Simple Poverty Scorecard®
Burkina Faso

Mark Schreiner

2 May 2011

Ce document est disponible en Français sur `SimplePovertyScorecard.com`
This document and related tools is available at `SimplePovertyScorecard.com`

Abstract

The Simple Poverty Scorecard® uses eight low-cost indicators from Burkina Faso's 2003 Household Living Standards Survey to estimate the likelihood that a household has consumption below a given poverty line. Field workers can collect responses in about ten minutes. The scorecard's accuracy is reported for a range of poverty lines. The scorecard is a practical way for pro-poor programs in Burkina Faso to measure poverty rates, to track changes in poverty rates over time, and to segment clients for differentiated treatment.

Acknowledgements

Catholic Relief Services funded this paper via a grant to Grameen Foundation (GF). Data are from the *Institut National de la Statistique et de la Démographie*. Thanks go to Sharlene Brown, Frank DeGiovanni, Boubacar Diallo, Mary Jo Kochendorfer, Jeff Toohig, and Preeti Wali. The Simple Poverty Scorecard® is the same as what GF calls the Progress out of Poverty Index®. The PPI® is a performance-management tool that GF promotes to help organizations achieve their social objectives more effectively.

Author

Mark Schreiner directs Microfinance Risk Management, L.L.C. He is also Senior Scholar at the Center for Social Development at Washington University in Saint Louis.

Simple Poverty Scorecard®

			Name	Identifier
Interview ID:				
Interview date:		Participant:		
Country:	BFA	Field agent:		
Scorecard:	001	Service point:		
Sampling wgt.:		Number of household members:		

Indicator	Value	Points	Score
1. How many household members are 14-years-old or younger?	A. Six or more	0	
	B. Five	5	
	C. Four	6	
	D. Three	10	
	E. Two	13	
	F. One	19	
	G. None	29	
2. In what languages can the male head/spouse read and write?	A. None, or no male head/spouse	0	
	B. French only	4	
	C. A non-French language (regardless of French literacy)	5	
3. Has the female head/spouse completed first grade?	A. No	0	
	B. No female head/spouse	0	
	C. Yes	9	
4. What is the main source of energy for lighting?	A. Firewood, or other	0	
	B. Candles, kerosene, or LPG	4	
	C. Flashlight, or batteries	5	
	D. Electricity, or solar energy	8	
5. What toilet arrangement does the household have?	A. No toilet arrangement, or other	0	
	B. Non-ventilated pit latrine	4	
	C. Ventilated pit latrine, or flush to a septic tank	15	
6. Does the household own a television?	A. No	0	
	B. Yes	10	
7. Does the household own a bed or a mattress?	A. No	0	
	B. Yes	3	
8. Does the household own a scooter or a motorcycle?	A. No	0	
	B. Yes	6	
9. Have any household members, in their main occupation in the last seven days, worked in agriculture, animal husbandry, fishing, or forestry?	A. Yes	0	
	B. No	8	
10. How many head of cattle or other large animals does the household now own?	A. None, or one	0	
	B. Two	2	
	C. Three to five	3	
	D. Six or more	7	

Score:

Simple Poverty Scorecard®
Burkina Faso

1. Introduction

This paper presents the Simple Poverty Scorecard®. Pro-poor programs in Burkina Faso can use it to estimate the likelihood that a household has consumption below a given poverty line, to estimate a population's poverty rate at a point in time, to track changes in a population's poverty rate over time, and to segment participants for differentiated treatment.

The direct approach to poverty measurement via surveys is difficult and costly. As a case in point, Burkina Faso's 2003 Household Living Standards Survey (EBCVM, *Enquête Burkinabé sur les Conditions de Vie des Ménages*) runs 28 pages and asks about more than 100 consumption items. For example, "In the last 15 days, has the household consumed any rice? What is the value of purchases of rice consumed and of stocks of rice purchased? What is the value of gifts of rice received and consumed? What is the value of self-produced rice consumed? Now then, has the household consumed any millet? . . . "

In contrast, the indirect approach via the Simple Poverty Scorecard® is simple, quick, and inexpensive. It uses ten verifiable indicators (such as "What toilet arrangement does the household have?" and "Does the household own a television?") to

get a score that is highly correlated with poverty status as measured by the exhaustive survey.

The Simple Poverty Scorecard® 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 these organizations are typically subjective and relative (such as participatory wealth ranking by skilled field agents) or blunt (such as rules based on land-ownership or housing quality). Measurements from these approaches are not comparable across organizations, they may be costly, and their accuracy is unknown.

Pro-poor organizations can use the scorecard to measure the share of their participants who are below a given poverty line, such as the Millennium Development Goals' $1.25/day line at 2005 purchase-power parity. USAID microenterprise partners can use poverty scoring to report how many of their participants are among the poorest half of people below the national poverty line. Organizations can also use the tool to measure movement across a poverty line. In all these cases, the scorecard provides an consumption-based, objective tool with known accuracy. While consumption surveys are costly even for governments, many small, local organizations may be able to implement an inexpensive poverty-assessment tool that can serve for monitoring and targeting.

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, not because they do not work, but because they are presented (when they are presented at all) as tables of regression coefficients incomprehensible to non-specialists (with cryptic indicator names such as "`LGHHSZ_2`", negative values, and many decimal places). Thanks to the predictive-modeling phenomenon known as the "flat maximum", simple poverty-assessment tools are often about as accurate as complex ones.

The technical approach here is also 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 techniques are simple and commonplace in statistical practice and in for-profit credit-risk scoring, they have rarely or never been applied to poverty-assessment tools.

The scorecard is based on the 2003 EBCVM conducted by Burkina Faso's *Institut National de la Statistique et de la Démographie* (INSD). 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 non-negative 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 consumption below a given poverty line.

Second, poverty scoring can be used to estimate the poverty rate of a group of households at a point in time. This estimate is simply the average of the poverty likelihoods among the households in the group.

Third, poverty scoring can be used to estimate changes in the poverty rate for a group of households (or for two independent samples of households that are representative of the same population) between two points in time. This estimate is the change in the average poverty likelihood of the group(s) of households over time.

Poverty scoring can also be used for targeting. To help managers choose an appropriate targeting cut-off for their purposes, 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 household consumption data and Burkina Faso's national poverty line. Scores from this one scorecard are calibrated to poverty likelihoods for seven poverty lines.

The scorecard is constructed and calibrated using half of the data from the 2003 EBCVM, and its accuracy is validated on the other half of the data.

While all three scoring estimators are *unbiased* (that is, they match the true value on average in repeated samples when applied to the same population from which

the scorecard was built), they are—like all predictive models—biased to some extent when applied to a different population.[1]

Thus, while the indirect scoring approach is less costly than the direct survey approach, it is also biased. There is bias because scoring must assume that the future relationships between indicators and poverty will be the same as in the data used to build the scorecard.[2] Of course, this assumption—ubiquitous and inevitable in predictive modeling—holds only partly.

When applied to the validation sample with bootstrap samples of n = 16,384, the difference between scorecard estimates of groups' poverty rates and the true rates at a point in time is –0.3 percentage points for the national line, and the average absolute difference across all seven lines is 0.2 percentage points. These differences are due to sampling variation and not bias; the average of each difference would be zero if the whole 2003 EBCVM were to be repeatedly redrawn and divided into sub-samples before repeating the entire process of building, calibrating, and validating scorecards.

The 90-percent confidence intervals for these estimates are ±0.6 percentage points or less. For n = 1,024, these intervals are ±2.6 percentage points or less.

[1] Important cases include nationally representative samples at a different point in time or non-nationally representative sub-groups (Tarozzi and Deaton, 2007).
[2] Bias may also result from changes in the quality of data collection, from changes to poverty lines, or from imperfect adjustment of poverty lines to account for differences in cost-of-living across time or regions.

Section 2 below describes data and poverty lines. Sections 3 and 4 describe scorecard construction and offer guidelines for use in practice. 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 over time, and Section 8 covers targeting. Section 9 places the new scorecard here in the context of existing exercises for Burkina Faso, and Section 10 is a summary.

2. Data and poverty lines

This section discusses the data used to construct and test the Simple Poverty Scorecard®. It also presents the poverty lines to which scores are calibrated.

2.1 Data

The scorecard is based on data from the 8,500 households in the EBCVM. Conducted from 10 May 2003 to 15 July 2003, this is Burkina Faso's most recent available national consumption survey.

For the purposes of poverty scoring, the households in the 2003 EBCVM are randomly divided into two sub-samples (Figure 2):

- *Construction* and *calibration* for selecting indicators and points and for associating scores with poverty likelihoods
- *Validation* for testing 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 group who live in households whose total household consumption (divided by the number of household 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 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 count more.

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 consumption above a poverty line (it is "non-poor") and that the second household has per-capita consumption below a poverty line (it is "poor"). The household-level rate counts both households as if they had only one person and so gives a poverty rate 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 of 2 $\div$ (1 + 2) = 67 percent.

Whether the household-level rate or the person-level rate is 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 people, regardless of how 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 may be relevant. For example, if a microlender has only one borrower in a household, then it might prefer to report household-level poverty rates.

Figure 2 reports poverty lines and poverty rates for Burkina Faso at both the household-level and the person-level. The scorecard is constructed using the 2003 EBCVM and household-level lines, scores are calibrated to household-level poverty

likelihoods, and accuracy is measured for household-level rates. This assumes that the household level is relevant for most pro-poor organizations.

Person-level poverty rates can be estimated by taking a household-size-weighted average of the household-level poverty likelihoods. It is possible to construct, calibrate, and validate a scorecard based on person-level weights, but it is not done here.

2.2.2 Poverty lines

Burkina Faso's national poverty line (sometimes called here "100% of the national line") is XOF226 per person per day. It is derived—using data from the 1994/5 Priorities Survey (*Enquête Prioritaire I*)—as the cost of a food basket with 2,283 Calories (INSD, 2003), plus (it is supposed) the non-food consumption observed for households whose food consumption is close to the caloric benchmark. The 1994/5 line is then adjusted for changes in prices up to 2003. It is not adjusted for cost-of-living differences across geographic regions. This national line gives a household-level poverty rate of 37.5 percent and a person-level poverty rate of 46.4 percent (Figure 2).

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

- 50% of national
- 75% of national
- 100% of national
- 150% of national
- USAID "extreme"
- $1.25/day 2005 PPP
- $2.50/day 2005 PPP

The USAID "extreme" line is defined as the median consumption of people (not households) below the national line (U.S. Congress, 2004).

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

- 2005 PPP exchange rate for "individual consumption expenditure by households" (World Bank, 2008): XOF242.42 per $1.00
- Average monthly Consumer Price Index for April through July of 2003 of 117.625[3]
- Average monthly CPI for 2005 of 123.583

Given this, the $1.25/day 2005 PPP line for Burkina Faso as a whole on average during the 2003 EBCVM is (Sillers, 2006):

$$(\text{2005 PPP exchange rate}) \cdot \$1.25 \cdot \left(\frac{\text{CPI}_{\text{Apr. 03-Jul. 03}}}{\text{CPI}_{\text{2005 average}}} \right) =$$

$$\left(\frac{\text{XOF242.42}}{\text{USD1.00}} \right) \cdot \$1.25 \cdot \left(\frac{117.625}{123.583} \right) = \text{XOF288.}$$

The $2.50/day 2005 PPP line is twice the $1.25/day line.

[3] INSD, `http://www.insd.bf/fr/Tableaux/T2203.htm`, retrieved 22 March 2011.

3. Scorecard construction

For Burkina Faso, about 85 potential indicators are initially prepared in the areas of:

- Family composition (such as household size)
- Education (such as literacy of the male head/spouse)
- Employment (such as whether any household member works in agriculture)
- Housing (such as wall material)
- Ownership of durable goods (such as televisions)
- Agriculture (such as ownership of cattle or other large animals)

Figure 3 lists all potential indicators, ranked by the entropy-based "uncertainty coefficient" that measures how well a given indicator predicts poverty on its own (Goodman and Kruskal, 1979).

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 bed or mattress 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 national poverty line and Logit regression on the construction/calibration sub-sample. Indicator selection uses both judgment and statistics. 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). These include 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.

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).

This algorithm is the Logit analogue to the common R^2-based stepwise 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 single scorecard here applies to all of Burkina Faso. Evidence from India and Mexico (Schreiner, 2006 and 2005a), Sri Lanka (Narayan and Yoshida, 2005), and Jamaica (Grosh and Baker, 1995) suggests that segmenting scorecards by urban/rural does not improve targeting accuracy much, although it may improve the accuracy of estimates of poverty rates for sub-groups (Tarozzi and Deaton, 2007).

4. Practical guidelines for scorecard use

The main challenge of scorecard design is not to maximize statistical accuracy but rather to improve the chances that the scorecard is actually adopted and used in practice (Schreiner, 2005b). When scoring projects fail, the reason is not usually statistical 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 about the same, 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 adopt it and use it properly. Of course, accuracy matters, but it is balanced against 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 make a lot of "extra" work and if the whole process generally seems to make sense.

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

- Only 10 indicators
- Only categorical indicators
- Simple weights (non-negative integers, no arithmetic beyond addition)

The scorecard is ready to be photocopied. A field agent using the paper scorecard would:

- Record participant identifiers and household size
- Read each question from the scorecard
- Circle the response and its point value
- Write the point value 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 filing or data entry and analysis

Of course, field agents must be trained. Quality outputs depend on quality inputs. If organizations or field agents gather their own data and 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 audits (Matul and Kline, 2003).[4] IRIS Center (2007a) and Toohig (2008) are useful nuts-and-bolts guides for budgeting, training field agents and supervisors, logistics, sampling, interviewing, piloting, recording data, and controlling quality.

[4] If an organization does not want field agents to know the points associated with indicators, then they can use the version of the scorecard without points and apply the points later in a spreadsheet or database at the central office.

In particular, while collecting scorecard indicators is relatively easier than alternatives, it is still absolutely difficult. Training and explicit definitions of terms and concepts in the scorecard is essential (see Appendix). For the example of Nigeria, Onwujekwe, Hanson, and Fox-Rushby (2006) found distressingly low inter-rater and test-retest correlations for indicators as seemingly simple and obvious as whether the household owns a car. At the same time, Grosh and Baker (1995) find that gross underreporting of assets does not affect targeting. For the first stage of targeting in a Mexican social program, 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 is done in Mexico in the second stage of its targeting process, most false self-reports can be corrected by field agents who verify responses with a home visit, and this is suggested for Burkina Faso.

In terms of sampling 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

In general, the sampling design should follow from the organization's goals for the exercise and the business questions that it seeks to inform.

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 downloaded to a database

Given a group of interest for a given question, 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

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 level of confidence and a desired confidence interval.

Frequency of application can be:

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

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

- With a different set of participants
- With the same set of participants

An example set of choices is illustrated by BRAC and ASA, two microlenders in Bangladesh. Each has more than 7 million participants, and each has stated their inention to use the Simple Poverty Scorecard® (Schreiner, 2013). Their design is that loan officers in a random sample of branches score all their participants each time they visit a homestead (about once a year) as part of their standard due diligence prior to loan disbursement. Responses are recorded on paper in the field before being sent to a central office to be entered into a database. ASA's and BRAC's sampling plans cover more than 50,000 participants (far more than most pro-poor organizations would need).

5. Estimates of household poverty likelihoods

The sum of scorecard points for a household is called the *score*. For Burkina Faso, 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 line, the scores themselves have only relative units. For example, doubling the score increases the likelihood of being above a given poverty line, but it does not double the likelihood.

To get absolute units, scores are 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 national line, scores of 15–19 have a poverty likelihood of 57.2 percent, and scores of 20–24 have a poverty likelihood of 43.7 percent (Figure 4).

The poverty likelihood associated with a score varies by poverty line. For example, scores of 15–19 are associated with a poverty likelihood of 57.2 percent for the national line but 72.4 percent for the $1.25/day 2005 PPP line.[5]

5.1 Calibrating scores with poverty likelihoods

A given score is 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.

[5] From Figure 4 on, many figures have seven versions, one for each of seven poverty lines. To keep them straight, they are grouped by poverty line. Single tables that pertain to all poverty lines are placed with the first group of tables for the national line.

For the example of the national line (Figure 5), there are 15,736 (normalized) households in the calibration sub-sample with a score of 15–19, of whom 9,004 (normalized) are below the poverty line. The estimated poverty likelihood associated with a score of 15–19 is then 57.2 percent, because 9,004 ÷ 15,736 = 57.2 percent.

To illustrate with the national line and a score of 20–24, there are 16,680 (normalized) households in the calibration sample, of whom 7,288 (normalized) are below the line (Figure 5). Thus, the poverty likelihood for this score is 7,288 ÷ 16,680 = 43.7 percent.

The same method is used to calibrate scores with estimated poverty likelihoods for the other six poverty lines.

Figure 6 shows, for all scores, the likelihood that consumption falls in a range demarcated by two adjacent poverty lines. For example, the daily consumption of someone with a score of 15–19 falls in the following ranges with probability:

- 11.3 percent below 50% of the national line
- 15.0 percent between 50% of the national line and the USAID "extreme" line
- 9.4 percent between the USAID "extreme" line and 75% of the national line
- 21.6 percent between 75% of the national line and 100% of the national line
- 15.2 percent between 100% of the national line and the $1.25/day 2005 PPP line
- 9.7 percent between the $1.25/day 2005 PPP line and 150% of the national line
- 13.7 percent between 150% of the national line and the $2.50/day 2005 PPP line
- 4.3 percent above the $2.50/day 2005 PPP line

Even though the scorecard is constructed partly based on judgment, the calibration process produces poverty likelihoods that are objective, that is, derived from survey data on consumption 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 constructed using only expert 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 the Burkina Faso 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. Non-specialists find it 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. In the field, going from scores to poverty likelihoods in this way requires no arithmetic at all, just a look-up table. This calibration can also improve accuracy, especially with large 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 that are representative of the same population from which the scorecard was constructed, then 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 value. 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.[6]

Of course, the relationship between indicators and poverty does change to some unknown extent with time and also across sub-groups in Burkina Faso's population. Thus, the scorecard will generally be biased when applied after July 2003 (when fieldwork for the 2003 EBCVM ended) or when applied with non-nationally representative sub-groups.

How accurate are estimates of households' poverty likelihoods? To measure, 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 and consumption 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.

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

For the national line, the average poverty likelihood across bootstrap samples for scores of 15–19 in the validation sample is too low by 0.6 percentage points. For scores of 20–24, the estimate is too high by 3.2 percentage points.[7]

The 90-percent confidence interval for the differences for scores of 15–19 is ±1.6 percentage points (Figure 7). This means that in 900 of 1,000 bootstraps, the difference between the estimate and the true value is between –2.2 and +1.0 percentage points (because –0.6 – 1.6 = –2.2, and –0.6 + 1.6 = +1.0). In 950 of 1,000 bootstraps (95 percent), the difference is –0.6 ±1.9 percentage points, and in 990 of 1,000 bootstraps (99 percent), the difference is –0.6 ±2.6 percentage points.

For most scores, Figure 7 shows some differences between estimated poverty likelihoods and true values. This is 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 Burkina Faso's population. For targeting, however, what matters is less the difference in all score ranges and more the difference in score ranges just above and 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.

[7] These differences are not zero, in spite of 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 process of scorecard construction and calibration.

In addition, if estimates of groups' poverty rates are to be usefully accurate, then errors for individual households must largely cancel out. This is generally the case, as discussed in the next section.

Another possible source of differences between estimates and true values is overfitting. By construction, the scorecard here is unbiased, but it may still be *overfit* when applied after the end of the EBCVM fieldwork in July 2003. That is, it may fit the 2003 EBCVM construction data so closely that it captures not only some timeless patterns but also some random patterns that, due to sampling variation, show up only in the construction data. Or the scorecard may be overfit in the sense that its bias is highly sensitive to changes over time in the relationship between indicators and poverty when it is applied to non-nationally representative samples.

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. Combining scorecards can also help, at the cost of greater complexity.

Most errors in individual households' likelihoods, however, cancel out in the estimates of groups' poverty rates (see later sections). Furthermore, at least some of the differences arise from non-scorecard sources such as changes in the relationship between indicators and poverty, sampling variation, changes in poverty lines, inconsistencies in data quality across time, and imperfections in cost-of-living adjustments across time and geographic regions. These factors can be addressed only by improving data quantity and quality (which is beyond the scope of the scorecard) or by reducing overfitting (which likely has limited returns, given the scorecard's parsimony).

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, 2011 and that they have scores of 20, 30, and 40, corresponding to poverty likelihoods of 43.7, 20.7, and 14.4 percent (national line, Figure 4). The group's estimated poverty rate is the households' average poverty likelihood of $(43.7 + 20.7 + 14.4) \div 3 = 26.3$ percent.[8]

6.1 Accuracy of estimated poverty rates at a point in time

For the Burkina Faso scorecard applied to the validation sample with $n =$ 16,384, the absolute differences between the estimated poverty rate at a point in time and the true rate are 0.3 percentage points or less (Figure 9, summarizing Figure 8 across poverty lines). The average absolute difference across the seven poverty lines is 0.2 percentage points. Part of these differences is due to sampling variation and the division of the 2003 EBCVM into two sub-samples.

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 ± 0.6 percentage points or less (Figure 9). This means that in 900 of 1,000 bootstraps of this size, the difference between the

[8] The group's poverty rate is *not* the poverty likelihood associated with the average score. Here, the poverty likelihood associated with the average score of 30 is 20.7 percent, which differs from the average of the three poverty likelihoods associated with each of the three scores (26.3 percent).

estimate and the true value is within 0.6 percentage points of the average difference. In the specific case of the national 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 –0.3 – 0.6 = –0.9 to –0.3 + 0.6 = +0.3 percentage points. This is because –0.3 is the average difference, and ±0.6 is its 90-percent confidence interval. The average difference is –0.3 because the average scorecard estimate is too low by 0.3 percentage points; the average estimated poverty rate for the validation sample is 37.1 percent, but the true value is 37.4 percent (Figure 2).

6.2 Formula for standard errors for estimates of poverty rates

How precise are the point-in-time estimates? Because they are averages of binary (0/1, or poor/non-poor) variables, the estimates (in "large" samples) have a Normal distribution and can be characterized by their average difference vis-à-vis true values, together 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 from indirect measurement via scorecards (Schreiner, 2008a), note that the textbook formula (Cochran, 1977) that relates confidence intervals with standard errors in the case of direct measurement of a proportion 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, $\sqrt{\frac{p \cdot (1-p)}{n}}$,

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

n is the sample size.

For example, this implies that for a sample n of 16,384 with 90-percent confidence ($z = 1.64$) and a poverty rate p of 37.6 percent (the poverty rate in the construction and calibration samples in Figure 2 for the national line), the confidence interval c is $+/- z \cdot \sqrt{\frac{p \cdot (1-p)}{n}} = +/- 1.64 \cdot \sqrt{\frac{0.376 \cdot (1-0.376)}{16{,}384}} = \pm 0.621$ percentage points.

Poverty scoring, however, does not measure poverty directly, so this formula is not immediately applicable. To derive a formula for the Burkina Faso scorecard, consider Figure 8, which reports empirical confidence intervals c for the differences for the scorecard applied to 1,000 bootstrap samples of various sizes from the validation

sample. For $n = 16{,}384$ and the national line, the 90-percent confidence interval is 0.635 percentage points.[9]

Thus, the 90-percent confidence interval with $n = 16{,}384$ is 0.635 percentage points for the Burkina Faso scorecard and 0.621 percentage points for direct measurement. The ratio of the two intervals is $0.635 \div 0.621 = 1.02$.

Now consider the same case, but with $n = 8{,}192$. The confidence interval under direct measurement is $+/-1.64 \cdot \sqrt{\frac{0.376 \cdot (1-0.376)}{8{,}192}} = \pm 0.878$ percentage points. The empirical confidence interval with the Burkina Faso scorecard (Figure 8) is 0.880 percentage points. Thus for $n = 8{,}192$, the ratio of the two intervals is $0.880 \div 0.878 = 1.00$.

This ratio of 1.00 for $n = 8{,}182$ is not far from the ratio of 1.02 for $n = 16{,}384$. Across all sample sizes of 256 or more in Figure 8, the average ratio turns out to be 1.03, implying that confidence intervals for indirect estimates of poverty rates via the Burkina Faso scorecard and this poverty line are slightly wider than for direct estimates via the 2003 EBCVM. This 1.03 appears in Figure 9 as the "α factor" because if $\alpha = 1.03$, then the formula relating confidence intervals c and standard errors σ for the Burkina Faso scorecard is $c = +/- z \cdot \alpha \cdot \sigma$. That is, formula for the standard error σ for point-in-time estimates of poverty rates via scoring is $\alpha \cdot \sqrt{\frac{p \cdot (1-p)}{n}}$.

[9] Due to rounding, Figure 8 displays 0.6, not 0.635.

In general, α can be more or less than 1.00. When α is more than 1.00, it means that the scorecard is less precise than direct measurement. This occurs for four of seven poverty lines in Figure 9.

The formula relating confidence intervals with standard errors for poverty scoring can be rearranged to give a formula for determining sample size before measurement.[10] If $\hat{p}$ is the expected poverty rate before measurement, then the formula for sample size n based on the desired confidence level that corresponds to z and the desired confidence interval $\pm c$ is $n = \left(\frac{\alpha \cdot z}{c}\right)^2 \cdot \hat{p} \cdot (1 - \hat{p})$.

To illustrate how to use this, suppose $c = 0.05275$ and $z = 1.64$ (90-percent confidence). Then the formula gives $n = \left(\frac{1.03 \cdot 1.64}{0.05275}\right)^2 \cdot 0.376 \cdot (1 - 0.376) = 241$, in the same ballpark as the sample size of 256 observed for these parameters in Figure 8 for the national line.

Of course, the α factors in Figure 9 are specific to Burkina Faso, its poverty lines, its poverty rates, and this scorecard. The derivation of the formulas, however, is valid for any poverty-assessment tool following the approach in this paper.

[10] IRIS Center (2007a and 2007b) says that a sample size of $n = 300$ is sufficient for USAID reporting. If a poverty-assessment tool 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 poverty-assessment tool could be more or less precise than direct measurement.

In practice after the end of fieldwork for the EBCVM in July 2003, an organization would select a poverty line (say, the national 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 37.5 percent national average in the 2003 EBCVM in Figure 2), look up α (here, 1.03), assume that the scorecard will still work in the future and/or for non-nationally representative sub-groups,[11] and then compute the required sample size. In this illustration, $n = \left(\frac{1.03 \cdot 1.64}{0.02}\right)^2 \cdot 0.375 \cdot (1 - 0.375) = 1{,}672$.

[11] 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 after July 2003 will resemble that in the 2003 EBCVM with deterioration to the extent that the relationships between indicators and poverty status change over time.

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 only from the 2003 EBCVM, this paper cannot test estimates of change over time for Burkina Faso, and it can only suggest approximate formulas for standard errors. Nevertheless, the relevant concepts are presented here because, in practice, pro-poor organizations can apply the scorecard to collect their own data and measure change through time.

7.1 Warning: Change is not impact

Scoring can estimate change. Of course, poverty could get better or worse, and scoring does not indicate what caused change. This point is often forgotten or confused, 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 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, 2011, a program samples three households who score 20, 30, and 40 and so have poverty likelihoods of 43.7, 20.7, and 14.4 percent (national line, Figure 4). The group's baseline estimated poverty rate is the households' average poverty likelihood of (43.7 + 20.7 + 14.4) ÷ 3 = 26.3 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, 2012, 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 25, 35, and 45 (poverty likelihoods of 34.6, 17.0, and 4.2 percent, national line, Figure 4). Their average poverty likelihood at follow-up is now (34.6 + 17.0 + 4.2) ÷ 3 = 18.6 percent, an improvement of 26.3 – 18.6 = 7.7 percentage points.[12]

This suggests that about one in 13 participants in this hypothetical example crossed the poverty line in 2011.[13] Among those who started below the line, about one in four (7.7 ÷ 26.3 = 29.3 percent) on net ended up above the line.[14]

[12] Of course, such a huge reduction in poverty in one year is unlikely, but this is just an example to show how poverty scoring can be used to estimate change.
[13] This is a net figure; some people start above the line and end below it, and vice versa.
[14] Poverty scoring does not reveal the reasons for this change.

7.3 Accuracy for estimated change in two independent samples

With only the 2003 EBCVM, it is not possible to measure the accuracy of scorecard estimates of changes in groups' poverty rates over time. In practice, of course, local pro-poor organizations can still apply the Burkina Faso scorecard to estimate change. The rest of this section suggests approximate formulas for standard errors and sample sizes that may be used until there is additional data.

For two equal-sized independent samples, the same logic as above can be used to derive a formula relating the confidence interval c with the standard error σ of a 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 above, n is the sample size at both baseline and follow-up,[15] and α is the average (across a range of bootstrapped sample sizes) of the ratio of the observed confidence interval from a scorecard and the theoretical confidence interval under direct measurement.

[15] This means that, for a given precision and with direct measurement, estimating the change in a poverty rate between two points in 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 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, 2010, 2009a, 2009b, 2009c, 2009d, 2009e, and 2008b; Schreiner and Woller, 2010a and 2010b; and Chen and Schreiner, 2009), the simple average of α across poverty lines and years for a given country and then across countries is 1.19. This is as reasonable a figure as any to use for Burkina Faso.

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 national line, $\alpha = 1.19$, and $\hat{p} = 0.375$ (from Figure 2). Then the baseline sample size is $n = 2 \cdot \left(\frac{1.19 \cdot 1.64}{0.02} \right)^2 \cdot 0.375 \cdot (1 - 0.375) =$ 4,464, and the follow-up sample size is also 4,464.

7.4 Accuracy for estimated change for one sample, scored twice

Analogous to previous derivations, the general formula relating the confidence interval c to the standard error σ when using a scorecard to estimate change for a single group of households, all of whom are scored at two points in time, is:[16]

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

where z, c, and α are defined as usual, 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.

The formula for standard errors can be rearranged to give a formula for sample size 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 change in the poverty rate will be zero, which implies $\hat{p}_{12} = \hat{p}_{21} = \hat{p}_*$, giving:

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

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

Because $\hat{p}_*$ could be anything between 0–0.5, more information is needed to apply 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 Burkina Faso scorecard is applied twice (once after July 2003 and then again later) is

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

In Peru (the only other country for which there is an 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 national line, and the sample will first be scored in 2011 and then again in 2014 ($y = 3$). The before-baseline poverty rate is 37.5 percent ($p_{2003} = 0.375$, Figure 2), and suppose $\alpha = 1.30$. Then the baseline sample size is $n = 2 \cdot \left(\frac{1.30 \cdot 1.64}{0.02} \right)^2 \cdot \{-0.02 + 0.016 \cdot 3 + 0.47 \cdot [0.375 \cdot (1 - 0.375)]\} = 3{,}140$. The same group of 3,140 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 consumption below a poverty line). Poverty status is a fact that depends on whether consumption 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 with the cut-off score; a higher cut-off has better inclusion (but greater leakage), while a lower cut-off has better exclusion (but higher undercoverage).

Programs 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 15–19, outcomes for the national line in the validation sample are:

- Inclusion: 21.7 percent are below the line and correctly targeted
- Undercoverage: 15.8 percent are below the line and mistakenly not targeted
- Leakage: 12.6 percent are above the line and mistakenly targeted
- Exclusion: 50.0 percent are above the line and correctly not targeted

Increasing the cut-off to 20–24 improves inclusion and undercoverage but worsens leakage and exclusion:

- Inclusion: 28.4 percent are below the line and correctly targeted
- Undercoverage: 9.0 percent are below the line and mistakenly not targeted
- Leakage: 22.5 percent are above the line and mistakenly targeted
- Exclusion: 40.1 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. A 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 correctly included or correctly 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 the Burkina Faso scorecard. For the national line in the validation sample, total net benefit is greatest (71.7) for a cut-off of 19 or less, with more than two in three households in Burkina Faso 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 valued 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).[17]

[17] Figure 11 also reports "BPAC", the Balanced Poverty Accuracy Criteria adopted by USAID for certifying poverty-assessment tools. IRIS Center (2005) says that BPAC considers accuracy in terms of the estimated poverty rate and in terms of targeting inclusion. 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 benefit, 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, for the Burkina Faso scorecard applied to the validation sample, the expected poverty rate among households who score at or below a given cut-off. For the example of the national line, targeting households who score 19 or less would target 34.3 percent of all households (second column) and lead to a poverty rate among those targeted of 63.3 percent (third column).

Figure 12 also reports two other measures of targeting accuracy. The first is a version of coverage ("% of poor who are targeted"). For the example of the national line in the validation sample and a cut-off of 19 or less, 57.9 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 national line in the validation sample and a cut-off of 19 or less, covering 1.7 poor households means leaking to 1 non-poor household.

9. Context of poverty-assessment tools in Burkina Faso

This section discusses an existing poverty-assessment tool for Burkina Faso.

Gwatkin *et al.* (2007) apply to Burkina Faso 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 9,097 households in Burkina Faso's 2003 DHS. The PCA index is like the poverty scorecard here except that, because the DHS does not collect data on income or consumption, it is based on a different conception of poverty, its accuracy vis-à-vis consumption-based poverty is unknown, and it can only be assumed to be a proxy for long-term wealth/economic status.[18] Well-known examples of the PCA asset-index approach include Ferguson *et al.* (2003), Sahn and Stifel (2000 and 2003), and Filmer and Pritchett (2001).

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.

[18] Still, because the indicators are similar and because the "flat maximum" is important, carefully built PCA indices and consumption-based poverty-assessment tools may pick up the same underlying construct (perhaps "permanent income", see Bollen, Glanville, and Stecklov, 2007). Tests of how well rankings by PCA indices correspond with rankings by consumption-based poverty-assessment tools include Howe *et al.* (2009), Filmer and Scott (2008), Lindelow (2006), Wagstaff and Watanabe (2003), and Montgomery *et al.* (2000).

The 14 indicators in Gwatkin *et al.* are similar to those in the scorecard here in terms of their simplicity, low cost, and verifiability:

- Characteristics of the residence:
  - — Presence of electricity
  - — Type of floor
  - — Type of cooking fuel
  - — Source of drinking water
  - — Type of toilet arrangement
- Ownership of consumer durables:
  - — Radios
  - — Televisions
  - — Refrigerators
  - — Telephones
  - — Bicycles
  - — Motorcycles or scooters
  - — Cars or trucks
- Presence of a domestic worker not related to the head
- Whether any family members work agricultural land owned by the family

In practice, the Gwatkin *et al.* index is more difficult and costly because it cannot be computed by hand in the field, as it has 66 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, consumption-based poverty line. Thus, while both approaches can rank households, only the Simple Poverty Scorecard® can estimate consumption-based poverty status.

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

10. Conclusion

Pro-poor programs in Burkina Faso can use the Simple Poverty Scorecard® to segment clients for differentiated treatment as well as to estimate:

- The likelihood that a household has consumption below a given poverty line
- The poverty rate of a population at a point in time
- The change in the poverty rate of a population over time

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

The scorecard is built with a sub-sample of data from the 2003 EBCVM, tested with a different sub-sample, and calibrated to seven poverty lines.

Accuracy and precision are 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 change are not the same as estimates of program impact. Targeting accuracy is also reported.

When the scorecard is applied to the validation sample with $n = 16{,}384$, the absolute difference between estimates versus true poverty rates for groups of households at a point in time is 0.3 percentage points or less and averages—across the seven poverty lines—about 0.2 percentage points. For $n = 16{,}384$ and 90-percent confidence, the precision of these differences is ± 0.6 percentage points or better.

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

Although the statistical technique is innovative, and although technical accuracy is important, the design of the scorecard here 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 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 Burkina Faso to estimate poverty rates, track changes in poverty rates over time, and target services. The same approach can be applied to any country with similar data.

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.

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", `microfinance.com/English/Papers/Scoring_SMEs_Hybrid.pdf`, retrieved 25 March 2011.

Chen, Shiyuan; and Mark Schreiner. (2009) "Simple Poverty Scorecard®: Vietnam", `microfinance.com/English/Papers/Scoring_Poverty_Vietnam_EN_2006.pdf`, retrieved 25 March 2011.

Coady, David; Grosh, Margaret; and John Hoddinott. (2002) "The Targeting of Transfers in Developing Countries: Review of Experience and Lessons", `info.worldbank.org/etools/docs/library/79646/Dc%202003/courses/dc2003/readings/targeting.pdf`, retrieved 25 March 2011.

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

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

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

Falkenstein, Eric. (2008) "DefProb™: A Corporate Probability-of-Default Model", `papers.ssrn.com/sol3/papers.cfm?abstract_id=1103404`, retrieved 25 March 2011.

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*.

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", World Bank Policy Research Working Paper No. 4605, Washington, D.C., `papers.ssrn.com/sol3/papers.cfm?abstract_id=1149108`, retrieved 25 March 2011.

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", `microfinance.com/English/Papers/Scoring_Poverty_Haiti_Fuller.pdf`, retrieved 25 March 2011.

Goodman, L.A.; and Kruskal, W.H. (1979) *Measures of Association for Cross Classification*.

Grootaert, Christiaan; and Jeanine Braithwaite. (1998) "Poverty Correlates and Indicator-Based Targeting in Eastern Europe and the Former Soviet Union", World Bank Policy Research Working Paper No. 1942, `go.worldbank.org/VPMWVLU8E0`, retrieved 25 March 2011.

Grosh, Margaret; and Judy L. Baker. (1995) "Proxy Means Tests for Targeting Social Programs: Simulations and Speculation", World Bank LSMS Working Paper No. 118, `go.worldbank.org/W9OWN57PD0`, retrieved 25 March 2011.

Gwatkin, Davidson R.; Rutstein, Shea; Johnson, Kiersten; Suliman, Eldaw; Wagstaff, Adam; and Agbessi Amouzou. (2007) "Socio-Economic Differences in Health, Nutrition, and Population: Burkina Faso", World Bank Country Reports on HNP and Poverty, `go.worldbank.org/T6LCN5A340`, retrieved 25 March 2011.

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.

Howe, Laura D.; Hargreaves, James R.; Gabrysch, Sabine; and Sharon R.A. Huttly. (2009) "Is the Wealth Index a Proxy for Consumption Expenditure? A Systematic Review", *Journal of Epidemiology and Community Health*, Vol. 63, pp. 871–880.

Institut National de la Statistique et la Démographie. (2003) "Burkina Faso: La Pauvreté en 2003", `www.insd.bf/fr/IMG/pdf/Profil_Pauvrete_2003.pdf`, retrieved 18 March 2011.

IRIS Center. (2007a) "Manual for the Implementation of USAID Poverty Assessment Tools", `povertytools.org/training_documents/Manuals/USAID_PAT_Manual_Eng.pdf`, retrieved 25 March 2011.

_____. (2007b) "Introduction to Sampling for the Implementation of PATs", `povertytools.org/training_documents/Sampling/Introduction_Sampling.ppt`, retrieved 25 March 2011.

_____. (2005) "Notes on Assessment and Improvement of Tool Accuracy", `povertytools.org/other_documents/AssessingImproving_Accuracy.pdf`, retrieved 25 March 2011.

Johnson, Glenn. (2007) "Lesson 3: Two-Way Tables—Dependent Samples", `www.stat.psu.edu/online/courses/stat504/03_2way/53_2way_compare.htm`, retrieved 25 March 2011.

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, ciep.itam.mx/~martinel/lies4.pdf, retrieved 25 March 2011.

Matul, Michal; and Sean Kline. (2003) "Scoring Change: Prizma's Approach to Assessing Poverty", Microfinance Centre for Central and Eastern Europe and the New Independent States Spotlight Note No. 4, www.mfc.org.pl/doc/Research/ImpAct/SN/MFC_SN04_eng.pdf, retrieved 25 March 2011.

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

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", World Bank Report No. SASPR–7, http://siteresources.worldbank.org/EXTSAREGTOPPOVRED/Resources/493440-1102216396155/572861-1102221461685/Proxy+Means+Test+for+Targeting+Welfare+Benefits.pdf, retrieved 25 March 2011.

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, measuredhs.com/pubs/pdf/CR6/CR6.pdf, retrieved 25 March 2011.

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*, support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_logistic_sect035.htm, retrieved 25 March 2011.

Schreiner, Mark. (2013) "Simple Poverty Scorecard®: Bangladesh", simplepovertyscorecard.com/BGD_2010_ENG.pdf, retrieved 30 June 2016.

_____. (2010) "Simple Poverty Scorecard®: Honduras", simplepovertyscorecard.com/HND_2007_ENG.pdf, retrieved 30 June 2016.

_____. (2009a) "Simple Poverty Scorecard®: Peru", simplepovertyscorecard.com/PER_2007_ENG.pdf, retrieved 30 June 2016.

_____. (2009b) "Simple Poverty Scorecard®: Philippines", simplepovertyscorecard.com/PHL_2002_ENG.pdf, retrieved 30 June 2016.

_____. (2009c) "Simple Poverty Scorecard®: Pakistan", simplepovertyscorecard.com/PAK_2005_ENG.pdf, retrieved 30 June 2016.

_____. (2009d) "Simple Poverty Scorecard®: Bolivia", simplepovertyscorecard.com/BOL_2007_ENG.pdf, retrieved 30 June 2016.

_____. (2009e) "Simple Poverty Scorecard®: Mexico", simplepovertyscorecard.com/MEX_2008_ENG.pdf, retrieved 30 June 2016.

_____. (2008a) "Simple Poverty Scorecard®: Peru", simplepovertyscorecard.com/PER_2003_ENG.pdf, retrieved 30 June 2016.

_____. (2008b) "Simple Poverty Scorecard®: India", simplepovertyscorecard.com/IND_2005_ENG.pdf, retrieved 30 June 2016.

_____. (2006) "Is One Simple Poverty Scorecard® Enough for India?", microfinance.com/English/Papers/Scoring_Poverty_India_Segments.pdf, retrieved 25 March 2011.

_____. (2005a) "Índice de Calificación de Pobreza™: México", simplepovertyscorecard.com/MEX_2002_SPA.pdf, retrieved 30 June 2016.

_____. (2005b) "IRIS Questions on the Simple Poverty Scorecard®", microfinance.com/English/Papers/Scoring_Poverty_Response_to_IRIS.pdf, retrieved 30 June 2016.

_____. (2002) *Scoring: The Next Breakthrough in Microfinance?* Consultative Group to Assist the Poor Occasional Paper No. 7, collab2.cgap.org//gm/document-1.9.29797/3276_076.pdf, retrieved 25 March 2011.

_____; Matul, Michal; Pawlak, Ewa; and Sean Kline. (2004) "Simple Poverty Scorecard®: Lessons from a Microlender in Bosnia-Herzegovina", microfinance.com/English/Papers/Scoring_Poverty_in_BiH_Short.pdf, retrieved 25 March 2011.

_____; and Gary Woller. (2010a) "Simple Poverty Scorecard®: Ghana", simplepovertyscorecard.com/GHA_2005_ENG.pdf, retrieved 30 June 2016.

_____; and Gary Woller. (2010b) "Simple Poverty Scorecard®: Guatemala", simplepovertyscorecard.com/GTM_2006_ENG.pdf, retrieved 30 June 2016.

Sillers, Don. (2006) "National and International Poverty Lines: An Overview", povertytools.org/other_documents/siller-povertylines.doc, retrieved 25 March 2011.

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, Alessandro; and Angus Deaton. (2007) "Using Census and Survey Data to Estimate Poverty and Inequality for Small Areas", *Review of Economics and Statistics*, Vol. 91, No. 4, Pages 773–792.

Toohig, Jeff. (2007) "Progress out of Poverty Index: Training Guide", Grameen Foundation, progressoutofpoverty.org/toolkit, retrieved 25 March 2011.

United States Congress. (2004) "Microenterprise Results and Accountability Act of 2004 (HR 3818 RDS)", November 20, povertytools.org/other_documents/HR%203818%20RDS.pdf, retrieved 25 March 2011.

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. (2008) "International Comparison Project: Tables of Results", siteresources.worldbank.org/ICPINT/Resources/icp-final-tables.pdf, retrieved 25 March 2011.

Zeller, Manfred. (2004) "Review of Poverty Assessment Tools", pdf.usaid.gov/pdf_docs/PNADH120.pdf, retrieved 25 March 2011.

Guidelines for the Interpretation of Scorecard Indicators

The following is taken from:

Institut National de la Statistique et de la Démographie. (2003) *Manuel de l'Enquêteur, Enquête Burkinabé sur les Conditions de Vie des Ménages, Première Phase*, Ministère de l'Economie et du Développement.

1. How many household members are 14-years-old or younger?

According to pp. 11–13 of the *Manuel* : "The *household* is a socio-economic unit within which different members, be they blood relatives or not, live in the same house or residence, share their resources in common, and together satisfy their nutritional and other basic life-needs, under the authority of one person called the household head. . . . There are many special cases of households, listed below:

"1. Any person who lives alone in a residence and provides for his/her own basic life needs (nutrition, shelter, clothing, etc.) should be considered to be a one-person household. If not, then the person should be associated with a household (in the local area) that provides for his/her basic needs. To determine which is the case, the interviewer can ask the person questions of this type: "Where do you normally eat your meals?" "Do you yourself pay for your own shelter?"

"2. Several wives with a single husband who live together and share meals together constitute a single household. If the wives live together but do not share meals, then they constitute distinct households, just as if they did not live together. Men with multiple wives should be counted in the household where they spent the previous night.

"3. Domestic servants (maids, servant boys, etc.) are not to be counted as part of the household where they work, even if they eat there and spend the night there. . . .

"Three criteria are used to determine whether a person should be considered to be a household member:

1. He/she must usually live with the household and eat meals there
2. He/she should recognize the authority of the head of the household
3. He/she must have been present in the household during at least six of the past twelve months (the household head is an exception to this rule). If someone is absent from the household for more than six of the past twelve months, then he/she should not be considered to be a member of the household (except the household head, who is still counted as a household member even if he/she has been absent for more than six of the past twelve months). If someone is present in the household for less than six of the past twelve months but nevertheless intends to continue to reside with the household from now on, then that person should be considered to be a household member. Examples of this type include:

   a. Newborns are considered to be household members even if they are less than six-months-old
   b. Newly married wives are considered to be household members even if they have not yet lived for at least six months in their new household

"*Present residents* are defined as anyone who normally lives with the household for more than six months and who spent the night before the interview with the household. . . . Furthermore, all of the following cases should be counted as 'present residents' even though they did not spend the night before the interview with the household:

- Doctors and other health-care workers
- Watchmen
- Night-shift factory workers
- Workers who are working at night—whether permanently or temporarily—such as truck drivers, travelers who happened to spend the previous night away from the household, etc.

"*Absent residents* are defined as anyone who did not spend the night before the interview with the household but whose absence will not last six months. Students who are frequently away from the household during the school year are to be considered as members of the household where they live during the school year. Students who are not present because they have left on vacation but who live with the household during the school year are to be considerd as absent residents.

"Take note: Wives who are members of the household who go home to stay with their families of origin are still considered to be household members, even if they stay away for more than six months.

"Visitors who are staying with the household at the time of the interview are not to be counted as members of the household. A *visitor* is defined as anyone who does not normally live with the household, that is, who is not there for at least six months out of twelve and who has no intention of staying for at least six months. . . .

"The household roster should be compiled carefully to be sure that all members—and only members—are included. In the case of polygamous households (or households with multiple mothers), the interviewer should compile a list of all the children and their association with a specific mother. To make sure than no one is left out, the interviewer should pay special attention to three classes of people who are at high risk of being mistakenly omitted. The first of these are people who are temporarily absent but who should be included as household members. The second is domestic servants and lodgers; they are often members of another household and therefore should not be counted. If, however, they do not qualify as members of some other household, then they should be counted as members of the household in which they work. The third group are infants, as they are easy to overlook.

2. <u>In what languages can the male head/spouse read and write?</u>

According to p. 13 of the *Manuel*, "the household head is the decision-maker for the household, and his/her authority is recognized by the other members of the household. In some cases, the household head may not be the main breadwinner."

For purposes of the poverty scorecard, the *male head/spouse* is defined as:

- The household head, if the household head is male
- The male partner of the household head, if the household head is female
- Non-existent, if no one fits the two criteria listed above

According to p. 16 of the *Manuel*, "The objective is to record functional literacy (the ability to read and write simple phrases used in daily life, as well as ordinary writing such as letters and newspapers). A person qualifies as *literate* regardless of the language in which he/she can read and write, whether or not the language is a common one or an official one, as long as the language has a written form. If a person can read but cannot write, then he/she is to be considered as illiterate and to correspond to the case of "None". To establish literacy, a person does not have to pass an actual test of reading and writing. Rather, the interviewer should use his/her judgment and the information provided by the respondent to make a decision about each particular person."

3. Has the female head/spouse completed first grade?

According to p. 13 of the *Manuel*, "the household head is the decision-maker for the household, and his/her authority is recognized by the other members of the household. In some cases, the household head may not be the main breadwinner."

For purposes of the poverty scorecard, the *female head/spouse* is defined as:

- The household head, if the household head is female
- The female partner of the household head, if the household head is male
- Non-existent, if no one fits the two criteria listed above

According to p. 13 of the *Manuel*, "To record a grade as having been completed, the person in question must have actually completed the grade. For example, someone who started fourth grade but never finished should be recorded as having completed third grade. In the same way, someone who is currently in second grade should be recorded as having completed first grade."

4. What is the main source of energy for lighting?

According to p. 38 of the *Manuel*, "If more than one source of energy is used for lighting, then the source that is used the most frequently is the one that should be recorded."

5. What toilet arrangement does the household have?

According to p. 38 of the *Manuel*, "The type of toilet arrangement is an important indicator of the hygienic conditions faced by the household. *Flush to a septic tank* is a modern arrangement with a specific piece of furniture, generally installed inside the residence in a room designed for that purpose, connected with a tank and a drain that leads to a septic tank for the dispersal of waste. A *ventilated pit latrine* is a covered pit with a pipe to vent foul smells. There may be one or two pits. "Improved" ordinary latrines are counted under the rubric of "ventilated pit latrines". A *non-ventilated pit latrine* is a pit covered with a slab without a vent. *No toilet arrangement* means excreting directly on the ground, in the bushes, or behind the residence."

For this indicator, the specific words used in the questionnaire for the 2003 EBCVM are not exactly the same as the words used in the enumerator's manual.

6. Does the household own a television?

According to p. 35 of the *Manuel*: "It does not matter which member of the household owns the asset, and the response should not be marked as 'Yes' if the asset is not in working order."

7. Does the household own a bed or a mattress?

According to p. 35 of the *Manuel*: "It does not matter which member of the household owns the asset, and the response should not be marked as 'Yes' if the asset is not in working order."

8. Does the household own a scooter or a motorcycle?

According to p. 35 of the *Manuel*: "It does not matter which member of the household owns the asset, and the response should not be marked as 'Yes' if the asset is not in working order."

9. Have any household members, in their main occupation in the last seven days, worked in agriculture, animal husbandry, fishing, or forestry?

According to p. 18 of the *Manuel*, "This questions records the high-level type of profession of the person. The types of profession are:

1. Agriculture, animal husbandry, fishing, or forestry
2. Management and administration: executives, middle managers, employees, clerks, laborers, junior managers, and professionals
3. Trade or commerce
4. Skilled crafts and manufacturing
5. Domestic and other services
6. Armed forces and security
7. Other careers and professions
8. No profession or non-specified occupations"

According to p. 38 of the *Manuel*, "the sub-types under *Agriculture, animal husbandry, fishing, or forestry*" are:

- Agriculture—farmer
- Vegetables and fruit to be sold at a profit as a business
- Subsistence farmer
- Orchard—nurseryman
- Lumberjack
- Cattle rancher
- Poultry breeder
- Shepherd
- Fisher
- Hunter
- Other professions in this group not otherwise classified"

10. How many head of cattle or other large animals does the household now own?

According to p. 18 of the *Manuel*, "This question deals with the number of head of cattle or other herds of large animals owned by the household at the time of the interview, whether the livestock is currently in the care of the household or whether they are being cared for by someone else. It does not matter which household member owns the livestock. Large animals such as camels, cows, donkeys, horses, etc. are to be counted, along with cattle, as 'large animals'. . . . Sheep, swine, and poultry are not counted as 'large animals'."

Figure 2: Sample sizes and poverty rates by sub-sample and by poverty line at both the household level and the person level

			Poverty rates (% with expenditure below a poverty line) and poverty lines (XOF/person/day)						
			National lines				USAID	Intl. 2005 PPP	
Sub-sample	**Level**	**Sample size**	**50%**	**75%**	**100%**	**150%**	**'Extreme'**	**$1.25/day**	**$2.50/day**
Poverty lines:	N/A	N/A	113	170	226	340	153	288	577
Poverty Rates:									
All Burkina Faso	Households	8,500	7.5	22.4	37.5	59.0	17.2	50.6	79.8
	People	N/A	10.9	29.3	46.4	68.8	23.2	60.3	87.0
Construction and calibration									
Selecting indicators and points, and associating scores with likelihoods	Households	4,193	7.5	22.5	37.6	59.0	17.2	50.6	79.9
	People	N/A	10.7	29.0	46.3	68.7	22.7	60.3	87.2
Validation									
Measuring accuracy	Households	4,307	7.5	22.3	37.4	59.0	17.3	50.5	79.6
	People	N/A	11.0	29.6	46.5	68.9	23.7	60.4	86.7

Source: 2003 EBCVM.

Figure 3: Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly linked with higher poverty likelihoods)
1,139	How many household members worked or been away from their usual work in the last seven days? (None; One; Two; Three; Four; Five; Six; Seven)
976	How many household members are 15-years-old or younger? (Six or more; Five; Four; Three; Two; One; None)
969	How many household members are 14-years-old or younger? (Six or more; Five; Four; Three; Two; One; None)
963	How many household members are 16-years-old or younger? (Six or more; Five; Four; Three; Two; One; None)
959	How many household members are 13-years-old or younger? (Five or more; Four; Three; Two; One; None)
940	How many household members are 17-years-old or younger? (Six or more; Five; Four; Three; Two; One; None)
939	How many members does the household have? (Ten or more; Nine; Eight; Seven; Six; Five; Four; Three; Two; One; None)
933	How many household members are 12-years-old or younger? (Five or more; Four; Three; Two; One; None)
921	How many household members are 11-years-old or younger? (Four or more; Three; Two; One; None)
913	How many household members are 18-years-old or younger? (Seven or more; Six; Five; Four; Three; Two; One; None)
825	If any household members worked in agriculture, animal husbandry, fishing, or forestry in their main occupation for the last seven days, how many head of cattle or other large animals does the household now own? (One; None; Two; Three to five; Six or more; No one in agriculture etc.)
743	Do all household members ages 7 to 12 currently go to school? (No; Yes; No members in this age range)
740	Do all household members ages 7 to 13 currently go to school? (No; Yes; No members in this age range)
735	Do all household members ages 7 to 14 currently go to school? (No; Yes; No members in this age range)

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

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly linked with higher poverty likelihoods)
724	Do all household members ages 7 to 15 currently go to school? (No; Yes; No members in this age range)
724	Do all household members ages 7 to 11 currently go to school? (No; Yes; No members in this age range)
687	Do all household members ages 7 to 16 currently go to school? (No; Yes; No members in this age range)
640	What has been the main occupation of the male head/spouse for the last seven days? (Agriculture, animal husbandry, fishing, or forestry; No male head/spouse; None; Trade, skilled crafts and manufacturing, domestic and other services, armed forces and security, other careers and professions, or no profession or non-specified occupations; Management and administration: executives, middle managers, employees, clerks, laborers, junior managers, and professionals)
631	What is the sector in which the male head/spouse works? (Agriculture, hunting, or forestry; None; Other; No male head/spouse; Public administration)
600	How many household members are 6-years-old or younger? (Three or more; Two; One; None)
591	Do all household members ages 7 to 17 currently go to school? (No; Yes; No members in this age range)
578	What is the main source of drinking water? (River, lake, stream, or other; Unimproved wells; Wells with a handpump; Tube wells; Public standpipe; Private indoor faucet, or shared indoor faucet)
573	What is the sector in which the female head/spouse works? (Agriculture, hunting, or forestry; None; No female head/spouse; Other)
553	What has been the main occupation of the female head/spouse for the last seven days? (Agriculture, animal husbandry, fishing, or forestry; None; No female head/spouse; Other)
552	What is the main source of energy for lighting? (Firewood, or other; Candles, kerosene, or LPG; Flashlight, or batteries; Electricity, or solar energy)
547	How many households work in a sector other than agriculture, hunting, or forestry? (None; One or more)
541	Do all household members ages 7 to 18 currently go to school? (No; Yes; No members in this age range)

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

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly linked with higher poverty likelihoods)
539	What toilet arrangement does the household have? (No toilet arrangement, or other; Non-ventilated pit latrine; Ventilated pit latrine, or flush to a septic tank)
529	How does the household dispose of its garbage? (Private refuse heap; Street; Pit; Public refuse heap; Bin; Curbside collection)
527	How many hectares of land does the household use, whether owned by the household or not? (None; 0 to <0.5; 0.5 to <1; 1 to <5; 5 to <10; 10 to <15; 15 to <20; 20 to <30; 30 to <45; 45 to <60; 60 or more)
526	Does the household own a television? (No; Yes)
498	Have any household members, in their main occupation for the last seven days, worked in agriculture, animal husbandry, fishing, or forestry? (Yes; No)
492	Does the household have an electrical connection? (No; Yes)
480	What is the highest grade that the male head/spouse has completed? (None, or pre-school; No male head/spouse; First to sixth grade; Seventh grade or higher)
476	What is the flooring material in the main residence? (Packed earth/sand; Cement, or tile)
472	Does the household own an iron, refrigerator/freezer, sewing machine, or stove? (No; Yes)
459	Can the female head/spouse read and write in any language? (No; No female head/spouse; Yes)
445	How many hectares of land does the household own? (None; 0 to <0.5; 0.5 to <1; 1 to <5; 5 to <10; 10 to <15; 15 to <20; 20 to <30; 30 to <45; 45 to <60; 60 or more)
420	How was the female head/spouse paid in her main line of work? (Apprectice or non-remunerated family worker, or self-employed; She did not work, or task-based (by the hour or day); No female head/spouse; Salary or wage in cash or kind)
387	What is the material of the walls of the residence? (Earth or mud bricks; Other)

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

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly linked with higher poverty likelihoods)
387	How was the male head/spouse paid in his main line of work? (Self-employed; No male head/spouse; Apprectice or non-remunerated family worker, he did not work, or task-based (by the hour or day); Salary or wage in cash or kind)
384	What is occupational position of the female head/spouse? (Unremunerated household worker; Self-employed; None; Other; No female head/spouse)
383	What is occupational position of the male head/spouse? (Self-employed; No male head/spouse; None; Other)
382	Has the female head/spouse completed first grade? (No; No female head/spouse; Yes)
378	Does the household or a household member own the residence? (Owner, nomad, or temporary residence; Occupied rent-free; Renter)
362	What is the marital status of the female head/spouse? (Polygamously married; Widow; Monogamously married; No female head/spouse; Single, never-married, divorced/separated, or cohabiting)
357	In what languages can the male head/spouse read and write? (None, or no male head/spouse; French only; A non-French language (regardless of French literacy))
352	Does the household own a scooter, motorcycle, car, or truck? (No; Yes)
351	What is the marital status of the male head/spouse? (Polygamously married; Monogamously married; Widower; No male head/spouse; Single, never-married, divorced or separated, or cohabiting)
346	Does the household own a motorcycle or scooter? (No; Yes)
334	What is the nature of the work in the main occupation of the female head/spouse for the last seven days? (Seasonal; None; Permanent; Day labor, other temporary, or no female head/spouse)
308	How old was the male head/spouse on his last birthday? (50 or older; 45 to 49; 40 to 44; 35 to 39; No male head/spouse; 30 to 34; 29 or younger)

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

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly linked with higher poverty likelihoods)
308	Has the female head/spouse worked or been away from her usual work in the last seven days? (Yes; No; No female head/spouse)
280	What is the main fuel used for cooking? (Firewood, crop residue/sawdust, animal waste, or other; Charcoal, paraffin or kerosene, LPG, or electricity)
280	What type of building is the residence? (Traditional house; Simple detached house; Condominium; Villa, or apartment building)
252	Does the household own a bed or a mattress? (No; Yes)
250	How old was the female head/spouse on her last birthday? (35 to 39; 45 to 49; 30 to 34; 50 or older; 40 to 44; 25 to 29; 20 to 24; 29 or younger; No female head/spouse)
248	Does the household have a telephone? (No; Yes)
223	What is the roofing material of the residence? (Earth; Straw; Metal sheets, cement/concrete, tiles, or other)
220	What is the nature of the work in the main occupation of the male head/spouse for the last seven days? (Seasonal; None; No male head/spouse; Permanent, day labor, or other temporary)
198	Does the household now own any cattle, other large animals, sheep, goats, or other mid-sized animals? (Yes; No)
173	How many sheep, goats, and other mid-sized animals does the household own now? (One or two; Three or four; Five to Seven; Eight to fifteen; None)
146	How many rooms does the residence have? (Six or more; Five; Four; Three; Two; One)
143	How many household members can read and write in any language? (None; One; Two; Three; Four or more)
142	What is the structure of household headship? (Both male and female heads/spouses; Only a female head/spouse; Only a male head/spouse)
127	How many head of cattle and other large animals does the household own now? (None, or one; Two; Three to five; Six or more)

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

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly linked with higher poverty likelihoods)
103	Does the household own a car or truck? (No; Yes)
89	Does any household member go to a private school (non-religious, Catholic, or Protestant)? (No; Yes)
83	Does the household own a radio? (No; Yes)
62	Does the household own a plow o cart? (Yes; No)
54	Does the household own a bicycle? (Yes; No)
54	Does the household own any draft animals? (Yes; No)
36	Has the male head/spouse worked or been away from his usual work in the last seven days? (No; No male head/spouse; Yes)
22	Does the residence have a kitchen? (No; Yes)
1	Does the household own a bicycle, scooter, motorcycle, car, or truck? (No; Yes)
0	Does the household use any land that it does not own? (No; Yes)

Source: 2003 EBCVM and the national poverty line

Tables for
the National Poverty Line

(and tables pertaining to all seven poverty lines)

Figure 4 (National 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	84.2
5–9	75.5
10–14	66.4
15–19	57.2
20–24	43.7
25–29	34.6
30–34	20.7
35–39	17.0
40–44	14.4
45–49	4.2
50–54	2.6
55–59	2.0
60–64	0.7
65–69	0.9
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Figure 5 (National line): Derivation of estimated poverty likelihoods associated with scores

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	1,064	÷	1,263	=	84.2
5–9	3,828	÷	5,071	=	75.5
10–14	8,101	÷	12,194	=	66.4
15–19	9,004	÷	15,736	=	57.2
20–24	7,288	÷	16,680	=	43.7
25–29	3,846	÷	11,122	=	34.6
30–34	2,104	÷	10,148	=	20.7
35–39	1,179	÷	6,922	=	17.0
40–44	749	÷	5,185	=	14.4
45–49	164	÷	3,960	=	4.2
50–54	75	÷	2,859	=	2.6
55–59	49	÷	2,421	=	2.0
60–64	12	÷	1,590	=	0.7
65–69	15	÷	1,686	=	0.9
70–74	0	÷	1,613	=	0.0
75–79	0	÷	703	=	0.0
80–84	0	÷	638	=	0.0
85–89	0	÷	87	=	0.0
90–94	0	÷	122	=	0.0
95–100	0	÷	0	=	0.0

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

Figure 6: Distribution of household poverty likelihoods across consumption ranges demarcated by per-adult-equivalent poverty lines

	Likelihood of having expenditure in range demarcated by poverty lines per day per person							
Score	<50% Natl. <XOF113	=>50% Natl. and <USAID =>XOF113 and <XOF153	=>USAID and <75% Natl. =>XOF153 and <XOF170	=>75% Natl. and <100% Natl. =>XOF170 and <XOF226	=>100% Natl. and <$1.25/day =>XOF226 and <XOF288	=>$1.25/day and <150% Natl. =>XOF288 and <XOF340	=>150% Natl. and <$2.50/day =>XOF340 and <XOF577	=>$2.50/day =>XOF577
0–4	29.7	26.4	7.1	21.1	10.1	0.0	5.7	0.0
5–9	28.2	21.6	8.2	17.5	13.3	5.0	5.5	0.8
10–14	14.4	19.1	10.4	22.5	13.4	7.7	10.2	2.3
15–19	11.3	15.0	9.4	21.6	15.2	9.7	13.7	4.3
20–24	6.9	11.8	6.2	18.8	18.0	9.5	20.8	8.0
25–29	4.6	7.1	3.9	19.0	14.2	10.2	28.3	12.8
30–34	3.0	4.3	2.3	11.2	16.2	11.1	32.4	19.6
35–39	1.2	2.6	1.8	11.5	13.1	10.5	30.2	29.2
40–44	1.2	2.1	3.1	8.1	11.1	9.1	35.7	29.6
45–49	0.2	0.0	1.0	3.0	7.0	9.6	30.9	48.4
50–54	0.0	1.0	0.5	1.2	4.9	3.1	28.6	60.7
55–59	0.0	0.0	0.0	2.0	1.4	3.8	26.6	66.2
60–64	0.0	0.0	0.0	0.7	1.7	0.8	17.4	79.4
65–69	0.0	0.0	0.0	0.9	0.0	3.3	13.3	82.6
70–74	0.0	0.0	0.0	0.0	0.0	1.0	3.7	95.3
75–79	0.0	0.0	0.0	0.0	0.0	0.0	4.4	95.6
80–84	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
85–89	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
90–94	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
95–100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0

All poverty likelihoods in percentage units.

Figure 7 (National line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) 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	–4.7	4.1	4.5	5.4
5–9	+2.2	2.8	3.2	4.1
10–14	+2.3	1.9	2.3	2.8
15–19	–0.6	1.6	1.9	2.6
20–24	+3.2	1.5	1.9	2.6
25–29	+0.0	1.9	2.2	2.8
30–34	–9.3	5.7	5.8	6.1
35–39	–2.3	2.3	2.7	3.6
40–44	+1.5	2.2	2.6	3.4
45–49	–5.0	3.6	3.8	4.2
50–54	+1.2	0.8	0.9	1.2
55–59	+1.5	0.6	0.6	0.8
60–64	+0.7	0.0	0.0	0.0
65–69	–0.6	1.2	1.4	2.0
70–74	+0.0	0.0	0.0	0.0
75–79	+0.0	0.0	0.0	0.0
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 8 (National 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	+2.2	65.9	72.9	80.5
4	+0.3	37.4	45.1	54.8
8	–0.8	28.2	33.6	40.8
16	–1.1	19.8	23.1	30.3
32	–0.7	13.5	16.3	21.5
64	–0.5	9.7	11.7	15.0
128	–0.4	7.2	8.4	10.9
256	–0.4	5.3	6.3	8.0
512	–0.4	3.8	4.5	5.7
1,024	–0.4	2.6	3.0	3.9
2,048	–0.4	1.8	2.1	2.6
4,096	–0.3	1.2	1.5	2.0
8,192	–0.3	0.9	1.1	1.4
16,384	–0.3	0.6	0.7	0.9

Figure 9 (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, scorecard applied to the validation sample

	Poverty line											
	National lines							USAID		Intl. 2005 PPP		
	50%		75%		100%		150%	'Extreme'		$1.25/day		$2.50/day
Estimate minus true value	–0.1		+0.1		–0.3		+0.2	–0.3		–0.1		+0.2
Precision of difference	0.4		0.5		0.6		0.6	0.5		0.6		0.4
α factor	1.07		1.04		1.03		0.92	1.04		0.97		0.84
Precision is measured as 90-percent confidence intervals in units of +/– percentage points.												
Differences and precision estimated from 500 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.												
The USAID "extreme" line is in per-person units.												

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

True poverty status	Targeting segment: Targeted	Targeting segment: Non-targeted
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 (National 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	1.1	36.3	0.1	62.4	63.5	–93.6
5–9	4.9	32.5	1.4	61.1	66.0	–70.0
10–14	12.6	24.9	5.9	56.6	69.2	–16.9
15–19	21.7	15.8	12.6	50.0	71.7	+49.4
20–24	28.4	9.0	22.5	40.1	68.5	+39.9
25–29	32.4	5.1	29.7	32.9	65.2	+20.7
30–34	35.2	2.3	37.1	25.5	60.6	+1.0
35–39	36.4	1.0	42.7	19.8	56.2	–14.1
40–44	37.0	0.4	47.3	15.3	52.3	–26.3
45–49	37.3	0.1	50.9	11.6	49.0	–36.0
50–54	37.4	0.0	53.7	8.8	46.2	–43.5
55–59	37.4	0.0	56.1	6.4	43.8	–49.9
60–64	37.4	0.0	57.7	4.8	42.2	–54.2
65–69	37.4	0.0	59.4	3.2	40.6	–58.6
70–74	37.4	0.0	61.0	1.6	39.0	–62.9
75–79	37.4	0.0	61.7	0.8	38.3	–64.8
80–84	37.4	0.0	62.3	0.2	37.7	–66.5
85–89	37.4	0.0	62.4	0.1	37.6	–66.7
90–94	37.4	0.0	62.6	0.0	37.4	–67.1
95–100	37.4	0.0	62.6	0.0	37.4	–67.1

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

Figure 12 (National line): For a given score cut-off, the percentage of all households who are targeted (that is, have a score equal to or less than the cut-off), the percentage of targeted households who are poor (that is, below the poverty line), the percentage of poor households who are targeted, and the number of poor households who are successful targeted (coverage) 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	1.3	88.7	3.0	7.9:1
5–9	6.3	77.4	13.1	3.4:1
10–14	18.5	67.9	33.6	2.1:1
15–19	34.3	63.3	57.9	1.7:1
20–24	50.9	55.8	76.0	1.3:1
25–29	62.1	52.1	86.4	1.1:1
30–34	72.2	48.7	93.9	0.9:1
35–39	79.1	46.0	97.2	0.9:1
40–44	84.3	43.9	98.9	0.8:1
45–49	88.3	42.3	99.7	0.7:1
50–54	91.1	41.0	99.9	0.7:1
55–59	93.6	40.0	99.9	0.7:1
60–64	95.2	39.3	99.9	0.6:1
65–69	96.8	38.7	100.0	0.6:1
70–74	98.4	38.0	100.0	0.6:1
75–79	99.2	37.8	100.0	0.6:1
80–84	99.8	37.5	100.0	0.6:1
85–89	99.9	37.5	100.0	0.6:1
90–94	100.0	37.4	100.0	0.6:1
95–100	100.0	37.4	100.0	0.6:1

Tables for
50% of the National Poverty Line

Figure 4 (50% of the national 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	29.7
5–9	28.2
10–14	14.4
15–19	11.3
20–24	6.9
25–29	4.6
30–34	3.0
35–39	1.2
40–44	1.2
45–49	0.2
50–54	0.0
55–59	0.0
60–64	0.0
65–69	0.0
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Figure 7 (50% of the national line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) 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	–3.2	5.7	7.0	9.2
5–9	+7.7	2.5	2.9	3.7
10–14	–3.9	2.7	2.8	3.2
15–19	+0.3	1.0	1.2	1.5
20–24	+0.2	0.8	0.9	1.3
25–29	+2.3	0.6	0.7	1.0
30–34	–1.7	1.3	1.4	1.6
35–39	–2.4	1.8	1.9	2.2
40–44	+1.1	0.2	0.2	0.2
45–49	+0.2	0.0	0.0	0.0
50–54	+0.0	0.0	0.0	0.0
55–59	+0.0	0.0	0.0	0.0
60–64	+0.0	0.0	0.0	0.0
65–69	+0.0	0.0	0.0	0.0
70–74	+0.0	0.0	0.0	0.0
75–79	+0.0	0.0	0.0	0.0
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 8 (50% of the national 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	–0.2	50.0	58.4	63.3
4	+0.0	21.3	26.2	36.2
8	–0.5	16.3	19.0	23.7
16	–0.1	11.2	13.2	17.6
32	+0.0	8.3	9.9	13.7
64	+0.1	6.1	7.2	8.8
128	–0.0	4.1	4.9	6.4
256	–0.0	3.0	3.5	4.4
512	–0.1	2.1	2.5	3.1
1,024	–0.1	1.4	1.6	2.2
2,048	–0.1	1.0	1.2	1.5
4,096	–0.1	0.7	0.8	1.1
8,192	–0.1	0.5	0.6	0.8
16,384	–0.1	0.4	0.4	0.5

Figure 11 (50% of the national 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	0.4	7.1	0.9	91.6	92.0	–78.1
5–9	1.4	6.1	4.9	87.6	89.0	+3.7
10–14	3.6	3.9	14.9	77.6	81.2	–99.3
15–19	5.4	2.1	28.9	63.6	69.0	–285.5
20–24	6.6	0.9	44.4	48.1	54.7	–492.5
25–29	6.8	0.7	55.2	37.3	44.1	–637.5
30–34	7.3	0.2	64.9	27.6	34.8	–767.0
35–39	7.5	0.0	71.7	20.8	28.3	–856.8
40–44	7.5	0.0	76.8	15.7	23.2	–925.8
45–49	7.5	0.0	80.8	11.7	19.2	–978.7
50–54	7.5	0.0	83.7	8.9	16.3	–1,016.9
55–59	7.5	0.0	86.1	6.4	13.9	–1,049.2
60–64	7.5	0.0	87.7	4.8	12.3	–1,070.4
65–69	7.5	0.0	89.3	3.2	10.7	–1,093.0
70–74	7.5	0.0	91.0	1.6	9.0	–1,114.5
75–79	7.5	0.0	91.7	0.8	8.3	–1,123.9
80–84	7.5	0.0	92.3	0.2	7.7	–1,132.4
85–89	7.5	0.0	92.4	0.1	7.6	–1,133.6
90–94	7.5	0.0	92.5	0.0	7.5	–1,135.2
95–100	7.5	0.0	92.5	0.0	7.5	–1,135.2

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

Figure 12 (50% of the national line): For a given score cut-off, the percentage of all households who are targeted (that is, have a score equal to or less than the cut-off), the percentage of targeted households who are poor (that is, below the poverty line), the percentage of poor households who are targeted, and the number of poor households who are successful targeted (coverage) 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	1.3	29.9	5.0	0.4:1
5–9	6.3	22.6	19.1	0.3:1
10–14	18.5	19.4	48.1	0.2:1
15–19	34.3	15.7	72.0	0.2:1
20–24	50.9	12.9	87.7	0.1:1
25–29	62.1	11.0	91.2	0.1:1
30–34	72.2	10.1	97.2	0.1:1
35–39	79.1	9.4	99.8	0.1:1
40–44	84.3	8.9	100.0	0.1:1
45–49	88.3	8.5	100.0	0.1:1
50–54	91.1	8.2	100.0	0.1:1
55–59	93.6	8.0	100.0	0.1:1
60–64	95.2	7.9	100.0	0.1:1
65–69	96.8	7.7	100.0	0.1:1
70–74	98.4	7.6	100.0	0.1:1
75–79	99.2	7.6	100.0	0.1:1
80–84	99.8	7.5	100.0	0.1:1
85–89	99.9	7.5	100.0	0.1:1
90–94	100.0	7.5	100.0	0.1:1
95–100	100.0	7.5	100.0	0.1:1

Tables for
75% of the National Poverty Line

Figure 4 (75% of the national 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	63.1
5–9	58.0
10–14	43.9
15–19	35.6
20–24	24.9
25–29	15.6
30–34	9.5
35–39	5.6
40–44	6.4
45–49	1.1
50–54	1.5
55–59	0.0
60–64	0.0
65–69	0.0
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Figure 7 (75% of the national line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) 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	–5.8	5.6	6.4	8.6
5–9	+7.7	3.0	3.6	4.7
10–14	–0.7	1.9	2.1	2.9
15–19	+0.8	1.6	2.0	2.6
20–24	+3.8	1.3	1.6	2.1
25–29	+0.3	1.5	1.8	2.3
30–34	–7.5	4.6	4.8	5.2
35–39	–4.4	3.1	3.3	3.7
40–44	+1.4	1.3	1.6	2.1
45–49	–2.5	2.0	2.1	2.4
50–54	+1.3	0.3	0.3	0.4
55–59	+0.0	0.0	0.0	0.0
60–64	+0.0	0.0	0.0	0.0
65–69	+0.0	0.0	0.0	0.0
70–74	+0.0	0.0	0.0	0.0
75–79	+0.0	0.0	0.0	0.0
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 8 (75% of the national 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	+1.5	59.5	71.2	78.4
4	+0.7	33.4	39.8	47.2
8	–0.1	26.3	30.2	38.3
16	–0.1	17.7	22.0	28.1
32	+0.4	12.9	15.5	20.3
64	+0.1	9.0	10.8	14.4
128	+0.2	6.0	7.1	9.9
256	+0.2	4.4	5.2	6.7
512	+0.1	3.1	3.7	4.9
1,024	+0.0	2.3	2.6	3.3
2,048	+0.0	1.6	1.9	2.4
4,096	+0.0	1.1	1.3	1.7
8,192	+0.1	0.8	1.0	1.2
16,384	+0.1	0.5	0.7	0.8

Figure 11 (75% of the national 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	0.9	21.5	0.4	77.3	78.1	–90.5
5–9	3.5	18.8	2.8	74.8	78.3	–56.0
10–14	8.9	13.4	9.6	68.1	77.0	+22.9
15–19	14.6	7.7	19.7	58.0	72.6	+12.0
20–24	18.1	4.2	32.8	44.8	62.9	–47.0
25–29	19.9	2.4	42.2	35.5	55.4	–88.7
30–34	21.4	1.0	50.8	26.8	48.2	–127.5
35–39	22.0	0.4	57.2	20.5	42.5	–155.8
40–44	22.2	0.1	62.1	15.6	37.8	–178.0
45–49	22.3	0.0	66.0	11.7	34.0	–195.2
50–54	22.3	0.0	68.8	8.9	31.2	–207.9
55–59	22.3	0.0	71.2	6.4	28.8	–218.8
60–64	22.3	0.0	72.8	4.8	27.2	–225.9
65–69	22.3	0.0	74.5	3.2	25.5	–233.4
70–74	22.3	0.0	76.1	1.6	23.9	–240.7
75–79	22.3	0.0	76.8	0.8	23.2	–243.8
80–84	22.3	0.0	77.4	0.2	22.6	–246.7
85–89	22.3	0.0	77.5	0.1	22.5	–247.0
90–94	22.3	0.0	77.7	0.0	22.3	–247.6
95–100	22.3	0.0	77.7	0.0	22.3	–247.6

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

Figure 12 (75% of the national line): For a given score cut-off, the percentage of all households who are targeted (that is, have a score equal to or less than the cut-off), the percentage of targeted households who are poor (that is, below the poverty line), the percentage of poor households who are targeted, and the number of poor households who are successful targeted (coverage) 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	1.3	68.3	3.9	2.2:1
5–9	6.3	55.2	15.6	1.2:1
10–14	18.5	48.2	40.0	0.9:1
15–19	34.3	42.6	65.4	0.7:1
20–24	50.9	35.5	81.0	0.6:1
25–29	62.1	32.1	89.1	0.5:1
30–34	72.2	29.6	95.7	0.4:1
35–39	79.1	27.8	98.4	0.4:1
40–44	84.3	26.3	99.4	0.4:1
45–49	88.3	25.3	99.9	0.3:1
50–54	91.1	24.5	100.0	0.3:1
55–59	93.6	23.9	100.0	0.3:1
60–64	95.2	23.5	100.0	0.3:1
65–69	96.8	23.1	100.0	0.3:1
70–74	98.4	22.7	100.0	0.3:1
75–79	99.2	22.5	100.0	0.3:1
80–84	99.8	22.4	100.0	0.3:1
85–89	99.9	22.4	100.0	0.3:1
90–94	100.0	22.3	100.0	0.3:1
95–100	100.0	22.3	100.0	0.3:1

Tables for
150% of the National Poverty Line

Figure 4 (150% of the national 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	94.3
5–9	93.8
10–14	87.5
15–19	82.1
20–24	71.2
25–29	59.0
30–34	48.0
35–39	40.6
40–44	34.7
45–49	20.7
50–54	10.7
55–59	7.3
60–64	3.3
65–69	4.2
70–74	1.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Figure 7 (150% of the national line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) 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	–2.4	2.1	2.3	3.0
5–9	+6.8	2.3	2.7	3.5
10–14	+0.7	1.3	1.5	1.9
15–19	–1.4	1.3	1.5	1.9
20–24	+1.7	1.6	1.9	2.5
25–29	+0.5	2.1	2.5	3.3
30–34	–5.9	4.0	4.2	4.6
35–39	+2.9	2.7	3.2	4.3
40–44	+1.5	2.9	3.5	4.9
45–49	–2.5	2.9	3.5	4.6
50–54	+1.2	2.3	2.8	3.5
55–59	+2.6	1.9	2.3	2.8
60–64	–0.7	2.6	3.0	4.0
65–69	+2.1	1.4	1.7	2.3
70–74	+1.0	0.0	0.0	0.0
75–79	+0.0	0.0	0.0	0.0
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 8 (150% of the national 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	–0.6	65.3	73.7	90.0
4	+0.0	35.0	42.4	51.6
8	–1.0	25.7	30.0	38.7
16	–0.3	18.4	22.8	29.0
32	–0.3	12.6	14.7	19.8
64	–0.1	9.3	10.9	14.3
128	+0.0	6.4	7.6	9.9
256	+0.2	4.6	5.6	7.0
512	+0.3	3.3	4.0	5.2
1,024	+0.2	2.3	2.7	3.8
2,048	+0.2	1.6	1.9	2.5
4,096	+0.2	1.2	1.4	1.7
8,192	+0.2	0.8	1.0	1.3
16,384	+0.2	0.6	0.7	0.9

Figure 11 (150% of the national 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	1.2	57.8	0.0	41.0	42.2	–95.8
5–9	5.7	53.2	0.6	40.4	46.1	–79.6
10–14	16.2	42.7	2.3	38.7	54.9	–41.1
15–19	29.4	29.6	4.9	36.1	65.5	+7.9
20–24	41.1	17.9	9.8	31.2	72.3	+56.1
25–29	47.8	11.1	14.2	26.8	74.6	+75.9
30–34	53.3	5.7	19.0	22.1	75.3	+67.9
35–39	55.9	3.0	23.2	17.8	73.8	+60.7
40–44	57.6	1.3	26.7	14.3	72.0	+54.7
45–49	58.5	0.5	29.8	11.3	69.8	+49.5
50–54	58.8	0.2	32.3	8.7	67.5	+45.1
55–59	58.9	0.1	34.7	6.4	65.3	+41.2
60–64	58.9	0.0	36.2	4.8	63.7	+38.6
65–69	59.0	0.0	37.9	3.2	62.1	+35.8
70–74	59.0	0.0	39.5	1.6	60.5	+33.1
75–79	59.0	0.0	40.2	0.8	59.8	+31.9
80–84	59.0	0.0	40.8	0.2	59.2	+30.8
85–89	59.0	0.0	40.9	0.1	59.1	+30.6
90–94	59.0	0.0	41.0	0.0	59.0	+30.4
95–100	59.0	0.0	41.0	0.0	59.0	+30.4

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

Figure 12 (150% of the national line): For a given score cut-off, the percentage of all households who are targeted (that is, have a score equal to or less than the cut-off), the percentage of targeted households who are poor (that is, below the poverty line), the percentage of poor households who are targeted, and the number of poor households who are successful targeted (coverage) 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	1.3	96.2	2.1	25.6:1
5–9	6.3	90.4	9.7	9.4:1
10–14	18.5	87.6	27.5	7.0:1
15–19	34.3	85.7	49.8	6.0:1
20–24	50.9	80.7	69.7	4.2:1
25–29	62.1	77.1	81.1	3.4:1
30–34	72.2	73.7	90.3	2.8:1
35–39	79.1	70.7	94.9	2.4:1
40–44	84.3	68.3	97.7	2.2:1
45–49	88.3	66.3	99.2	2.0:1
50–54	91.1	64.5	99.7	1.8:1
55–59	93.6	63.0	99.9	1.7:1
60–64	95.2	61.9	99.9	1.6:1
65–69	96.8	60.9	100.0	1.6:1
70–74	98.4	59.9	100.0	1.5:1
75–79	99.2	59.5	100.0	1.5:1
80–84	99.8	59.1	100.0	1.4:1
85–89	99.9	59.0	100.0	1.4:1
90–94	100.0	59.0	100.0	1.4:1
95–100	100.0	59.0	100.0	1.4:1

Tables for
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	56.0
5–9	49.8
10–14	33.5
15–19	26.3
20–24	18.7
25–29	11.7
30–34	7.3
35–39	3.8
40–44	3.3
45–49	0.2
50–54	1.0
55–59	0.0
60–64	0.0
65–69	0.0
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Figure 7 (USAID "extreme" line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) 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	–7.8	6.8	7.3	9.1
5–9	+4.3	3.0	3.6	4.8
10–14	–1.5	1.8	2.1	2.9
15–19	–1.1	1.5	1.8	2.4
20–24	+1.8	1.3	1.5	2.0
25–29	+4.1	1.0	1.2	1.6
30–34	–3.8	2.6	2.8	3.0
35–39	–5.8	3.8	4.0	4.5
40–44	+0.9	0.9	1.0	1.3
45–49	–1.3	1.1	1.2	1.4
50–54	+0.8	0.3	0.3	0.4
55–59	+0.0	0.0	0.0	0.0
60–64	+0.0	0.0	0.0	0.0
65–69	+0.0	0.0	0.0	0.0
70–74	+0.0	0.0	0.0	0.0
75–79	+0.0	0.0	0.0	0.0
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 8 (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	+0.7	53.6	65.5	76.1
4	–0.1	31.2	35.1	46.1
8	–0.7	23.7	27.6	34.1
16	–0.5	16.5	19.9	24.9
32	–0.3	11.4	13.6	18.0
64	–0.2	8.0	9.7	12.0
128	–0.1	5.7	6.6	9.0
256	–0.1	3.9	4.6	6.1
512	–0.2	2.8	3.2	4.1
1,024	–0.3	2.1	2.3	2.9
2,048	–0.3	1.4	1.8	2.2
4,096	–0.3	1.0	1.2	1.5
8,192	–0.3	0.7	0.9	1.1
16,384	–0.3	0.5	0.6	0.7

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	0.8	16.5	0.5	82.2	83.0	–88.1
5–9	3.1	14.2	3.2	79.5	82.6	–45.3
10–14	7.4	9.9	11.2	71.6	78.9	+35.4
15–19	11.9	5.4	22.4	60.3	72.2	–29.6
20–24	14.6	2.7	36.4	46.3	60.9	–110.4
25–29	15.5	1.8	46.5	36.2	51.7	–169.3
30–34	16.5	0.8	55.7	27.0	43.6	–222.1
35–39	17.1	0.2	62.0	20.7	37.8	–258.9
40–44	17.2	0.1	67.1	15.6	32.8	–288.1
45–49	17.3	0.0	71.0	11.7	29.0	–310.7
50–54	17.3	0.0	73.9	8.9	26.1	–327.2
55–59	17.3	0.0	76.3	6.4	23.7	–341.2
60–64	17.3	0.0	77.9	4.8	22.1	–350.4
65–69	17.3	0.0	79.5	3.2	20.5	–360.2
70–74	17.3	0.0	81.2	1.6	18.8	–369.5
75–79	17.3	0.0	81.9	0.8	18.1	–373.6
80–84	17.3	0.0	82.5	0.2	17.5	–377.2
85–89	17.3	0.0	82.6	0.1	17.4	–377.7
90–94	17.3	0.0	82.7	0.0	17.3	–378.5
95–100	17.3	0.0	82.7	0.0	17.3	–378.5

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 (that is, have a score equal to or less than the cut-off), the percentage of targeted households who are poor (that is, below the poverty line), the percentage of poor households who are targeted, and the number of poor households who are successful targeted (coverage) 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	1.3	62.7	4.6	1.7:1
5–9	6.3	49.4	18.1	1.0:1
10–14	18.5	39.8	42.6	0.7:1
15–19	34.3	34.6	68.6	0.5:1
20–24	50.9	28.6	84.3	0.4:1
25–29	62.1	25.0	89.8	0.3:1
30–34	72.2	22.9	95.7	0.3:1
35–39	79.1	21.6	98.8	0.3:1
40–44	84.3	20.4	99.6	0.3:1
45–49	88.3	19.6	99.9	0.2:1
50–54	91.1	19.0	100.0	0.2:1
55–59	93.6	18.5	100.0	0.2:1
60–64	95.2	18.2	100.0	0.2:1
65–69	96.8	17.9	100.0	0.2:1
70–74	98.4	17.6	100.0	0.2:1
75–79	99.2	17.4	100.0	0.2:1
80–84	99.8	17.3	100.0	0.2:1
85–89	99.9	17.3	100.0	0.2:1
90–94	100.0	17.3	100.0	0.2:1
95–100	100.0	17.3	100.0	0.2:1

Tables for
$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	94.3
5–9	88.8
10–14	79.8
15–19	72.4
20–24	61.7
25–29	48.8
30–34	36.9
35–39	30.1
40–44	25.6
45–49	11.1
50–54	7.6
55–59	3.4
60–64	2.5
65–69	0.9
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

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) 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	+3.7	3.5	4.1	5.0
5–9	+5.6	2.5	2.9	3.6
10–14	+0.6	1.6	1.9	2.5
15–19	–2.2	1.8	1.9	2.1
20–24	+2.4	1.6	1.9	2.7
25–29	+1.1	2.0	2.3	3.2
30–34	–5.9	4.0	4.2	4.4
35–39	–2.5	2.6	3.1	3.8
40–44	+3.1	2.6	3.3	4.3
45–49	–3.0	2.6	2.9	3.8
50–54	+0.9	2.1	2.5	3.1
55–59	+0.6	1.5	1.8	2.3
60–64	+2.5	0.0	0.0	0.0
65–69	–1.1	1.4	1.7	2.3
70–74	+0.0	0.0	0.0	0.0
75–79	+0.0	0.0	0.0	0.0
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 8 ($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	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+1.8	67.8	74.9	88.8
4	–0.3	38.4	43.8	55.7
8	–1.6	27.7	32.9	42.1
16	–0.9	19.2	23.1	30.2
32	–0.5	13.5	16.0	20.0
64	–0.2	9.5	11.4	14.2
128	–0.1	6.8	8.0	10.4
256	+0.0	5.1	6.0	8.1
512	+0.0	3.5	4.3	5.4
1,024	–0.1	2.5	3.1	3.9
2,048	–0.1	1.8	2.1	2.7
4,096	–0.1	1.2	1.5	1.9
8,192	–0.1	0.9	1.1	1.4
16,384	–0.1	0.6	0.8	1.0

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	1.2	49.4	0.1	49.4	50.5	–95.2
5–9	5.5	45.1	0.9	48.6	54.1	–76.7
10–14	15.0	35.5	3.5	45.9	60.9	–33.7
15–19	26.7	23.8	7.6	41.9	68.6	+20.7
20–24	36.7	13.8	14.3	35.2	71.9	+71.8
25–29	42.1	8.4	19.9	29.6	71.7	+60.6
30–34	46.4	4.1	25.8	23.7	70.1	+48.9
35–39	48.6	1.9	30.5	19.0	67.6	+39.6
40–44	49.7	0.8	34.6	14.9	64.6	+31.5
45–49	50.2	0.3	38.0	11.4	61.7	+24.7
50–54	50.4	0.1	40.7	8.8	59.2	+19.4
55–59	50.5	0.0	43.1	6.4	56.9	+14.7
60–64	50.5	0.0	44.7	4.8	55.3	+11.6
65–69	50.5	0.0	46.3	3.2	53.7	+8.3
70–74	50.5	0.0	47.9	1.6	52.1	+5.1
75–79	50.5	0.0	48.6	0.8	51.4	+3.7
80–84	50.5	0.0	49.3	0.2	50.7	+2.4
85–89	50.5	0.0	49.4	0.1	50.6	+2.3
90–94	50.5	0.0	49.5	0.0	50.5	+2.0
95–100	50.5	0.0	49.5	0.0	50.5	+2.0

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 (that is, have a score equal to or less than the cut-off), the percentage of targeted households who are poor (that is, below the poverty line), the percentage of poor households who are targeted, and the number of poor households who are successful targeted (coverage) 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	1.3	91.1	2.3	10.2:1
5–9	6.3	86.2	10.8	6.2:1
10–14	18.5	80.9	29.7	4.2:1
15–19	34.3	77.9	52.9	3.5:1
20–24	50.9	72.0	72.6	2.6:1
25–29	62.1	67.9	83.4	2.1:1
30–34	72.2	64.2	91.8	1.8:1
35–39	79.1	61.4	96.2	1.6:1
40–44	84.3	58.9	98.4	1.4:1
45–49	88.3	56.9	99.5	1.3:1
50–54	91.1	55.3	99.8	1.2:1
55–59	93.6	53.9	99.9	1.2:1
60–64	95.2	53.0	99.9	1.1:1
65–69	96.8	52.2	100.0	1.1:1
70–74	98.4	51.3	100.0	1.1:1
75–79	99.2	50.9	100.0	1.0:1
80–84	99.8	50.6	100.0	1.0:1
85–89	99.9	50.6	100.0	1.0:1
90–94	100.0	50.5	100.0	1.0:1
95–100	100.0	50.5	100.0	1.0:1

Tables for
$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	100.0
5–9	99.2
10–14	97.8
15–19	95.7
20–24	92.0
25–29	87.3
30–34	80.4
35–39	70.8
40–44	70.4
45–49	51.6
50–54	39.3
55–59	33.9
60–64	20.6
65–69	17.4
70–74	4.7
75–79	4.4
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

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) 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	+3.5	1.5	1.8	2.4
10–14	+0.9	0.7	0.8	1.0
15–19	+0.6	0.7	0.9	1.0
20–24	+2.9	1.0	1.2	1.5
25–29	–2.0	1.6	1.6	1.9
30–34	–4.7	3.1	3.2	3.5
35–39	–1.2	2.5	2.9	3.7
40–44	+0.5	2.8	3.3	4.4
45–49	–1.9	3.6	4.5	5.6
50–54	–0.5	4.3	5.0	6.6
55–59	+6.5	3.9	4.9	6.7
60–64	+2.3	4.1	4.8	6.1
65–69	+10.5	2.6	3.1	4.0
70–74	–8.6	6.5	7.0	8.0
75–79	+1.1	2.6	3.0	3.8
80–84	–7.5	5.8	6.5	7.3
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 ($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	–0.8	64.4	74.0	87.4
4	–0.1	25.2	30.4	38.6
8	–0.3	18.0	22.2	28.8
16	–0.2	13.2	15.5	20.8
32	+0.1	9.3	11.2	14.5
64	+0.1	6.8	7.9	10.1
128	+0.1	4.9	5.8	7.7
256	+0.2	3.5	4.1	5.4
512	+0.1	2.3	2.8	3.7
1,024	+0.2	1.8	2.1	2.8
2,048	+0.2	1.2	1.4	1.9
4,096	+0.2	0.9	1.1	1.4
8,192	+0.2	0.6	0.7	0.9
16,384	+0.2	0.4	0.5	0.7

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	1.3	78.3	0.0	20.4	21.7	–96.8
5–9	6.2	73.4	0.2	20.2	26.4	–84.3
10–14	18.0	61.6	0.5	19.9	37.9	–54.1
15–19	33.0	46.7	1.3	19.1	52.0	–15.6
20–24	47.9	31.7	3.1	17.3	65.2	+24.1
25–29	57.8	21.8	4.3	16.1	73.9	+50.5
30–34	66.3	13.3	5.9	14.5	80.8	+74.0
35–39	71.4	8.2	7.8	12.6	84.0	+89.1
40–44	75.0	4.7	9.4	11.0	86.0	+88.2
45–49	77.2	2.4	11.1	9.3	86.5	+86.0
50–54	78.3	1.3	12.9	7.5	85.8	+83.8
55–59	78.9	0.7	14.6	5.8	84.7	+81.6
60–64	79.2	0.4	15.9	4.5	83.7	+80.0
65–69	79.4	0.2	17.5	2.9	82.3	+78.1
70–74	79.5	0.1	18.9	1.5	81.0	+76.2
75–79	79.6	0.0	19.6	0.8	80.4	+75.4
80–84	79.6	0.0	20.2	0.2	79.8	+74.6
85–89	79.6	0.0	20.3	0.1	79.7	+74.5
90–94	79.6	0.0	20.4	0.0	79.6	+74.4
95–100	79.6	0.0	20.4	0.0	79.6	+74.4

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 (that is, have a score equal to or less than the cut-off), the percentage of targeted households who are poor (that is, below the poverty line), the percentage of poor households who are targeted, and the number of poor households who are successful targeted (coverage) 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	1.3	100.0	1.6	Only poor targeted
5–9	6.3	97.6	7.8	40.3:1
10–14	18.5	97.1	22.6	33.8:1
15–19	34.3	96.2	41.4	25.1:1
20–24	50.9	94.0	60.1	15.6:1
25–29	62.1	93.1	72.6	13.5:1
30–34	72.2	91.9	83.3	11.3:1
35–39	79.1	90.2	89.7	9.2:1
40–44	84.3	88.9	94.2	8.0:1
45–49	88.3	87.4	96.9	6.9:1
50–54	91.1	85.9	98.3	6.1:1
55–59	93.6	84.4	99.1	5.4:1
60–64	95.2	83.3	99.5	5.0:1
65–69	96.8	82.0	99.7	4.5:1
70–74	98.4	80.8	99.9	4.2:1
75–79	99.2	80.2	99.9	4.1:1
80–84	99.8	79.8	100.0	3.9:1
85–89	99.9	79.7	100.0	3.9:1
90–94	100.0	79.6	100.0	3.9:1
95–100	100.0	79.6	100.0	3.9:1