 3 (cont.): Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (responses are ordered starting with those associated with higher poverty likelihoods)
147	Does the household have a color television? (No; Yes)
125	Did the female head/spouse do any work in the past week? (No; Yes; No female head/spouse)
106	What is the main construction material of the external walls of the residence? (Other; Masonry/bricks/tile, adobe, stone, pre-molded concrete or the like, including marble, metal, glass, or paneling)
67	What is the main source of energy for lighting the residence? (Other; Electricity (from public grid, generator, or solar))
63	What is the structure of household headship? (Both male and female heads/spouses; Female head/spouse only; Male head/spouse only)
62	Does the household have a water filter? (No; Yes)
53	Can all household members read and write? (No; Yes)
47	How many household members work as employees without a written contract? (One or more; None)
45	How many rooms in the residence are used as bedrooms? (Two; Two or more; One)
41	Does the household have a stove? (No; Yes, with one burner; Yes, with two or more burners)
36	How many household members do any work? (One; None; Three or more; Two)
31	In their main occupation, how many household members are service workers or sales people/shopkeepers? (None; One; Two or more)
21	How many household members are self-employed (with or without employees)? (None or one; Two or more)
20	Does the household own the land where the residence sits? (No; Yes)
6	How many household members in their main job work in industry/manufacturing or construction? (None; One or more)
1	Did the male head/spouse do any work in the past week? (Yes; No; No male head/spouse)

Source: 2008 Brazil PNAD and the half-minimum wage poverty line.

Half-Minimum-Wage Poverty Line

(and tables pertaining to all nine poverty lines)

Figure 4 (Half-minimum-wage line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	95.0
5–9	93.4
10–14	89.4
15–19	81.1
20–24	68.7
25–29	54.2
30–34	41.1
35–39	26.1
40–44	17.4
45–49	12.4
50–54	6.9
55–59	3.4
60–64	2.1
65–69	1.0
70–74	1.1
75–79	0.1
80–84	0.1
85–89	0.0
90–94	0.0
95–100	0.0

Figure 5 (Half-minimum-wage line): Derivation of estimated poverty likelihoods associated with scores

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	637	÷	670	=	95.0
5–9	1,195	÷	1,279	=	93.4
10–14	1,961	÷	2,193	=	89.4
15–19	2,938	÷	3,622	=	81.1
20–24	3,636	÷	5,291	=	68.7
25–29	3,646	÷	6,725	=	54.2
30–34	3,164	÷	7,701	=	41.1
35–39	2,440	÷	9,336	=	26.1
40–44	1,786	÷	10,277	=	17.4
45–49	1,220	÷	9,874	=	12.4
50–54	648	÷	9,353	=	6.9
55–59	272	÷	8,002	=	3.4
60–64	152	÷	7,254	=	2.1
65–69	62	÷	6,414	=	1.0
70–74	52	÷	4,862	=	1.1
75–79	2	÷	3,580	=	0.1
80–84	1	÷	2,038	=	0.1
85–89	0	÷	1,028	=	0.0
90–94	0	÷	382	=	0.0
95–100	0	÷	118	=	0.0

Number of all households normalized to sum to 100,000.

Figure 6 (All poverty lines): Distribution of household poverty likelihoods across ranges demarcated by poverty lines

	Likelihood of expenditure in range demarcated by poverty lines per day per adult equivalent								
	<$1.25/day	=>$1.25/day and <1/4	=>1/4 and <$2.50/day	=>$2.50/day and <$3.75/day	=>$3.75/day and <1/2	=>1/2 and <$5.00/day	=>$5.00/day and <1	=>1 and <2	=>2
Score	<BRL2.18	=>BRL2.18 and <BRL3.43	=>BRL3.43 and <BRL4.35	=>BRL4.35 and <BRL6.53	=>BRL6.53 and <BRL6.90	=>BRL6.90 and <BRL8.70	=>BRL8.70 and <BRL13.83	=>BRL13.83 and <BRL27.67	=>BRL27.67
0–4	46.4	25.3	10.2	11.9	1.3	4.0	0.5	0.5	0.0
5–9	34.2	31.2	12.4	14.2	1.4	4.0	2.3	0.4	0.0
10–14	24.0	27.6	14.4	21.3	2.1	4.9	5.2	0.5	0.0
15–19	14.0	20.9	14.0	27.0	5.1	9.2	8.1	1.4	0.1
20–24	10.2	14.4	12.5	26.8	4.7	11.6	15.9	3.5	0.3
25–29	7.1	9.0	7.9	23.6	6.7	13.3	24.7	7.2	0.6
30–34	4.6	5.9	5.0	18.0	7.7	12.2	31.7	13.7	1.3
35–39	3.3	2.9	2.5	11.1	6.4	11.0	38.1	21.3	3.5
40–44	2.0	1.8	1.3	6.8	5.4	8.6	35.8	31.9	6.3
45–49	1.9	0.6	0.6	4.6	4.6	7.7	31.9	37.6	10.4
50–54	1.5	0.2	0.3	1.9	3.0	3.7	24.9	46.6	17.9
55–59	1.0	0.2	0.0	0.8	1.4	2.2	18.8	45.0	30.6
60–64	1.1	0.0	0.1	0.3	0.6	1.7	11.5	43.4	41.2
65–69	0.4	0.0	0.0	0.2	0.3	0.9	7.0	34.1	57.1
70–74	0.6	0.0	0.0	0.2	0.3	0.2	2.7	25.8	70.3
75–79	0.0	0.0	0.0	0.1	0.0	0.0	1.3	18.0	80.6
80–84	0.0	0.0	0.0	0.0	0.1	0.2	0.5	9.5	89.7
85–89	0.0	0.0	0.0	0.0	0.0	0.0	1.4	6.1	92.5
90–94	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.7	94.3
95–100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0

Note: All poverty likelihoods in percentage units.

The USAID "extreme" line is omitted because it is almost the same as the $2.50/day 2005 PPP line.

Figure 7 (Half-minimum-wage line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) from the validation sample, with confidence intervals, scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	**Diff.**	**90-percent**	**95-percent**	**99-percent**
0–4	+1.1	4.0	4.7	6.4
5–9	+2.5	3.5	4.1	5.3
10–14	+3.2	3.2	3.8	5.3
15–19	+2.5	2.8	3.3	4.2
20–24	+1.2	2.9	3.4	4.4
25–29	+1.8	2.7	3.2	4.0
30–34	+3.0	2.3	2.7	3.9
35–39	–1.0	2.1	2.4	3.1
40–44	–2.4	2.1	2.2	2.5
45–49	+1.5	1.3	1.6	2.2
50–54	+0.9	1.1	1.3	1.7
55–59	–0.3	0.9	1.1	1.4
60–64	+0.4	0.7	0.8	1.0
65–69	–0.7	0.7	0.8	1.1
70–74	+0.6	0.4	0.5	0.6
75–79	–0.3	0.4	0.4	0.6
80–84	+0.1	0.0	0.0	0.0
85–89	+0.0	0.0	0.0	0.0
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 8 (All poverty lines): Differences, precision of differences, and the α factor for bootstrapped estimates of poverty rates for groups of households at a point in time for the scorecard applied to the validation sample

	Poverty line								
	National (minimum wages)				USAID	International 2005 PPP			
	1/2	1/4	1	2	'Extreme'	\$1.25/day	\$2.50/day	\$3.75/day	\$5.00/day
Estimate minus true value									
2005/6 scorecard applied to 2005/6 validation	+0.5	+0.3	–0.0	–0.1	+0.2	–0.1	+0.3	+0.5	+0.3
Precision of difference									
2005/6 scorecard applied to 2005/6 validation	0.4	0.3	0.5	0.4	0.3	0.3	0.3	0.4	0.5
α factor									
2005/6 scorecard applied to 2005/6 validation	0.79	0.89	0.79	0.83	0.87	0.96	0.86	0.78	0.78

Precision is measured as 90-percent confidence intervals in units of +/– percentage points.
Differences and precision estimated from 1,000 bootstraps of size $n = 16{,}384$.
α is estimated from 1,000 bootstrap samples of $n = 256, 512, 1{,}024, 2{,}048, 4{,}096, 8{,}192$, and $16{,}384$.

Figure 9 (Half-minimum-wage line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, scorecard applied to the validation sample

Sample Size n	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
	Diff.	90-percent	95-percent	99-percent
1	+0.3	57.5	68.4	92.1
4	+0.6	29.0	34.9	49.6
8	+0.9	20.1	24.9	33.0
16	+0.8	13.8	16.3	23.3
32	+0.8	9.5	11.0	14.6
64	+0.7	6.8	8.3	10.8
128	+0.5	4.8	5.6	7.6
256	+0.4	3.4	4.2	5.9
512	+0.5	2.5	3.0	3.9
1,024	+0.5	1.8	2.2	2.9
2,048	+0.4	1.2	1.5	1.9
4,096	+0.5	0.9	1.0	1.3
8,192	+0.4	0.6	0.8	1.1
16,384	+0.5	0.4	0.5	0.7

Figure 10 (All poverty lines): Possible outcomes from targeting by poverty score

		Targeting segment	
		Targeted	**Non-targeted**
True poverty status	**Below poverty line**	**Inclusion** Under poverty line Correctly Targeted	**Undercoverage** Under poverty line Mistakenly Non-targeted
	Above poverty line	**Leakage** Above poverty line Mistakenly Targeted	**Exclusion** Above poverty line Correctly Non-targeted

Figure 11 (Half-minimum-wage line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.6	23.2	0.0	76.1	76.8	–94.6
5–9	1.8	22.0	0.2	76.0	77.8	–84.3
10–14	3.7	20.1	0.4	75.7	79.4	–67.1
15–19	6.6	17.3	1.2	75.0	81.5	–39.9
20–24	10.2	13.7	2.9	73.3	83.4	–2.5
25–29	13.7	10.1	6.0	70.1	83.9	+40.7
30–34	16.8	7.0	10.7	65.5	82.3	+55.1
35–39	19.4	4.4	17.4	58.8	78.1	+26.8
40–44	21.5	2.3	25.6	50.6	72.0	–7.6
45–49	22.6	1.2	34.4	41.8	64.4	–44.3
50–54	23.2	0.6	43.1	33.1	56.3	–81.0
55–59	23.5	0.3	50.8	25.4	48.9	–113.2
60–64	23.7	0.1	57.9	18.3	42.0	–143.1
65–69	23.8	0.0	64.2	12.0	35.7	–169.6
70–74	23.8	0.0	69.1	7.1	30.9	–189.9
75–79	23.8	0.0	72.6	3.6	27.4	–204.9
80–84	23.8	0.0	74.7	1.5	25.3	–213.5
85–89	23.8	0.0	75.7	0.5	24.3	–217.8
90–94	23.8	0.0	76.1	0.1	23.9	–219.4
95–100	23.8	0.0	76.2	0.0	23.8	–219.9

Inclusion, undercoverage, leakage, and exclusion normalized to sum to 100.

Figure 12 (Half-minimum-wage line): For a given score cut-off, the percentage of all households who are targeted, the percentage of targeted households who are poor, the percentage of poor households who are targeted, and the number of poor households who are successful targeted (inclusion) per non-poor household mistakenly targeted (leakage), scorecard applied to the validation sample

Targeting cut-off	% all households who are targeted	% targeted who are poor	% of poor who are targeted	Poor households targeted per non-poor household targeted
0–4	0.7	93.5	2.6	14.4:1
5–9	1.9	92.3	7.6	12.0:1
10–14	4.1	89.2	15.5	8.3:1
15–19	7.8	84.5	27.5	5.4:1
20–24	13.1	77.8	42.6	3.5:1
25–29	19.8	69.5	57.7	2.3:1
30–34	27.5	61.1	70.5	1.6:1
35–39	36.8	52.7	81.4	1.1:1
40–44	47.1	45.6	90.2	0.8:1
45–49	57.0	39.7	94.9	0.7:1
50–54	66.3	35.0	97.5	0.5:1
55–59	74.3	31.7	98.8	0.5:1
60–64	81.6	29.0	99.4	0.4:1
65–69	88.0	27.0	99.8	0.4:1
70–74	92.9	25.6	99.9	0.3:1
75–79	96.4	24.7	100.0	0.3:1
80–84	98.5	24.2	100.0	0.3:1
85–89	99.5	23.9	100.0	0.3:1
90–94	99.9	23.8	100.0	0.3:1
95–100	100.0	23.8	100.0	0.3:1

Quarter-Minimum-Wage Poverty Line

Figure 4 (Quarter-minimum-wage line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	71.7
5–9	65.4
10–14	51.6
15–19	35.0
20–24	24.6
25–29	16.1
30–34	10.5
35–39	6.2
40–44	3.9
45–49	2.6
50–54	1.7
55–59	1.2
60–64	1.1
65–69	0.4
70–74	0.6
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Figure 5 (Quarter-minimum-wage line): Derivation of estimated poverty likelihoods associated with scores

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	480	÷	670	=	71.7
5–9	837	÷	1,279	=	65.4
10–14	1,132	÷	2,193	=	51.6
15–19	1,266	÷	3,622	=	35.0
20–24	1,303	÷	5,291	=	24.6
25–29	1,079	÷	6,725	=	16.1
30–34	806	÷	7,701	=	10.5
35–39	575	÷	9,336	=	6.2
40–44	396	÷	10,277	=	3.9
45–49	254	÷	9,874	=	2.6
50–54	162	÷	9,353	=	1.7
55–59	95	÷	8,002	=	1.2
60–64	83	÷	7,254	=	1.1
65–69	26	÷	6,414	=	0.4
70–74	31	÷	4,862	=	0.6
75–79	0	÷	3,580	=	0.0
80–84	0	÷	2,038	=	0.0
85–89	0	÷	1,028	=	0.0
90–94	0	÷	382	=	0.0
95–100	0	÷	118	=	0.0

Number of all households normalized to sum to 100,000.

Figure 7 (Quarter-minimum-wage line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) from the validation sample, with confidence intervals, scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+1.3	7.8	9.1	11.6
5–9	+8.5	5.9	7.0	10.0
10–14	+5.1	4.6	5.7	7.5
15–19	+0.6	3.4	4.2	5.4
20–24	–1.5	2.6	3.1	4.2
25–29	+1.2	1.8	2.2	2.9
30–34	+1.3	1.4	1.6	2.2
35–39	+0.0	1.0	1.2	1.8
40–44	–1.1	1.0	1.1	1.5
45–49	+0.1	0.7	0.8	1.0
50–54	+0.3	0.5	0.6	0.8
55–59	–0.0	0.5	0.7	0.9
60–64	+0.4	0.5	0.5	0.7
65–69	–0.5	0.6	0.6	0.8
70–74	+0.2	0.4	0.5	0.6
75–79	–0.2	0.2	0.3	0.3
80–84	+0.0	0.0	0.0	0.0
85–89	+0.0	0.0	0.0	0.0
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 9 (Quarter-minimum-wage line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	–0.7	50.0	59.5	78.0
4	+0.1	20.8	26.5	37.4
8	+0.4	14.3	17.7	24.6
16	+0.6	9.0	11.6	15.8
32	+0.5	6.8	8.4	10.7
64	+0.4	4.8	5.6	7.3
128	+0.4	3.5	4.1	5.3
256	+0.3	2.5	3.0	3.9
512	+0.3	1.9	2.2	2.8
1,024	+0.3	1.3	1.5	2.0
2,048	+0.3	0.9	1.1	1.5
4,096	+0.3	0.6	0.7	1.0
8,192	+0.3	0.5	0.5	0.7
16,384	+0.3	0.3	0.4	0.5

Figure 11 (Quarter-minimum-wage line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.5	8.1	0.2	91.3	91.7	–86.6
5–9	1.2	7.3	0.7	90.7	92.0	–62.9
10–14	2.3	6.3	1.9	89.6	91.8	–25.0
15–19	3.5	5.0	4.2	87.3	90.8	+32.7
20–24	5.0	3.6	8.1	83.4	88.4	+5.2
25–29	6.0	2.5	13.8	77.7	83.7	–61.6
30–34	6.8	1.8	20.7	70.8	77.5	–142.9
35–39	7.4	1.1	29.4	62.0	69.4	–245.3
40–44	7.9	0.6	39.2	52.3	60.2	–359.8
45–49	8.1	0.4	48.8	42.6	50.8	–472.7
50–54	8.3	0.2	58.0	33.5	41.7	–580.5
55–59	8.4	0.1	65.9	25.5	33.9	–673.3
60–64	8.4	0.1	73.1	18.3	26.8	–757.7
65–69	8.5	0.0	79.5	12.0	20.5	–832.3
70–74	8.5	0.0	84.3	7.1	15.7	–889.1
75–79	8.5	0.0	87.9	3.6	12.1	–931.0
80–84	8.5	0.0	89.9	1.5	10.1	–954.9
85–89	8.5	0.0	91.0	0.5	9.0	–966.9
90–94	8.5	0.0	91.4	0.1	8.6	–971.4
95–100	8.5	0.0	91.5	0.0	8.5	–972.8

Inclusion, undercoverage, leakage, and exclusion normalized to sum to 100.

Figure 12 (Quarter-minimum-wage line): For a given score cut-off, the percentage of all households who are targeted, the percentage of targeted households who are poor, the percentage of poor households who are targeted, and the number of poor households who are successful targeted (inclusion) per non-poor household mistakenly targeted (leakage), scorecard applied to the validation sample

Targeting cut-off	% all households who are targeted	% targeted who are poor	% of poor who are targeted	Poor households targeted per non-poor household targeted
0–4	0.7	70.4	5.5	2.4:1
5–9	1.9	62.5	14.3	1.7:1
10–14	4.1	54.4	26.4	1.2:1
15–19	7.8	45.7	41.6	0.8:1
20–24	13.1	38.1	58.3	0.6:1
25–29	19.8	30.3	70.3	0.4:1
30–34	27.5	24.6	79.4	0.3:1
35–39	36.8	20.0	86.5	0.3:1
40–44	47.1	16.7	92.5	0.2:1
45–49	57.0	14.3	95.4	0.2:1
50–54	66.3	12.5	97.3	0.1:1
55–59	74.3	11.3	98.4	0.1:1
60–64	81.6	10.3	99.0	0.1:1
65–69	88.0	9.7	99.7	0.1:1
70–74	92.9	9.2	99.9	0.1:1
75–79	96.4	8.8	100.0	0.1:1
80–84	98.5	8.7	100.0	0.1:1
85–89	99.5	8.6	100.0	0.1:1
90–94	99.9	8.5	100.0	0.1:1
95–100	100.0	8.5	100.0	0.1:1

One-Minimum-Wage Poverty Line

Figure 4 (One-minimum-wage line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	99.5
5–9	99.6
10–14	99.5
15–19	98.5
20–24	96.2
25–29	92.2
30–34	85.0
35–39	75.3
40–44	61.8
45–49	52.0
50–54	35.6
55–59	24.4
60–64	15.4
65–69	8.9
70–74	3.9
75–79	1.4
80–84	0.8
85–89	1.4
90–94	0.0
95–100	0.0

Figure 5 (One-minimum-wage line): Derivation of estimated poverty likelihoods associated with scores

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	667	÷	670	=	99.5
5–9	1,274	÷	1,279	=	99.6
10–14	2,181	÷	2,193	=	99.5
15–19	3,566	÷	3,622	=	98.5
20–24	5,088	÷	5,291	=	96.2
25–29	6,201	÷	6,725	=	92.2
30–34	6,547	÷	7,701	=	85.0
35–39	7,028	÷	9,336	=	75.3
40–44	6,350	÷	10,277	=	61.8
45–49	5,134	÷	9,874	=	52.0
50–54	3,326	÷	9,353	=	35.6
55–59	1,952	÷	8,002	=	24.4
60–64	1,114	÷	7,254	=	15.4
65–69	568	÷	6,414	=	8.9
70–74	191	÷	4,862	=	3.9
75–79	49	÷	3,580	=	1.4
80–84	16	÷	2,038	=	0.8
85–89	14	÷	1,028	=	1.4
90–94	0	÷	382	=	0.0
95–100	0	÷	118	=	0.0

Number of all households normalized to sum to 100,000.

Figure 7 (One-minimum-wage line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) from the validation sample, with confidence intervals, scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	**Diff.**	**90-percent**	**95-percent**	**99-percent**
0–4	–0.5	0.2	0.2	0.2
5–9	–0.2	0.3	0.4	0.6
10–14	+0.2	0.7	0.8	0.9
15–19	–0.1	0.9	1.0	1.4
20–24	+0.6	1.2	1.5	1.8
25–29	+0.3	1.4	1.7	2.2
30–34	+0.6	1.8	2.2	3.0
35–39	–0.5	1.9	2.3	2.9
40–44	–2.4	2.3	2.5	3.2
45–49	+2.4	2.3	2.7	3.4
50–54	–1.1	2.3	2.6	3.6
55–59	–0.7	2.0	2.3	3.2
60–64	+0.5	1.8	2.2	2.7
65–69	+0.9	1.5	1.7	2.2
70–74	+0.4	1.1	1.3	1.7
75–79	–0.7	1.0	1.2	1.6
80–84	–0.6	1.2	1.4	1.8
85–89	+1.0	0.6	0.7	0.9
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 9 (One-minimum-wage line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, scorecard applied to the validation sample

Sample Size n	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
	Diff.	90-percent	95-percent	99-percent
1	–0.7	63.1	75.4	91.7
4	–0.3	32.8	38.8	51.4
8	+0.1	23.0	27.9	36.6
16	–0.1	15.5	19.0	24.9
32	–0.3	11.3	13.7	18.6
64	–0.2	8.1	9.6	12.7
128	–0.1	5.6	6.7	9.3
256	–0.1	4.1	4.9	6.2
512	–0.1	2.9	3.4	4.7
1,024	–0.1	2.0	2.5	3.1
2,048	–0.1	1.4	1.7	2.1
4,096	–0.0	1.0	1.2	1.6
8,192	–0.1	0.7	0.8	1.1
16,384	–0.0	0.5	0.6	0.8

Figure 11 (One-minimum-wage line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.7	51.1	0.0	48.2	48.9	–97.4
5–9	1.9	49.9	0.0	48.2	50.1	–92.5
10–14	4.1	47.7	0.0	48.2	52.3	–84.1
15–19	7.7	44.1	0.1	48.1	55.8	–70.2
20–24	12.8	39.1	0.3	47.9	60.6	–50.2
25–29	18.9	32.9	0.8	47.3	66.3	–25.3
30–34	25.5	26.4	2.0	46.2	71.6	+2.2
35–39	32.6	19.2	4.2	44.0	76.5	+33.9
40–44	39.3	12.5	7.8	40.4	79.6	+66.7
45–49	44.2	7.6	12.7	35.5	79.7	+75.5
50–54	47.8	4.0	18.5	29.6	77.4	+64.2
55–59	49.8	2.0	24.5	23.7	73.6	+52.8
60–64	51.0	0.9	30.6	17.6	68.5	+40.9
65–69	51.5	0.3	36.5	11.7	63.2	+29.6
70–74	51.7	0.1	41.1	7.0	58.8	+20.6
75–79	51.8	0.0	44.6	3.5	55.3	+13.8
80–84	51.8	0.0	46.7	1.5	53.3	+10.0
85–89	51.8	0.0	47.7	0.5	52.3	+8.0
90–94	51.8	0.0	48.1	0.1	51.9	+7.2
95–100	51.8	0.0	48.2	0.0	51.8	+7.0

Inclusion, undercoverage, leakage, and exclusion normalized to sum to 100.

Figure 12 (One-minimum-wage line): For a given score cut-off, the percentage of all households who are targeted, the percentage of targeted households who are poor, the percentage of poor households who are targeted, and the number of poor households who are successful targeted (inclusion) per non-poor household mistakenly targeted (leakage), scorecard applied to the validation sample

Targeting cut-off	% all households who are targeted	% targeted who are poor	% of poor who are targeted	Poor households targeted per non-poor household targeted
0–4	0.7	100.0	1.3	Only poor targeted
5–9	1.9	99.9	3.8	812.3:1
10–14	4.1	99.4	7.9	179.1:1
15–19	7.8	99.0	14.8	97.1:1
20–24	13.1	97.7	24.6	42.3:1
25–29	19.8	95.7	36.5	22.3:1
30–34	27.5	92.7	49.1	12.6:1
35–39	36.8	88.5	62.9	7.7:1
40–44	47.1	83.4	75.8	5.0:1
45–49	57.0	77.7	85.4	3.5:1
50–54	66.3	72.0	92.2	2.6:1
55–59	74.3	67.1	96.2	2.0:1
60–64	81.6	62.5	98.3	1.7:1
65–69	88.0	58.5	99.4	1.4:1
70–74	92.9	55.7	99.8	1.3:1
75–79	96.4	53.7	99.9	1.2:1
80–84	98.5	52.6	100.0	1.1:1
85–89	99.5	52.1	100.0	1.1:1
90–94	99.9	51.9	100.0	1.1:1
95–100	100.0	51.8	100.0	1.1:1

Two-Minimum-Wage Poverty Line

Figure 4 (Two-minimum-wage line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	100.0
15–19	99.9
20–24	99.7
25–29	99.4
30–34	98.7
35–39	96.5
40–44	93.7
45–49	89.6
50–54	82.1
55–59	69.4
60–64	58.8
65–69	42.9
70–74	29.8
75–79	19.4
80–84	10.3
85–89	7.5
90–94	5.7
95–100	0.0

Figure 5 (Two-minimum-wage line): Derivation of estimated poverty likelihoods associated with scores

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	670	÷	670	=	100.0
5–9	1,279	÷	1,279	=	100.0
10–14	2,193	÷	2,193	=	100.0
15–19	3,617	÷	3,622	=	99.9
20–24	5,275	÷	5,291	=	99.7
25–29	6,684	÷	6,725	=	99.4
30–34	7,599	÷	7,701	=	98.7
35–39	9,013	÷	9,336	=	96.5
40–44	9,626	÷	10,277	=	93.7
45–49	8,851	÷	9,874	=	89.6
50–54	7,680	÷	9,353	=	82.1
55–59	5,550	÷	8,002	=	69.4
60–64	4,265	÷	7,254	=	58.8
65–69	2,753	÷	6,414	=	42.9
70–74	1,446	÷	4,862	=	29.8
75–79	694	÷	3,580	=	19.4
80–84	210	÷	2,038	=	10.3
85–89	77	÷	1,028	=	7.5
90–94	22	÷	382	=	5.7
95–100	0	÷	118	=	0.0

Number of all households normalized to sum to 100,000.

Figure 7 (Two-minimum-wage line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) from the validation sample, with confidence intervals, scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+0.0	0.0	0.0	0.0
5–9	+0.0	0.0	0.0	0.0
10–14	+0.2	0.3	0.3	0.5
15–19	–0.1	0.1	0.1	0.1
20–24	+0.0	0.4	0.4	0.5
25–29	–0.1	0.4	0.5	0.6
30–34	+0.0	0.6	0.7	0.9
35–39	+0.2	0.9	1.0	1.2
40–44	–0.7	1.0	1.2	1.6
45–49	+0.9	1.4	1.6	2.3
50–54	+1.7	1.8	2.1	2.7
55–59	–3.6	2.8	3.1	3.6
60–64	+1.0	2.5	3.0	3.8
65–69	+0.9	2.6	3.1	3.8
70–74	–1.1	2.9	3.4	4.9
75–79	–1.1	2.8	3.3	4.7
80–84	–1.4	3.2	3.8	4.6
85–89	+3.1	2.5	3.1	4.0
90–94	+2.7	3.6	4.4	5.4
95–100	+0.0	0.0	0.0	0.0

Figure 9 (Two-minimum-wage line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, scorecard applied to the validation sample

Sample Size n	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
	Diff.	90-percent	95-percent	99-percent
1	–0.9	60.6	76.2	88.6
4	+0.5	30.7	37.0	46.7
8	+0.4	21.1	24.6	32.9
16	–0.1	14.7	17.6	22.3
32	–0.3	10.2	11.9	16.6
64	–0.2	7.5	8.8	11.1
128	–0.2	5.3	6.3	8.7
256	–0.1	3.6	4.3	5.7
512	–0.1	2.5	3.1	4.1
1,024	–0.1	1.8	2.2	3.0
2,048	–0.1	1.3	1.5	1.9
4,096	–0.1	0.9	1.0	1.4
8,192	–0.1	0.6	0.7	1.0
16,384	–0.1	0.4	0.5	0.7

Figure 11 (Two-minimum-wage line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.7	76.9	0.0	22.4	23.1	–98.3
5–9	1.9	75.7	0.0	22.4	24.3	–95.0
10–14	4.1	73.5	0.0	22.4	26.5	–89.3
15–19	7.8	69.8	0.0	22.4	30.1	–80.0
20–24	13.0	64.6	0.0	22.4	35.4	–66.4
25–29	19.7	57.9	0.1	22.3	42.1	–49.1
30–34	27.3	50.3	0.2	22.2	49.5	–29.4
35–39	36.3	41.3	0.5	21.9	58.2	–5.8
40–44	46.0	31.6	1.1	21.3	67.3	+20.0
45–49	54.8	22.8	2.2	20.2	75.0	+44.0
50–54	62.3	15.3	4.0	18.4	80.7	+65.8
55–59	68.1	9.5	6.2	16.2	84.4	+83.6
60–64	72.3	5.3	9.2	13.2	85.5	+88.1
65–69	75.0	2.6	13.0	9.4	84.5	+83.3
70–74	76.6	1.0	16.3	6.1	82.7	+79.0
75–79	77.3	0.3	19.1	3.3	80.6	+75.3
80–84	77.5	0.1	20.9	1.5	79.0	+73.0
85–89	77.6	0.0	21.9	0.5	78.1	+71.8
90–94	77.6	0.0	22.3	0.1	77.7	+71.3
95–100	77.6	0.0	22.4	0.0	77.6	+71.1

Inclusion, undercoverage, leakage, and exclusion normalized to sum to 100.

Figure 12 (Two-minimum-wage line): For a given score cut-off, the percentage of all households who are targeted, the percentage of targeted households who are poor, the percentage of poor households who are targeted, and the number of poor households who are successful targeted (inclusion) per non-poor household mistakenly targeted (leakage), scorecard applied to the validation sample

Targeting cut-off	% all households who are targeted	% targeted who are poor	% of poor who are targeted	Poor households targeted per non-poor household targeted
0–4	0.7	100.0	0.9	Only poor targeted
5–9	1.9	100.0	2.5	Only poor targeted
10–14	4.1	99.9	5.3	862.2:1
15–19	7.8	99.9	10.0	1,617.2:1
20–24	13.1	99.8	16.8	623.5:1
25–29	19.8	99.7	25.4	330.0:1
30–34	27.5	99.4	35.2	165.9:1
35–39	36.8	98.6	46.8	73.0:1
40–44	47.1	97.7	59.3	42.5:1
45–49	57.0	96.1	70.6	25.0:1
50–54	66.3	94.0	80.3	15.6:1
55–59	74.3	91.7	87.8	11.0:1
60–64	81.6	88.7	93.2	7.8:1
65–69	88.0	85.3	96.7	5.8:1
70–74	92.9	82.5	98.7	4.7:1
75–79	96.4	80.2	99.6	4.0:1
80–84	98.5	78.7	99.9	3.7:1
85–89	99.5	78.0	100.0	3.5:1
90–94	99.9	77.7	100.0	3.5:1
95–100	100.0	77.6	100.0	3.5:1

USAID "Extreme" Poverty Line

Figure 4 (USAID "extreme" line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	80.2
5–9	77.2
10–14	65.0
15–19	47.0
20–24	36.1
25–29	23.2
30–34	15.2
35–39	8.3
40–44	5.1
45–49	3.1
50–54	2.1
55–59	1.2
60–64	1.2
65–69	0.4
70–74	0.6
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Figure 5 (USAID "extreme" line): Derivation of estimated poverty likelihoods associated with scores

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	537	÷	670	=	80.2
5–9	987	÷	1,279	=	77.2
10–14	1,425	÷	2,193	=	65.0
15–19	1,702	÷	3,622	=	47.0
20–24	1,909	÷	5,291	=	36.1
25–29	1,563	÷	6,725	=	23.2
30–34	1,167	÷	7,701	=	15.2
35–39	776	÷	9,336	=	8.3
40–44	520	÷	10,277	=	5.1
45–49	306	÷	9,874	=	3.1
50–54	193	÷	9,353	=	2.1
55–59	95	÷	8,002	=	1.2
60–64	86	÷	7,254	=	1.2
65–69	28	÷	6,414	=	0.4
70–74	31	÷	4,862	=	0.6
75–79	0	÷	3,580	=	0.0
80–84	0	÷	2,038	=	0.0
85–89	0	÷	1,028	=	0.0
90–94	0	÷	382	=	0.0
95–100	0	÷	118	=	0.0

Number of all households normalized to sum to 100,000.

Figure 7 (USAID "extreme" line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) from the validation sample, with confidence intervals, scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	**Diff.**	**90-percent**	**95-percent**	**99-percent**
0–4	+0.9	6.9	8.0	10.4
5–9	+7.5	5.7	6.7	8.3
10–14	+2.4	4.4	5.2	6.8
15–19	+0.3	3.5	4.2	5.5
20–24	+0.1	2.8	3.4	4.8
25–29	+1.0	2.2	2.5	3.3
30–34	+1.9	1.7	2.0	2.4
35–39	–0.3	1.3	1.6	2.1
40–44	–1.3	1.2	1.3	1.6
45–49	+0.3	0.7	0.9	1.1
50–54	+0.4	0.5	0.6	0.9
55–59	–0.2	0.6	0.7	0.9
60–64	+0.3	0.5	0.6	0.8
65–69	–0.5	0.5	0.6	0.8
70–74	+0.2	0.4	0.5	0.6
75–79	–0.3	0.3	0.4	0.5
80–84	+0.0	0.0	0.0	0.0
85–89	+0.0	0.0	0.0	0.0
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 9 (USAID "extreme" line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	–1.0	50.0	60.5	80.9
4	+0.1	22.8	28.6	38.7
8	+0.5	16.5	19.6	26.6
16	+0.6	10.5	12.4	17.0
32	+0.5	7.4	9.0	12.2
64	+0.4	5.3	6.6	8.8
128	+0.3	4.0	4.8	6.0
256	+0.3	3.0	3.5	4.5
512	+0.2	2.0	2.3	3.2
1,024	+0.3	1.4	1.6	2.1
2,048	+0.2	1.0	1.2	1.5
4,096	+0.2	0.7	0.8	1.1
8,192	+0.2	0.5	0.6	0.8
16,384	+0.2	0.3	0.4	0.5

Figure 11 (USAID “extreme” line): Households by targeting classification and score, along with “Total Accuracy” and BPAC, scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.5	10.8	0.1	88.5	89.1	–89.4
5–9	1.4	9.9	0.5	88.1	89.6	–70.2
10–14	2.8	8.5	1.3	87.3	90.1	–38.7
15–19	4.5	6.8	3.2	85.4	90.0	+8.5
20–24	6.5	4.8	6.6	82.1	88.6	+42.0
25–29	8.0	3.3	11.8	76.9	84.9	–3.8
30–34	9.1	2.2	18.4	70.3	79.4	–62.2
35–39	9.9	1.4	26.9	61.8	71.7	–137.2
40–44	10.6	0.7	36.5	52.2	62.8	–222.0
45–49	10.9	0.4	46.1	42.6	53.5	–306.5
50–54	11.1	0.2	55.2	33.4	44.5	–387.4
55–59	11.2	0.1	63.1	25.5	36.7	–457.2
60–64	11.2	0.1	70.3	18.3	29.6	–520.6
65–69	11.3	0.0	76.7	12.0	23.3	–576.8
70–74	11.3	0.0	81.5	7.1	18.5	–619.5
75–79	11.3	0.0	85.1	3.6	14.9	–651.0
80–84	11.3	0.0	87.1	1.5	12.9	–668.9
85–89	11.3	0.0	88.2	0.5	11.8	–678.0
90–94	11.3	0.0	88.5	0.1	11.5	–681.4
95–100	11.3	0.0	88.7	0.0	11.3	–682.4

Inclusion, undercoverage, leakage, and exclusion normalized to sum to 100.

Figure 12 (USAID "extreme" line): For a given score cut-off, the percentage of all households who are targeted, the percentage of targeted households who are poor, the percentage of poor households who are targeted, and the number of poor households who are successful targeted (inclusion) per non-poor household mistakenly targeted (leakage), scorecard applied to the validation sample

Targeting cut-off	% all households who are targeted	% targeted who are poor	% of poor who are targeted	Poor households targeted per non-poor household targeted
0–4	0.7	79.0	4.7	3.8:1
5–9	1.9	73.4	12.6	2.8:1
10–14	4.1	67.6	24.7	2.1:1
15–19	7.8	58.4	40.0	1.4:1
20–24	13.1	49.7	57.2	1.0:1
25–29	19.8	40.5	70.7	0.7:1
30–34	27.5	33.1	80.3	0.5:1
35–39	36.8	27.0	87.7	0.4:1
40–44	47.1	22.5	93.5	0.3:1
45–49	57.0	19.1	96.2	0.2:1
50–54	66.3	16.7	97.8	0.2:1
55–59	74.3	15.0	98.7	0.2:1
60–64	81.6	13.8	99.2	0.2:1
65–69	88.0	12.8	99.7	0.1:1
70–74	92.9	12.2	99.9	0.1:1
75–79	96.4	11.8	100.0	0.1:1
80–84	98.5	11.5	100.0	0.1:1
85–89	99.5	11.4	100.0	0.1:1
90–94	99.9	11.3	100.0	0.1:1
95–100	100.0	11.3	100.0	0.1:1

$1.25/Day 2005 PPP Poverty Line

Figure 4 ($1.25/day 2005 PPP line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	46.4
5–9	34.2
10–14	24.0
15–19	14.0
20–24	10.2
25–29	7.1
30–34	4.6
35–39	3.3
40–44	2.0
45–49	1.9
50–54	1.5
55–59	1.0
60–64	1.1
65–69	0.4
70–74	0.6
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Figure 5 ($1.25/day 2005 PPP line): Derivation of estimated poverty likelihoods associated with scores

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	311	÷	670	=	46.4
5–9	438	÷	1,279	=	34.2
10–14	526	÷	2,193	=	24.0
15–19	508	÷	3,622	=	14.0
20–24	541	÷	5,291	=	10.2
25–29	474	÷	6,725	=	7.1
30–34	352	÷	7,701	=	4.6
35–39	304	÷	9,336	=	3.3
40–44	210	÷	10,277	=	2.0
45–49	191	÷	9,874	=	1.9
50–54	141	÷	9,353	=	1.5
55–59	78	÷	8,002	=	1.0
60–64	83	÷	7,254	=	1.1
65–69	25	÷	6,414	=	0.4
70–74	31	÷	4,862	=	0.6
75–79	0	÷	3,580	=	0.0
80–84	0	÷	2,038	=	0.0
85–89	0	÷	1,028	=	0.0
90–94	0	÷	382	=	0.0
95–100	0	÷	118	=	0.0

Number of all households normalized to sum to 100,000.

Figure 7 ($1.25/day 2005 PPP line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) from the validation sample, with confidence intervals, scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	**Diff.**	**90-percent**	**95-percent**	**99-percent**
0–4	+0.5	8.3	10.1	13.3
5–9	+3.7	5.6	6.8	9.0
10–14	+1.1	3.9	4.6	6.3
15–19	–1.7	2.5	3.0	4.1
20–24	–0.5	1.8	2.1	2.8
25–29	+0.5	1.3	1.5	2.0
30–34	+0.4	0.9	1.1	1.5
35–39	–0.4	0.8	1.0	1.3
40–44	–1.2	1.0	1.0	1.3
45–49	+0.1	0.6	0.7	0.9
50–54	+0.4	0.4	0.5	0.7
55–59	+0.0	0.5	0.6	0.8
60–64	+0.5	0.4	0.5	0.6
65–69	–0.4	0.5	0.6	0.7
70–74	+0.2	0.4	0.5	0.6
75–79	–0.2	0.2	0.3	0.3
80–84	+0.0	0.0	0.0	0.0
85–89	+0.0	0.0	0.0	0.0
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 9 ($1.25/day 2005 PPP line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, scorecard applied to the validation sample

Sample Size n	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
	Diff.	90-percent	95-percent	99-percent
1	–0.5	7.0	56.9	66.1
4	–0.2	15.6	19.6	29.3
8	+0.1	10.5	12.5	17.3
16	+0.1	7.0	8.9	12.0
32	+0.1	5.1	6.2	8.6
64	+0.1	3.7	4.3	5.6
128	+0.1	2.7	3.1	4.0
256	+0.0	2.0	2.3	3.0
512	–0.0	1.4	1.7	2.3
1,024	–0.0	1.0	1.2	1.5
2,048	–0.0	0.7	0.8	1.1
4,096	–0.1	0.5	0.6	0.8
8,192	–0.1	0.4	0.4	0.5
16,384	–0.1	0.3	0.3	0.4

Figure 11 ($1.25/day 2005 PPP line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.3	4.0	0.4	95.3	95.6	–77.7
5–9	0.7	3.6	1.2	94.4	95.1	–39.0
10–14	1.2	3.1	2.9	92.7	93.9	+23.3
15–19	1.8	2.5	6.0	89.7	91.5	–37.3
20–24	2.4	2.0	10.7	85.0	87.4	–146.0
25–29	2.8	1.5	17.0	78.7	81.5	–290.8
30–34	3.2	1.2	24.3	71.3	74.5	–460.2
35–39	3.5	0.8	33.3	62.3	65.9	–667.1
40–44	3.8	0.5	43.3	52.4	56.2	–896.1
45–49	4.0	0.3	52.9	42.7	46.7	–1,119.2
50–54	4.1	0.2	62.2	33.5	37.6	–1,331.8
55–59	4.2	0.1	70.1	25.6	29.8	–1,514.4
60–64	4.3	0.1	77.3	18.3	22.6	–1,680.4
65–69	4.3	0.0	83.7	12.0	16.3	–1,827.0
70–74	4.3	0.0	88.5	7.1	11.5	–1,938.4
75–79	4.3	0.0	92.1	3.6	7.9	–2,020.7
80–84	4.3	0.0	94.1	1.5	5.9	–2,067.6
85–89	4.3	0.0	95.2	0.5	4.8	–2,091.3
90–94	4.3	0.0	95.5	0.1	4.5	–2,100.1
95–100	4.3	0.0	95.7	0.0	4.3	–2,102.8

Inclusion, undercoverage, leakage, and exclusion normalized to sum to 100.

Figure 12 ($1.25/day 2005 PPP line): For a given score cut-off, the percentage of all households who are targeted, the percentage of targeted households who are poor, the percentage of poor households who are targeted, and the number of poor households who are successful targeted (inclusion) per non-poor household mistakenly targeted (leakage), scorecard applied to the validation sample

Targeting cut-off	% all households who are targeted	% targeted who are poor	% of poor who are targeted	Poor households targeted per non-poor household targeted
0–4	0.7	44.6	6.9	0.8:1
5–9	1.9	35.9	16.1	0.6:1
10–14	4.1	29.3	27.9	0.4:1
15–19	7.8	23.2	41.4	0.3:1
20–24	13.1	18.2	54.7	0.2:1
25–29	19.8	14.2	64.7	0.2:1
30–34	27.5	11.5	72.6	0.1:1
35–39	36.8	9.5	80.7	0.1:1
40–44	47.1	8.2	88.4	0.1:1
45–49	57.0	7.1	92.7	0.1:1
50–54	66.3	6.3	95.5	0.1:1
55–59	74.3	5.7	97.2	0.1:1
60–64	81.6	5.2	98.2	0.1:1
65–69	88.0	4.9	99.3	0.1:1
70–74	92.9	4.7	99.8	0.0:1
75–79	96.4	4.5	100.0	0.0:1
80–84	98.5	4.4	100.0	0.0:1
85–89	99.5	4.4	100.0	0.0:1
90–94	99.9	4.3	100.0	0.0:1
95–100	100.0	4.3	100.0	0.0:1

$2.50/Day 2005 PPP Poverty Line

Figure 4 ($2.50/day 2005 PPP line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	81.8
5–9	77.8
10–14	66.1
15–19	49.0
20–24	37.2
25–29	23.9
30–34	15.4
35–39	8.6
40–44	5.2
45–49	3.2
50–54	2.1
55–59	1.2
60–64	1.2
65–69	0.4
70–74	0.6
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Figure 5 ($2.50/day 2005 PPP line): Derivation of estimated poverty likelihoods associated with scores

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	548	÷	670	=	81.8
5–9	995	÷	1,279	=	77.8
10–14	1,449	÷	2,193	=	66.1
15–19	1,775	÷	3,622	=	49.0
20–24	1,966	÷	5,291	=	37.2
25–29	1,608	÷	6,725	=	23.9
30–34	1,188	÷	7,701	=	15.4
35–39	804	÷	9,336	=	8.6
40–44	530	÷	10,277	=	5.2
45–49	312	÷	9,874	=	3.2
50–54	193	÷	9,353	=	2.1
55–59	95	÷	8,002	=	1.2
60–64	86	÷	7,254	=	1.2
65–69	28	÷	6,414	=	0.4
70–74	31	÷	4,862	=	0.6
75–79	0	÷	3,580	=	0.0
80–84	0	÷	2,038	=	0.0
85–89	0	÷	1,028	=	0.0
90–94	0	÷	382	=	0.0
95–100	0	÷	118	=	0.0

Number of all households normalized to sum to 100,000.

Figure 7 ($2.50/day 2005 PPP line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) from the validation sample, with confidence intervals, scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+1.6	6.7	8.1	10.2
5–9	+6.9	5.6	6.5	8.3
10–14	+2.2	4.4	5.2	7.0
15–19	+0.9	3.5	4.3	5.6
20–24	+0.1	2.9	3.4	4.9
25–29	+1.1	2.2	2.6	3.3
30–34	+1.9	1.7	2.0	2.4
35–39	–0.2	1.3	1.6	2.1
40–44	–1.2	1.2	1.3	1.6
45–49	+0.3	0.7	0.9	1.2
50–54	+0.3	0.5	0.7	0.9
55–59	–0.2	0.6	0.7	0.9
60–64	+0.3	0.5	0.6	0.8
65–69	–0.5	0.5	0.6	0.8
70–74	+0.2	0.4	0.5	0.6
75–79	–0.3	0.3	0.4	0.5
80–84	+0.0	0.0	0.0	0.0
85–89	+0.0	0.0	0.0	0.0
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 9 ($2.50/day 2005 PPP line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	–0.7	50.0	60.9	81.5
4	+0.1	23.1	28.9	39.6
8	+0.5	16.6	19.7	26.5
16	+0.6	10.7	12.6	17.5
32	+0.6	7.5	8.9	12.1
64	+0.5	5.4	6.6	9.3
128	+0.3	4.0	4.7	6.2
256	+0.3	3.0	3.4	4.5
512	+0.3	2.0	2.4	3.2
1,024	+0.3	1.4	1.6	2.2
2,048	+0.3	1.0	1.2	1.5
4,096	+0.3	0.7	0.8	1.1
8,192	+0.3	0.5	0.6	0.8
16,384	+0.3	0.3	0.4	0.5

Figure 11 ($2.50/day 2005 PPP line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.5	11.1	0.1	88.3	88.8	–89.6
5–9	1.5	10.1	0.5	87.9	89.4	–70.7
10–14	2.8	8.7	1.3	87.1	90.0	–39.7
15–19	4.6	7.0	3.1	85.3	89.9	+6.9
20–24	6.7	4.9	6.4	82.0	88.7	+44.8
25–29	8.2	3.4	11.6	76.8	85.1	+0.3
30–34	9.3	2.3	18.2	70.2	79.6	–56.6
35–39	10.2	1.4	26.6	61.8	72.0	–129.7
40–44	10.9	0.7	36.2	52.2	63.0	–212.6
45–49	11.2	0.4	45.8	42.6	53.8	–295.1
50–54	11.3	0.2	55.0	33.4	44.8	–374.1
55–59	11.4	0.1	62.9	25.5	37.0	–442.3
60–64	11.5	0.1	70.1	18.3	29.8	–504.3
65–69	11.6	0.0	76.4	12.0	23.5	–559.1
70–74	11.6	0.0	81.3	7.1	18.7	–600.9
75–79	11.6	0.0	84.8	3.6	15.2	–631.6
80–84	11.6	0.0	86.9	1.5	13.1	–649.2
85–89	11.6	0.0	87.9	0.5	12.1	–658.1
90–94	11.6	0.0	88.3	0.1	11.7	–661.4
95–100	11.6	0.0	88.4	0.0	11.6	–662.4

Inclusion, undercoverage, leakage, and exclusion normalized to sum to 100.

Figure 12 ($2.50/day 2005 PPP line): For a given score cut-off, the percentage of all households who are targeted, the percentage of targeted households who are poor, the percentage of poor households who are targeted, and the number of poor households who are successful targeted (inclusion) per non-poor household mistakenly targeted (leakage), scorecard applied to the validation sample

Targeting cut-off	% all households who are targeted	% targeted who are poor	% of poor who are targeted	Poor households targeted per non-poor household targeted
0–4	0.7	79.8	4.6	3.9:1
5–9	1.9	74.5	12.5	2.9:1
10–14	4.1	68.8	24.6	2.2:1
15–19	7.8	59.7	40.0	1.5:1
20–24	13.1	50.9	57.4	1.0:1
25–29	19.8	41.5	70.9	0.7:1
30–34	27.5	33.9	80.4	0.5:1
35–39	36.8	27.7	87.8	0.4:1
40–44	47.1	23.0	93.6	0.3:1
45–49	57.0	19.6	96.2	0.2:1
50–54	66.3	17.1	97.9	0.2:1
55–59	74.3	15.4	98.7	0.2:1
60–64	81.6	14.1	99.2	0.2:1
65–69	88.0	13.1	99.7	0.2:1
70–74	92.9	12.5	99.9	0.1:1
75–79	96.4	12.0	100.0	0.1:1
80–84	98.5	11.8	100.0	0.1:1
85–89	99.5	11.7	100.0	0.1:1
90–94	99.9	11.6	100.0	0.1:1
95–100	100.0	11.6	100.0	0.1:1

\$3.75/Day 2005 PPP Poverty Line

Figure 4 ($3.75/day 2005 PPP line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	93.7
5–9	92.0
10–14	87.3
15–19	76.0
20–24	64.0
25–29	47.6
30–34	33.4
35–39	19.7
40–44	12.0
45–49	7.8
50–54	4.0
55–59	2.0
60–64	1.5
65–69	0.7
70–74	0.8
75–79	0.1
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Figure 5 ($3.75/day 2005 PPP line): Derivation of estimated poverty likelihoods associated with scores

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	628	÷	670	=	93.7
5–9	1,177	÷	1,279	=	92.0
10–14	1,915	÷	2,193	=	87.3
15–19	2,754	÷	3,622	=	76.0
20–24	3,386	÷	5,291	=	64.0
25–29	3,198	÷	6,725	=	47.6
30–34	2,574	÷	7,701	=	33.4
35–39	1,842	÷	9,336	=	19.7
40–44	1,228	÷	10,277	=	12.0
45–49	770	÷	9,874	=	7.8
50–54	369	÷	9,353	=	4.0
55–59	160	÷	8,002	=	2.0
60–64	111	÷	7,254	=	1.5
65–69	43	÷	6,414	=	0.7
70–74	39	÷	4,862	=	0.8
75–79	2	÷	3,580	=	0.1
80–84	0	÷	2,038	=	0.0
85–89	0	÷	1,028	=	0.0
90–94	0	÷	382	=	0.0
95–100	0	÷	118	=	0.0

Number of all households normalized to sum to 100,000.

Figure 7 ($3.75/day 2005 PPP line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample ($n =$ 16,384) from the validation sample, with confidence intervals, scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	**Diff.**	**90-percent**	**95-percent**	**99-percent**
0–4	+0.4	4.1	4.9	6.6
5–9	+5.1	4.2	5.0	6.5
10–14	+3.7	3.4	4.1	5.7
15–19	+0.2	3.0	3.4	4.5
20–24	+0.8	3.0	3.5	4.7
25–29	+1.6	2.6	3.0	4.0
30–34	+2.3	2.2	2.6	3.6
35–39	–0.4	1.8	2.2	2.8
40–44	–1.1	1.4	1.6	2.2
45–49	+1.3	1.1	1.3	1.7
50–54	+0.5	0.8	1.0	1.3
55–59	–0.4	0.7	0.9	1.1
60–64	+0.2	0.6	0.7	0.9
65–69	–0.5	0.6	0.7	0.9
70–74	+0.4	0.4	0.5	0.6
75–79	–0.2	0.3	0.4	0.5
80–84	+0.0	0.0	0.0	0.0
85–89	+0.0	0.0	0.0	0.0
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 9 ($3.75/day 2005 PPP line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+1.3	57.1	72.1	91.7
4	+0.7	25.5	32.7	42.8
8	+1.0	18.1	21.9	29.6
16	+0.9	12.6	15.3	19.8
32	+0.8	8.6	10.7	13.8
64	+0.8	6.3	7.4	9.7
128	+0.5	4.8	5.3	6.9
256	+0.5	3.2	3.7	5.1
512	+0.5	2.2	2.8	3.6
1,024	+0.5	1.6	1.9	2.6
2,048	+0.4	1.1	1.3	1.8
4,096	+0.5	0.8	0.9	1.2
8,192	+0.4	0.6	0.7	1.0
16,384	+0.5	0.4	0.5	0.6

Figure 11 ($3.75/day 2005 PPP line): Households by targeting classification and score, along with “Total Accuracy” and BPAC, scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.6	19.5	0.0	79.8	80.5	–93.6
5–9	1.7	18.4	0.2	79.7	81.4	–81.6
10–14	3.6	16.5	0.6	79.3	82.9	–61.6
15–19	6.3	13.8	1.4	78.5	84.8	–29.9
20–24	9.7	10.4	3.4	76.5	86.2	+13.1
25–29	12.8	7.3	6.9	72.9	85.8	+62.2
30–34	15.3	4.8	12.1	67.8	83.1	+39.6
35–39	17.3	2.8	19.6	60.3	77.6	+2.7
40–44	18.7	1.5	28.4	51.5	70.1	–41.4
45–49	19.4	0.8	37.6	42.3	61.6	–87.1
50–54	19.7	0.4	46.6	33.3	53.0	–131.8
55–59	19.9	0.2	54.4	25.5	45.4	–170.6
60–64	20.0	0.1	61.6	18.3	38.3	–206.2
65–69	20.1	0.0	67.9	12.0	32.0	–237.8
70–74	20.1	0.0	72.8	7.1	27.2	–261.9
75–79	20.1	0.0	76.3	3.6	23.7	–279.6
80–84	20.1	0.0	78.4	1.5	21.6	–289.7
85–89	20.1	0.0	79.4	0.5	20.6	–294.8
90–94	20.1	0.0	79.8	0.1	20.2	–296.7
95–100	20.1	0.0	79.9	0.0	20.1	–297.3

Inclusion, undercoverage, leakage, and exclusion normalized to sum to 100.

Figure 12 ($3.75/day 2005 PPP line): For a given score cut-off, the percentage of all households who are targeted, the percentage of targeted households who are poor, the percentage of poor households who are targeted, and the number of poor households who are successful targeted (inclusion) per non-poor household mistakenly targeted (leakage), scorecard applied to the validation sample

Targeting cut-off	% all households who are targeted	% targeted who are poor	% of poor who are targeted	Poor households targeted per non-poor household targeted
0–4	0.7	93.0	3.1	13.3:1
5–9	1.9	89.5	8.7	8.5:1
10–14	4.1	86.5	17.8	6.4:1
15–19	7.8	81.6	31.5	4.4:1
20–24	13.1	74.2	48.2	2.9:1
25–29	19.8	64.9	63.8	1.8:1
30–34	27.5	55.8	76.3	1.3:1
35–39	36.8	46.9	85.8	0.9:1
40–44	47.1	39.6	92.8	0.7:1
45–49	57.0	34.0	96.2	0.5:1
50–54	66.3	29.7	98.1	0.4:1
55–59	74.3	26.8	99.0	0.4:1
60–64	81.6	24.5	99.5	0.3:1
65–69	88.0	22.8	99.8	0.3:1
70–74	92.9	21.6	99.9	0.3:1
75–79	96.4	20.9	100.0	0.3:1
80–84	98.5	20.4	100.0	0.3:1
85–89	99.5	20.2	100.0	0.3:1
90–94	99.9	20.1	100.0	0.3:1
95–100	100.0	20.1	100.0	0.3:1

$5.00/Day 2005 PPP Poverty Line

Figure 4 ($5.00/day 2005 PPP line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	99.0
5–9	97.4
10–14	94.3
15–19	90.3
20–24	80.3
25–29	67.5
30–34	53.3
35–39	37.2
40–44	26.0
45–49	20.1
50–54	10.6
55–59	5.6
60–64	3.8
65–69	1.8
70–74	1.3
75–79	0.1
80–84	0.3
85–89	0.0
90–94	0.0
95–100	0.0

Figure 5 ($5.00/day 2005 PPP line): Derivation of estimated poverty likelihoods associated with scores

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	664	÷	670	=	99.0
5–9	1,245	÷	1,279	=	97.4
10–14	2,067	÷	2,193	=	94.3
15–19	3,272	÷	3,622	=	90.3
20–24	4,248	÷	5,291	=	80.3
25–29	4,539	÷	6,725	=	67.5
30–34	4,103	÷	7,701	=	53.3
35–39	3,468	÷	9,336	=	37.2
40–44	2,670	÷	10,277	=	26.0
45–49	1,984	÷	9,874	=	20.1
50–54	995	÷	9,353	=	10.6
55–59	447	÷	8,002	=	5.6
60–64	277	÷	7,254	=	3.8
65–69	117	÷	6,414	=	1.8
70–74	62	÷	4,862	=	1.3
75–79	4	÷	3,580	=	0.1
80–84	6	÷	2,038	=	0.3
85–89	0	÷	1,028	=	0.0
90–94	0	÷	382	=	0.0
95–100	0	÷	118	=	0.0

Number of all households normalized to sum to 100,000.

Figure 7 ($5.00/day 2005 PPP line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) from the validation sample, with confidence intervals, scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	**Diff.**	**90-percent**	**95-percent**	**99-percent**
0–4	+3.8	3.6	4.4	5.5
5–9	–0.1	1.9	2.3	3.2
10–14	+1.0	2.2	2.6	3.5
15–19	+2.3	2.3	2.7	3.6
20–24	+1.1	2.5	3.0	3.9
25–29	+0.7	2.5	3.0	3.8
30–34	+0.6	2.4	2.9	3.5
35–39	–1.7	2.2	2.7	3.4
40–44	–1.5	1.9	2.3	2.8
45–49	+2.4	1.7	2.0	2.8
50–54	+0.5	1.4	1.7	2.0
55–59	+0.0	1.1	1.3	1.7
60–64	+0.5	0.9	1.0	1.4
65–69	–0.5	0.8	1.0	1.3
70–74	+0.5	0.5	0.6	0.7
75–79	–0.5	0.5	0.7	0.8
80–84	+0.1	0.3	0.4	0.5
85–89	+0.0	0.0	0.0	0.0
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 9 ($5.00/day 2005 PPP line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+0.0	63.7	73.7	87.4
4	+0.5	28.3	37.1	50.4
8	+0.7	20.3	24.7	34.7
16	+0.7	15.1	18.1	23.3
32	+0.5	10.4	12.4	16.3
64	+0.5	7.5	8.8	11.6
128	+0.3	5.2	6.0	7.3
256	+0.2	3.6	4.3	6.3
512	+0.3	2.6	3.3	4.1
1,024	+0.3	1.8	2.3	2.9
2,048	+0.2	1.3	1.6	2.0
4,096	+0.2	0.9	1.1	1.5
8,192	+0.2	0.7	0.8	1.0
16,384	+0.3	0.5	0.6	0.7

Figure 11 ($5.00/day 2005 PPP line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.6	29.8	0.0	69.5	70.2	–95.7
5–9	1.9	28.5	0.1	69.5	71.4	–87.4
10–14	3.9	26.5	0.2	69.4	73.3	–73.5
15–19	7.1	23.3	0.6	68.9	76.1	–51.0
20–24	11.4	19.1	1.7	67.9	79.2	–19.8
25–29	15.9	14.5	3.9	65.7	81.6	+17.3
30–34	20.1	10.4	7.4	62.1	82.2	+56.2
35–39	23.8	6.7	13.1	56.5	80.3	+57.1
40–44	26.7	3.8	20.4	49.1	75.8	+32.9
45–49	28.5	2.0	28.5	41.1	69.5	+6.4
50–54	29.5	1.0	36.8	32.7	62.2	–21.0
55–59	30.0	0.5	44.4	25.2	55.1	–45.8
60–64	30.2	0.2	51.4	18.2	48.4	–68.8
65–69	30.4	0.1	57.6	11.9	42.3	–89.3
70–74	30.4	0.0	62.4	7.1	37.5	–105.2
75–79	30.4	0.0	66.0	3.6	34.0	–116.9
80–84	30.4	0.0	68.0	1.5	32.0	–123.5
85–89	30.4	0.0	69.1	0.5	30.9	–126.9
90–94	30.4	0.0	69.4	0.1	30.6	–128.2
95–100	30.4	0.0	69.6	0.0	30.4	–128.6

Inclusion, undercoverage, leakage, and exclusion normalized to sum to 100.

Figure 12 ($5.00/day 2005 PPP line): For a given score cut-off, the percentage of all households who are targeted, the percentage of targeted households who are poor, the percentage of poor households who are targeted, and the number of poor households who are successful targeted (inclusion) per non-poor household mistakenly targeted (leakage), scorecard applied to the validation sample

Targeting cut-off	% all households who are targeted	% targeted who are poor	% of poor who are targeted	Poor households targeted per non-poor household targeted
0–4	0.7	95.2	2.1	20.0:1
5–9	1.9	96.7	6.2	29.6:1
10–14	4.1	95.1	12.9	19.3:1
15–19	7.8	91.9	23.5	11.4:1
20–24	13.1	87.1	37.3	6.7:1
25–29	19.8	80.5	52.3	4.1:1
30–34	27.5	73.0	65.9	2.7:1
35–39	36.8	64.5	78.1	1.8:1
40–44	47.1	56.6	87.6	1.3:1
45–49	57.0	50.0	93.5	1.0:1
50–54	66.3	44.5	96.9	0.8:1
55–59	74.3	40.3	98.4	0.7:1
60–64	81.6	37.0	99.3	0.6:1
65–69	88.0	34.5	99.8	0.5:1
70–74	92.9	32.8	99.9	0.5:1
75–79	96.4	31.6	100.0	0.5:1
80–84	98.5	30.9	100.0	0.4:1
85–89	99.5	30.6	100.0	0.4:1
90–94	99.9	30.5	100.0	0.4:1
95–100	100.0	30.4	100.0	0.4:1

Appendix A:
Guide to Interpretation of Scorecard Indicators

The following information comes from:

Instituto Brasileiro de Geografia e Estatística. (2008) "Pesquisa Nacional por Amostra de Domicílios 2008: Notas Metodológicas, Pesquisa Básica", (the manual").

1. How many members does the household have?

According to p. 24 of the manual, a household is considered to be "those people joined by blood, domestic dependence, or living arrangements, who live in the same residence and who make up a single living unit."

According to p. 14 of the manual, *household members* are "those people that usually live in a residence (be it a single-family or collective residence) and who, on the date of the interview, are present or have been temporarily absent for a period of not longer than 12 months prior to that date."

According to the PNAD database, the household does not include boarders, domestic employees, nor family members of domestic employees.

2. Do any household members ages 5 to 18 go to private school or pre-school?

There is no additional information about this indicator.

3. How many years of schooling has the female head/spouse completed?

According to p. 28 of the manual, the number of years of schooling completed is determined by "the grades that the person is attending or has attended, looking only at the highest level or grade completed. Each grade completed is counted as one year. Having completed first grade is counted as one year. Having completed the first middle cycle is counted as five years. Having completed the middle course of study (second middle cycle) is counted as nine years. Having completed high school is counted as 12 years."

4. <u>How many household members work as employees with a written contract or as civil servants for the government or military?</u>

There is no additional information about this indicator.

5. <u>In their main occupation, how many household members are managers, administrators, professionals in the arts and sciences, mid-level technicians, or clerks?</u>

According to p. 32 of the manual, this is based on the International Standard Classification of Occupations (ISCO–88).

6. <u>How many rooms does the residence have?</u>

There is no additional information about this indicator.

7. <u>How does the household dispose of sewage?</u>

According to pp. 20–21 of the manual, the response concepts are defined as follows:

Public sewage/rainwater system: When waste water and solid wastes are channeled into a collection system that serves to carry sewage away from an area, region, or city, regardless of whether the wastes are eventually subject to treatment

Septic tank connected to public sewage/rainwater system: When waste water and solid wastes are deposited into a septic tank in which they receive treatment before being channeled into a collection system for an area, region, or city

Septic tank not connected to public sewage/rainwater system: When waste water and solid wastes are deposited into a septic tank and are not later channeled into a collection system for an area, region, or city

Simple hole, or directly into river, lake, or ocean: When waste is deposited into a handmade hole (simple pit, well, borehole, etc.) or directly into a river, lake, ocean, or other natural body of water

Ditch, other, or no bathroom: When waste is deposited directly on an open-air ditch, when waste is disposed of in some other manner which is not covered by the other options, or when the household has no fixed place where human waste is deposited

8. Does the household have a refrigerator?

According to p. 23 of the manual, this indicator relates to the existence of "a refrigerator with two doors (that is, an apparatus with two distinct compartments, one for refrigeration of food and the other for freezing food), or a refrigerator with a single compartment."

9. Does the household have a washing machine?

According to p. 23 of the manual, this indicator relates to the existence of "a washing machine, that is, an appliance that automatically performs all the steps of washing clothes, from submersion in water through agitation and water removal via spinning."

10. Does the household have a cellular or land-line telephone?

According to p. 22 of the manual, "register whether the household has a conventional land-line telephone, including telephones that are shared with other households, whether in the residence or not, including "party lines", etc. Also register whether any household member has a cellular telephone."