

A Simple Poverty Scorecard for Sierra Leone

Mark Schreiner

29 March 2011

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

Abstract

This study uses Sierra Leone's 2003/4 Integrated Household Survey to construct an easy-to-use scorecard that estimates the likelihood that a household has expenditure below a given poverty line. The scorecard uses ten simple indicators that field agents can quickly collect and verify. Poverty scores can be computed on paper in the field in five to ten minutes. The scorecard's accuracy and precision are reported for a range of poverty lines. The poverty scorecard is a practical way for pro-poor programs in Sierra Leone to estimate poverty rates, track changes in poverty rates over time, and target services.

Acknowledgements

Funding is from Ford Foundation via a grant to Grameen Foundation. Data are from Statistics Sierra Leone, with thanks to Mohamed King Koroma and Olive Odia.

Expenditure data come from Geoffrey Greenwell and Rose Mungai. Thanks also to Sharlene Brown, Frank DeGiovanni, Olivier Dupriez, Dane Shikman, and Jeff Toohig. The poverty scorecard is the same as what Grameen Foundation calls the Progress out of Poverty Index™. The PPI™ is a performance-management tool that Grameen Foundation promotes to help institutions achieve their social objectives more effectively.

Author

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

Figure 1: A simple poverty scorecard for Sierra Leone

<u>Entity</u>	<u>Name</u>	<u>ID</u>	<u>Date</u> (DD/MM/YY)
Member:	_____	_____	Joined: _____
Field agent:	_____	_____	Today: _____
Service point:	_____	_____	Household size: _____

Indicator	Value	Points	Score
1. How many members does the household have?	A. Ten or more	0	
	B. Seven, eight, or nine	9	
	C. Six	13	
	D. Five	16	
	E. Four	21	
	F. One, two, or three	28	
2. Are all household members ages 6 to 13 in school now?	A. No	0	
	B. Yes, or no one aged 6 to 13	5	
3. What was the activity of the female head/spouse in her main occupation in the past 12 months?	A. No female head/spouse	0	
	B. Agriculture, forestry, mining, or quarrying	3	
	C. Other, or does not work	9	
4. How many rooms does the household occupy (exclude bathrooms, toilets, kitchen, pantry, hall, and storage)?	A. One	0	
	B. Two	4	
	C. Three or more	7	
5. What is the main flooring material?	A. Earth/mud, stone/brick, or other	0	
	B. Wood, or cement/concrete	3	
6. What is the main construction material of the outside walls?	A. Stone/burnt bricks, or other	0	
	B. Mud/mud bricks, or wood	11	
	C. Cement/sandcrete, or corrugated iron sheets	14	
7. What type of toilet is used by the household?	A. Bush/river, none, or other	0	
	B. Bucket, common pit, or VIP	1	
	C. Private pit, common flush, or flush toilet	7	
8. What is the main source of lighting for the dwelling?	A. Generator, kerosene, gas lamp, candles/torch light, or other	0	
	B. Electricity (mains)	6	
9. What is the main fuel used by the household for cooking?	A. Wood, or other	0	
	B. Charcoal	4	
	C. Gas, kerosene, or electricity	6	
10. How many radios, radio cassettes, record players, or 3-in-1 radio cassettes do members of the household own?	A. None	0	
	B. One	4	
	C. Two or more	14	

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

Simple poverty scorecard for Sierra Leone (no points)

<u>Entity</u>	<u>Name</u>	<u>ID</u>	<u>Date</u> (DD/MM/YY)
Member:	_____	_____	Joined: _____
Field agent:	_____	_____	Today: _____
Service point:	_____	_____	Household size: _____

Indicator	Value
1. How many members does the household have?	A. Ten or more B. Seven, eight, or nine C. Six D. Five E. Four F. One, two, or three
2. Are all household members ages 6 to 13 in school now?	A. No B. Yes, or no one aged 6 to 13
3. What was the activity of the female head/spouse in her main occupation in the past 12 months?	A. No female head/spouse B. Agriculture, forestry, mining, or quarrying C. Other, or does not work
4. How many rooms does the household occupy (exclude bathrooms, toilets, kitchen, pantry, hall, and storage)?	A. One B. Two C. Three or more
5. What is the main flooring material?	A. Earth/mud, stone/brick, or other B. Wood, or cement/concrete
6. What is the main construction material of the outside walls?	A. Stone/burnt bricks, or other B. Mud/mud bricks, or wood C. Cement/sandcrete, or corrugated iron sheets
7. What type of toilet is used by the household?	A. Bush/river, none, or other B. Bucket, common pit, or VIP C. Private pit, common flush, or flush toilet
8. What is the main source of lighting for the dwelling?	A. Generator, kerosene, gas lamp, candles/torch light, or other B. Electricity (mains)
9. What is the main fuel used by the household for cooking?	A. Wood, or other B. Charcoal C. Gas, kerosene, or electricity
10. How many radios, radio cassettes, record players, or 3-in-1 radio cassettes do members of the household own?	A. None B. One C. Two or more

A Simple Poverty Scorecard for Sierra Leone

1. Introduction

This paper presents an easy-to-use poverty scorecard that pro-poor programs in Sierra Leone can use to estimate the likelihood that a household has expenditure below a given poverty line, to estimate groups' poverty rates at a point in time, to track changes in groups' poverty rates over time, and to target services.

The direct approach to poverty measurement via surveys is difficult and costly, asking households about a lengthy list of expenditure items. As a case in point, Sierra Leone's 2003/4 Integrated Household Survey (IHS) runs 136 pages. The expenditure module covers almost 500 items, and each household is visited seven times over the course of a month. An example set of questions are "Did the household consume any home-produced maize-cob (fresh) in the last 12 months? If yes, how much was consumed since my last visit? For how much could you sell one unit of maize-cob (fresh) now? Now then, did the household consume any maize-flour/dough in the last 12 months? . . .".

In contrast, the indirect approach via poverty scoring is simple, quick, and inexpensive. It uses ten verifiable indicators (such as "What is the main flooring material?" or "Are all household members ages 6 to 13 in school now?") to get a score that is highly correlated with poverty status as measured by the exhaustive survey.

The poverty scorecard here differs from “proxy means tests” (Coady, Grosh, and Hoddinott, 2002) in that it is tailored to the capabilities and purposes not of national governments but rather of local, pro-poor organizations. The feasible poverty-measurement options for 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 and precision are unknown.

Pro-poor organizations can use the poverty scorecard to measure the share of their participants below a given poverty line, such as the Millennium Development Goals’ \$1.25/day at 2005 purchase-power parity. USAID microenterprise partners can use it to report how many of its participants are among the poorest half of people below the national poverty line. Organizations can also use it to measure movement across a poverty line. In all these cases, the poverty scorecard provides an expenditure-based, objective tool with known accuracy. While expenditure surveys are costly even for governments, some small, local organizations may be able to implement an inexpensive scorecard 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 at the local level. This is 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 scorecards are usually about as accurate as complex ones.

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

The scorecard (Figure 1) is based on the 2003/4 IHS conducted by Statistics Sierra Leone. Indicators are selected to be:

- Inexpensive to collect, easy to answer quickly, and simple to verify
- Associated 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-adult-equivalent or per-capita expenditure below a given poverty line.

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

Third, poverty scoring can 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 organizations 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 expenditure data and Sierra Leone'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 households in the 2003/4 IHS, and its accuracy is validated on the other half.

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 relationship between indicators and poverty will be the same as in the data used to build the scorecard. 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 average difference between scorecard estimates of groups' poverty rates and the true rates at a point in time is +1.1 percentage points. These differences are due to sampling variation and not bias; the average of each difference would be zero if the whole 2003/4 IHS were to be repeatedly redrawn and divided into sub-samples before repeating the entire process of construction and calibration.

The 90-percent confidence intervals for these estimates are ± 0.8 percentage points or less. For $n = 1,024$, these intervals are ± 3.0 percentage points or less.

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 is a summary.

¹ Important examples include nationally representative samples after 2003/4 or non-nationally representative sub-groups (Tarozzi and Deaton, 2007).

2. Data and poverty lines

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

2.1 Data

The scorecard is based on data from the 3,702 households in the 2003/4 IHS surveyed by Statistics Sierra Leone from 24 April 2003 to 26 April 2004, excluding 18 households with missing values for aggregate expenditure. Expenditure data were graciously provided by Geoffrey Greenwell and by Rose Mungai of the World Bank. The World Bank expenditure data is used here.

For the purposes of poverty scoring, the households in the 2003/4 IHS 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 measuring accuracy with data not used in construction or calibration

2.2 Poverty rates and poverty lines

2.2.1 Rates

As a general definition, a *poverty rate* is the share of people in a group who live in households whose total household expenditure (divided by the number of household members or by the number of adult equivalents) 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 or by the number of adult equivalents 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 expenditure (or per-adult-equivalent expenditure) above a poverty line (it is “non-poor”) and that the second household has per-capita expenditure (or per-adult-equivalent expenditure) 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 more 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 those people are arranged in households, so governments typically report person-level poverty rates.

If an organization has only one “participant” per household, however, then the household-level rate 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 Sierra Leone at both the household- and person-level for the country as a whole, for its four regions, and for the construction/calibration and validation sub-samples used for scoring.

The poverty scorecard is constructed using the 2003/4 IHS and household-level lines, scores are calibrated to household-level poverty likelihoods, and accuracy is measured for household-level rates. Person-level poverty rates can be estimated as the household-size-weighted average of the household-level poverty likelihoods. It is also possible to construct a scorecard based on person-level lines, calibrate scores to person-level likelihoods, and measure accuracy for person-level rates, but it is not done here.

2.2.2 Poverty lines

Sierra Leone’s food (“core”) poverty line is defined as the cost of 2,700 kilocalories from a food basket consistent with that consumed by the poorest 20 percent of people in the 2003/4 IHS (Greenwell, 2005). Prices are adjusted to May 2003 across regions in Sierra Leone and across months of the IHS with an index for a basket of food and non-food relevant for the poorest 20 percent of people. The average food line is SLL1,133 per adult equivalent per day, giving a household-level poverty rate for all of Sierra Leone of 19.1 percent and person-level rate of 22.4 percent (Figure 2).

The poverty lines and poverty rates here differ somewhat from those in Greenwell (2005), Statistics Sierra Leone (2007), and World Bank (2009) because:

- The 2003/4 IHS data provided by Statistics Sierra Leone omits aggregate household expenditure, price deflators, and poverty lines
- The data used here from the World Bank for expenditure, price indices, and poverty lines has undocumented adjustments
- Correct sampling weights were not available to Greenwell (2005)
- Statistics Sierra Leone (2007) and World Bank (2009) sometimes report poverty rates that (incorrectly) weigh households equally or by adult equivalents

The national poverty line (sometimes called here “100% of the national line”) is defined as average total expenditure (food plus non-food) for households whose food expenditure is with ± 10 percent of the food poverty line (Greenwell, 2005). For Sierra Leone as a whole, the national line is SLL2,363 per adult equivalent per day, giving a household-level poverty rate of 61.9 percent and a person-level rate of 66.3 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:

- Food
- 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 expenditure of people (not adult equivalents nor households) below the national line (U.S. Congress, 2004). This median line is defined for each of Sierra Leone’s four regions.

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

- 2005 PPP exchange rate for “individual consumption expenditure by households” (World Bank, 2008): SLL 1396.21 per \$1.00
- Average all-Sierra Leone consumer price index for 2005 of 127.0233²
- Average all-Sierra Leone CPI for May 2003 of 100.0

Given this, the \$1.25/day 2005 PPP line for Sierra Leone as a whole during the 2003/4 IHS is (Sillers, 2006):

$$\begin{aligned} & (\text{2005 PPP exchange rate}) \cdot \$1.25 \cdot \left(\frac{\text{CPI}_{\text{May 2003}}}{\text{CPI}_{\text{2005 average}}} \right) = \\ & \left(\frac{\text{SLL}1,396.21}{\$1.00} \right) \cdot \$1.25 \cdot \left(\frac{100.0}{127.0233} \right) = \text{SLL}1,374. \end{aligned}$$

This line is adjusted for each household by multiplying by its price deflator and then dividing by the person-weighted average of all household-specific price deflators.

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

² This is based on monthly inflation rates for Sierra Leone’s four regions in the Bank of Sierra Leone’s *Annual Reports*. These monthly rates are converted to a CPI with a base of 100.0 in May 2003, weighting each region by its population share in the 2003/4 IHS.

3. Scorecard construction

For Sierra Leone, 75 potential indicators are initially prepared in the areas of:

- Family composition (such as household size)
- Education (such as school attendance by children)
- Employment (such as the main occupation of the female head/spouse)
- Housing (such as flooring material)
- Ownership of durable goods (such as radios and other music players)
- Agriculture (such as ownership of land or livestock)

Figure 3 lists the candidate indicators ordered by the entropy-based “uncertainty coefficient” (Goodman and Kruskal, 1979) that measures how well an indicator predicts poverty on its own.

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, the main fuel used for cooking 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 with the construction 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), including improvement in accuracy, likelihood of

acceptance by users (determined by simplicity, cost of collection, and “face validity” in terms of experience, theory, and common sense), sensitivity to changes in poverty status, variety among indicators, and verifiability.

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

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 poverty scorecard here applies to all of Sierra Leone. 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 (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 scoring 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 have similar targeting accuracy, 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 in Figure 1 is ready to be photocopied and can be used with a simple spreadsheet database (Microfinance Risk Management, L.L.C., 2011) that records identifying information, dates, indicator values, scores, and poverty likelihoods.

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

Of course, field agents must be trained. The quality of outputs depends on the quality of 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).³ IRIS Center (2007a) and Toohig (2008) are useful

³ If an organization does not want field agents to know the point values 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.

nuts-and-bolts guides for budgeting, training field agents and supervisors, logistics, sampling, interviewing, piloting, recording data, and controlling quality.

In particular, while collecting scorecard indicators is relatively easier than alternatives, it is still absolutely difficult. Training and explicit definitions for the 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 an automobile. 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 conditional cash-transfer program in Mexico, 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 poverty scoring in Sierra Leone.

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 questions that the organization wants to inform with the results of the poverty-scoring exercise.

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 well-defined group that is relevant to a particular business 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 who each have more than 7 million participants and who are applying a poverty scorecard similar to the one here (Chen and Schreiner, 2009b). Their design is that loan officers in a random sample of branches score all 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 50,000–100,000 participants each (far more than the typical pro-poor organization would need).

5. Estimates of household poverty likelihoods

The sum of scorecard points for a household is called the *score*. For Sierra Leone, 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 must be converted to *poverty likelihoods*, that is, probabilities of being below a poverty line. This is done via simple look-up tables. For the example of the national line, scores of 50–54 have a poverty likelihood of 57.3 percent, and scores of 55–59 have a poverty likelihood of 45.0 percent (Figure 4).

The poverty likelihood associated with a score varies by poverty line. For example, scores of 50–54 are associated with a poverty likelihood of 57.3 percent for the national line but 12.8 percent for the food line.⁴

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.

⁴ Starting with Figure 4, many figures have seven versions, one for each of the seven poverty lines. To keep them straight, they are grouped by poverty line. Single tables pertaining to all poverty lines are placed with the tables for the national line.

For the example of the national line (Figure 5), there are 10,858 (normalized) households in the calibration sub-sample with a score of 50–54, of whom 6,216 (normalized) are below the poverty line. The estimated poverty likelihood associated with a score of 50–54 is then 57.3 percent, because $6,216 \div 10,858 = 57.3$ percent.

To illustrate with the national line and a score of 55–59, there are 8,659 (normalized) households in the calibration sample, of whom 3,893 (normalized) are below the line (Figure 5). Thus, the poverty likelihood for this score is $3,893 \div 8,659 = 45.0$ percent.

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

Figures 6a and 6b show, for all scores, the likelihood that expenditure falls in a range demarcated by two adjacent poverty lines.⁵ For example, the daily expenditure of an adult equivalent in a household with a score of 50–54 falls in the following ranges with probability (Figure 6a):

- 12.8 percent below the food line
- 23.7 percent between the food line and 75% of the national line
- 20.8 percent between 75% of the national line and 100% of the national line
- 31.7 percent between 100% of the national line and 150% of the national line
- 11.0 percent above 150% of the national line

⁵ Figure 6a is for the per-adult-equivalent national lines, and Figure 6b is for the USAID “extreme” per-person line and the 2005 PPP per-person lines.

For the poverty lines in per-capita terms, a household with a score of 50–54 falls in the following ranges with probability (Figure 6b):

- 17.7 percent below the USAID “extreme” line
- 19.3 percent between the USAID “extreme” and the \$1.25/day 2005 PPP lines
- 51.3 percent between the \$1.25/day and the \$2.50/day 2005 PPP lines
- 11.8 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 expenditure 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 Sierra Leone poverty scorecard are transformed coefficients from a Logit regression, scores are not converted to poverty likelihoods via the Logit formula of $2.718281828^{\text{score}} \times (1 + 2.718281828^{\text{score}})^{-1}$. 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 approach can also improve accuracy, especially with large samples.

5.2 Accuracy of estimates of households' poverty likelihoods

If the relationships between indicators and poverty do not change and if 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 poverty likelihood. The scorecard also produces unbiased estimates of poverty rates at a point in time, as well as unbiased estimates of changes in poverty rates between two points in time.⁶

Of course, the relationship between indicators and poverty does change to some unknown extent with time and also across sub-groups in Sierra Leone's population, so the scorecard will generally be biased when applied after April 2004 (the last month of fieldwork for the 2003/4 IHS) or when applied with non-nationally representative sub-groups.

How accurate are estimates of households' poverty likelihoods? To get a measurement of accuracy under the assumption that the scorecard is applied to a nationally representative sample in the period from April 2003 to April 2004, the

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

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 expenditure below a poverty line
- For each score, record the difference between the estimated poverty likelihood (Figure 4) and the true poverty likelihood in the bootstrap sample
- Repeat the previous three steps 1,000 times
- For each score, report the average difference between estimated and true poverty likelihoods across the 1,000 bootstrap samples
- For each score, report the two-sided interval containing the central 900, 950, or 990 differences between estimated and true poverty likelihoods

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

For the national line, the average poverty likelihood across bootstrap samples for scores of 50–54 in the validation sample is too low by 8.6 percentage points. For scores of 45–49, the estimate is too high by 2.8 percentage points.⁷

The 90-percent confidence interval for the differences for scores of 50–54 is ± 5.4 percentage points (Figure 7). This means that in 900 of 1,000 bootstraps, the difference between the estimate and the true value is between -14.0 and -3.2 percentage points (because $-8.6 - 5.4 = -14.0$, and $-8.6 + 5.4 = -3.2$). In 950 of 1,000 bootstraps

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, calibration, and validation.

(95 percent), the difference is -8.6 ± 5.6 percentage points, and in 990 of 1,000 bootstraps (99 percent), the difference is -8.6 ± 5.9 percentage points.

For several score ranges, Figure 7 shows differences—sometimes large ones—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 Sierra Leone’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.

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

Another possible source of differences between estimates and true values is overfitting. The scorecard here is unbiased, but it may still be *overfit* when applied after the end of the IHS fieldwork in April 2004. That is, it may fit the data from the 2003/4 IHS 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 2003/4 IHS. Or the scorecard may be overfit in the sense that it is sensitive to small changes in the relationships between indicators and poverty over time or when applied to non-nationally representative samples.

Overfitting can be mitigated by simplifying the scorecard and by not relying only on data but 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 come from non-scorecard sources such as changes in the relationships between indicators and poverty, changes in poverty lines, inconsistencies in data quality across time, and imperfections in cost-of-living adjustments across time and geography. 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 97.6, 90.5 and 81.3 percent (national line, Figure 4). The group's estimated poverty rate is the households' average poverty likelihood of $(97.6 + 90.5 + 81.3) \div 3 = 89.8$ percent.⁸

6.1 Accuracy of estimated poverty rates at a point in time

For the Sierra Leone scorecard applied to the validation sample with $n = 16,384$, the difference between the estimated poverty rate at a point in time for the national line and the true rate is +1.8 percentage points (Figure 9, summarizing Figure 8 across poverty lines). Across all seven lines, estimates differ from true values on average by 1.1 percentage points. At least part of these differences is due to sampling variation in the validation sample and in the division of the 2003/4 IHS 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 ± 0.8 percentage points or less (Figure 9). This means that in 900 of 1,000 bootstraps of this size, the difference

⁸ The group's poverty rate is *not* the poverty likelihood associated with the average score. Here, the average score of 30 is associated with a poverty likelihood of 90.5 percent. This obviously is different from the 89.8 percent that is the average of the three poverty likelihoods associated with each of the three scores.

between the estimate and the true value is within 0.8 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 $+1.8 - 0.6 = +1.2$ to $+1.8 + 0.6 = +2.4$ percentage points. This is because $+1.8$ is the average difference, and ± 0.6 is its 90-percent confidence interval. The average difference is $+1.8$ because the average scorecard estimate is too high by 1.8 percentage points; the average estimated poverty rate for the validation sample is 63.4 percent, but the true value is 61.6 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 poverty scorecards (Schreiner, 2008a), note that the textbook formula (Cochran, 1977) that relates confidence intervals with standard errors in the case of direct measurement of rates is $c = +/- z \cdot \sigma$, where:

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

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

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

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 62.1 percent (the poverty rate in the construction/calibration sample in Figure 2 for the national line), the confidence

interval c is $+/- z \cdot \sqrt{\frac{p \cdot (1 - p)}{n}} = +/- 1.64 \cdot \sqrt{\frac{0.621 \cdot (1 - 0.621)}{16,384}} = +/- 0.622$ percentage

points.

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

sample. For $n = 16,384$ and the national line, the 90-percent confidence interval is 0.595 percentage points.⁹

Thus, the 90-percent confidence interval with $n = 16,384$ is 0.595 percentage points for the Sierra Leone poverty scorecard and 0.622 percentage points for direct measurement. The ratio of the two intervals is $0.595 \div 0.622 = 0.96$.

Now consider the same case, but with $n = 8,192$. The confidence interval under direct measurement is $\pm 1.64 \cdot \sqrt{\frac{0.621 \cdot (1 - 0.621)}{8,192}} = \pm 0.879$ percentage points. The empirical confidence interval with the Sierra Leone poverty scorecard (Figure 8) is 0.835 percentage points. Thus for $n = 8,192$, the ratio of the two intervals is $0.835 \div 0.879 = 0.95$.

This ratio of 0.95 for $n = 8,182$ is not far from the ratio of 0.96 for $n = 16,384$. Across all sample sizes of 256 or more in Figure 8, the average ratio turns out to be 0.94, implying that confidence intervals for indirect estimates of poverty rates via the Sierra Leone scorecard and this poverty line are slightly narrower than confidence intervals for direct estimates via the 2003/4 IHS. This 0.94 appears in Figure 9 as the “ α factor” because if $\alpha = 0.94$, then the formula relating confidence intervals c and standard errors σ for the Sierra Leone poverty scorecard is $c = \pm z \cdot \alpha \cdot \sigma$. That is, formula for the standard error σ for point-in-time estimates of poverty rates via scoring

$$\text{is } \alpha \cdot \sqrt{\frac{p \cdot (1 - p)}{n}}.$$

⁹ Due to rounding, Figure 8 displays 0.6, not 0.595.

In general, α can be more or less than 1.00. When α is less than 1.00, it means that the scorecard is more precise than direct measurement, and vice versa when α is more than 1.00. The α factor is less than 1.00 for five of the 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 $c = 0.04620$ and $z = 1.64$ (90-percent confidence). Then the formula gives $n = \left(\frac{0.94 \cdot 1.64}{0.04620}\right)^2 \cdot 0.621 \cdot (1 - 0.621) = 263$, close to 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 Sierra Leone, its poverty lines, its poverty rates, and this scorecard. The derivation of the formulas, however, is valid for any poverty scorecard 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 scorecard is as precise as direct measurement, if the expected (before measurement) poverty rate is 50 percent, and if the confidence level is 90 percent, then $n = 300$ implies a confidence interval of ± 2.2 percentage points. In fact, USAID has not specified confidence levels or intervals. Furthermore, the expected poverty rate may not be 50 percent, and the scorecard could be more or less precise than direct measurement.

In practice after the end of fieldwork for the IHS in April 2004, 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 61.9-percent national average in the 2003/4 IHS in Figure 2), look up α (here, 0.94), 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, $n = \left(\frac{0.94 \cdot 1.64}{0.02} \right)^2 \cdot 0.619 \cdot (1 - 0.619) = 1,402$.

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 April 2004 will resemble that in the 2003/4 IHS 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 2003/4 IHS, this paper cannot test estimates of change over time for Sierra Leone, 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 97.6, 90.5, and 81.3 percent (national line, Figure 4). The group's baseline estimated poverty rate is the households' average poverty likelihood of $(97.6 + 90.5 + 81.3) \div 3 = 89.8$ 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 98.6, 85.2, and 74.1 percent, national line, Figure 4). Their average poverty likelihood at follow-up is now $(98.6 + 85.2 + 74.1) \div 3 = 86.0$ percent, an improvement of $89.8 - 86.0 = 3.8$ percentage points.¹²

This suggests that about one in 26 participants in this hypothetical example crossed the poverty line in 2011.¹³ Among those who started below the line, about one in 25 ($3.8 \div 89.8 = 4.2$ 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.

7.3 Accuracy for estimated change in two independent samples

With only the 2003/4 IHS, 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 Sierra Leone poverty 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 poverty scorecard's estimate of the change in poverty rates over time:

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

z , c , and p are defined as 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 poverty scorecard and the theoretical confidence interval under direct measurement.

¹⁴ Poverty scoring does not reveal the reasons for this change.

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 poverty scorecard, where \hat{p} is based on previous measurements and is assumed equal at both baseline and follow-up:

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

For countries for which this α has been measured (Schreiner, 2010, 2009a, 2009b, 2009c, 2009d, 2009e, and 2008b; Schreiner and Woller, 2010a and 2010b; and Chen and Schreiner, 2009a and 2009b), 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 Sierra Leone.

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 $\hat{p} = 0.619$ (from Figure 2). Then the baseline sample size is $n = 2 \cdot \left(\frac{1.19 \cdot 1.64}{0.02} \right)^2 \cdot 0.619 \cdot (1 - 0.619) = 4,492$, and the follow-up sample size is also 4,492.

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 poverty 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 Sierra Leone scorecard is applied twice (once after April 2004 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 61.9 percent ($p_{2003/4} = 0.619$, Figure 2), and suppose $\alpha = 1.30$. Then the baseline sample size is

$$n = 2 \cdot \left(\frac{1.30 \cdot 1.64}{0.02} \right)^2 \cdot \{-0.02 + 0.016 \cdot 3 + 0.47 \cdot [0.619 \cdot (1 - 0.619)]\} = 3,156. \text{ The same}$$

group of 3,156 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* (expenditure below a poverty line). Poverty status is a fact that depends on whether expenditure is below a poverty line as directly measured by a survey. In contrast, targeting status is a program’s policy choice that depends on a cut-off and on an indirect estimate from a scorecard.

Targeting is successful when households truly below a poverty line are targeted (*inclusion*) and when households truly above a poverty line are not targeted (*exclusion*). Of course, no scorecard is perfect, and targeting is unsuccessful when households truly below a poverty line are not targeted (*undercoverage*) or when households truly above a poverty line are targeted (*leakage*). Figure 10 depicts these four possible targeting outcomes. Targeting accuracy varies by 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 50–54, outcomes for the national line in the validation sample are:

- Inclusion: 51.8 percent are below the line and correctly targeted
- Undercoverage: 9.8 percent are below the line and mistakenly not targeted
- Leakage: 15.5 percent are above the line and mistakenly targeted
- Exclusion: 22.8 percent are above the line and correctly not targeted

Increasing the cut-off to 55–59 improves inclusion and undercoverage but worsens leakage and exclusion:

- Inclusion: 56.1 percent are below the line and correctly targeted
- Undercoverage: 5.5 percent are below the line and mistakenly not targeted
- Leakage: 19.9 percent are above the line and mistakenly targeted
- Exclusion: 18.5 percent are above the line and correctly not targeted

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

$$\begin{array}{rcl}
 (\text{Benefit per household correctly included} & \times & \text{Households correctly included}) & - \\
 (\text{Cost per household mistakenly not covered} & \times & \text{Households mistakenly not covered}) & - \\
 (\text{Cost per household mistakenly leaked} & \times & \text{Households mistakenly leaked}) & + \\
 (\text{Benefit per household correctly excluded} & \times & \text{Households correctly excluded}). &
 \end{array}$$

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 and

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:

$$\begin{array}{rclcl}
 \text{Total Accuracy} = & 1 & \times & \text{Households correctly included} & - \\
 & 0 & \times & \text{Households mistakenly undercovered} & - \\
 & 0 & \times & \text{Households mistakenly leaked} & + \\
 & 1 & \times & \text{Households correctly excluded.} &
 \end{array}$$

Figure 11 shows “Total Accuracy” for all cut-offs for the Sierra Leone scorecard. For the national line in the validation sample, total net benefit is greatest (74.6) for a cut-off of 54 or less or of 59 or less, with about three in four households in Sierra Leone 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 $(2 \times \text{Households correctly included}) + (1 \times \text{Households correctly excluded})$.¹⁷

As an alternative to assigning benefits and costs to targeting outcomes and then choosing a cut-off to maximize total net benefits, a program could set a cut-off to

¹⁷ Figure 11 also reports “BPAC”, the Balanced Poverty Accuracy Criteria adopted by USAID for certifying poverty scorecards. IRIS Center (2005) says that BPAC considers accuracy in terms of estimated poverty rates and in terms of targeting inclusion. $\text{BPAC} = (\text{Inclusion} - |\text{Undercoverage} - \text{Leakage}|) \times [100 \div (\text{Inclusion} + \text{Undercoverage})]$.

achieve a desired poverty rate among targeted households. The third column of Figure 12 (“% targeted who are poor”) shows, for the Sierra Leone 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 54 or less would target 67.3 percent of all households (second column) and produce a poverty rate among those targeted of 76.9 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 54 or less, 84.0 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 54 or less, covering 3.3 poor households means leaking to 1 non-poor household.

9. Conclusion

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

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

The scorecard is built with half of the data from Sierra Leone's 2003/4 IHS, tested on the other half of the 2003/4 IHS, 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 changes 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 average difference between estimates and true poverty rates for groups of households at a point in time across all seven poverty lines is +1.1 percentage points. For $n = 16,384$ and 90-percent confidence, the precision of these differences is ± 0.8 percentage points or better.

If a program wants to use the scorecard for targeting, then the results here provide the information needed to select a cut-off that fits its 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 poverty scorecard is kept simple, using ten indicators that are inexpensive to collect and that are straightforward to verify. Points are all zeros or positive integers, and scores range from 0 to 100. Scores are related to poverty likelihoods via simple look-up tables, and targeting cut-offs are likewise simple to apply. The design attempts to facilitate adoption by helping managers understand and trust scoring and by allowing non-specialists to generate scores quickly in the field.

In summary, the poverty scorecard is a practical, objective way for pro-poor programs in Sierra Leone to estimate poverty rates, track changes in poverty rates over time, and target services. The same approach can be applied in any country with similar data from a national income or expenditure survey.

References

- Adams, Niall M.; and David J. Hand. (2000) “Improving the Practice of Classifier Performance Assessment”, *Neural Computation*, Vol. 12, pp. 305–311.
- Baesens, Bart; Van Gestel, Tony; Viaene, Stijn; Stepanova, Maria; Suykens, Johan A. K.; and Jan Vanthienen. (2003) “Benchmarking State-of-the-Art Classification Algorithms for Credit Scoring”, *Journal of the Operational Research Society*, Vol. 54, pp. 627–635.
- Caire, Dean. (2004) “Building Credit Scorecards for Small-Business Lending in Developing Markets”,
http://www.microfinance.com/English/Papers/Scoring_SMEs_Hybrid.pdf,
retrieved 18 March 2011.
- Chen, Shiyuan; and Mark Schreiner. (2009a) “A Simple Poverty Scorecard for Vietnam”,
http://www.microfinance.com/English/Papers/Scoring_Poverty_Vietnam_EN_2006.pdf, retrieved 18 March 2011.
- (2009b) “A Simple Poverty Scorecard for Bangladesh”,
http://www.microfinance.com/English/Papers/Scoring_Poverty_Bangladesh_2005_EN.pdf, retrieved 18 March 2011.
- Coady, David; Grosh, Margaret; and John Hoddinott. (2002) “The Targeting of Transfers in Developing Countries: Review of Experience and Lessons”,
http://info.worldbank.org/etools/docs/library/79646/Dc%202003/course_s/dc2003/readings/targeting.pdf, retrieved 18 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) “DefProbTM: A Corporate Probability-of-Default Model”,
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1103404, retrieved 18 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”, http://www.microfinance.com/English/Papers/Scoring_Poverty_Haiti_Fuller.pdf, retrieved 18 March 2011.
- Goodman, Leo A.; and Kruskal, William H. (1979) *Measures of Association for Cross Classification*, New York, NY: Springer-Verlag.
- Greenwell, Geoffrey. (2005) “Consumption Aggregates and Poverty Estimates for the Sierra Leone Integrated Household Survey”, Freetown: Statistics Sierra Leone.
- 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, Washington, D.C., <http://go.worldbank.org/VPMWVLU8E0>, retrieved 18 March 2011.
- Grosh, Margaret; and Judy L. Baker. (1995) “Proxy Means Tests for Targeting Social Programs: Simulations and Speculation”, LSMS Working Paper No. 118, Washington, D.C.: World Bank, <http://go.worldbank.org/W90WN57PD0>, retrieved 18 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.
- IRIS Center. (2007a) “Manual for the Implementation of USAID Poverty Assessment Tools”, http://pdf.usaid.gov/pdf_docs/PNADQ620.pdf, retrieved 18 March 2011.
- (2007b) “Introduction to Sampling for the Implementation of PATs”, http://www.povertytools.org/training_documents/Sampling/Introduction_Sampling.ppt, retrieved 8 March 2011.
- (2005) “Notes on Assessment and Improvement of Tool Accuracy”, http://www.povertytools.org/other_documents/AssessingImproving_Accuracy.pdf, retrieved 18 March 2011.

- Johnson, Glenn. (2007) “Lesson 3: Two-Way Tables—Dependent Samples”, http://www.stat.psu.edu/online/courses/stat504/03_2way/53_2way_compare.htm, accessed 8 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.
- Lovie, Alexander D.; and Patricia Lovie. (1986) “The Flat-Maximum Effect and Linear Scoring Models for Prediction”, *Journal of Forecasting*, Vol. 5, pp. 159–168.
- Martinelli, César; and Susan W. Parker. (2007) “Deception and Misreporting in a Social Program”, Centro de Investigación Económica and Instituto Tecnológico Autónomo de México, <http://ciep.itam.mx/~martinel/lies4.pdf>, retrieved 18 March 2011.
- Matul, Michal; and Sean Kline. (2003) “Scoring Change: Prizma’s Approach to Assessing Poverty”, Spotlight Note No. 4, Warsaw, Poland: Microfinance Centre for Central and Eastern Europe and the New Independent States, http://www.mfc.org.pl/doc/Research/ImpAct/SN/MFC_SN04_eng.pdf, retrieved 18 March 2011.
- McNemar, Quinn. (1947) “Note on the Sampling Error of the Difference between Correlated Proportions or Percentages”, *Psychometrika*, Vol. 17, pp. 153–157.
- Microfinance Risk Management, L.L.C. (2011) “Data-Entry Software for a Simple Poverty Scorecard for Sierra Leone”, http://www.microfinance.com/#Sierra_Leone, retrieved 29 March 2011.
- Myers, James H.; and Edward W. Forgy. (1963) “The Development of Numerical Credit-Evaluation Systems”, *Journal of the American Statistical Association*, Vol. 58, No. 303, pp. 779–806.
- Narayan, Ambar; and Nobuo Yoshida. (2005) “Proxy Means Tests for Targeting Welfare Benefits in Sri Lanka”, Report No. SASPR–7, Washington, D.C.: World Bank, http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2005/08/09/000090341_20050809094744/Rendered/PDF/332580PAPEROSASPR17.pdf, retrieved 18 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.

- SAS Institute Inc. (2004) “The LOGISTIC Procedure: Rank Correlation of Observed Responses and Predicted Probabilities”, in *SAS/STAT User’s Guide, Version 9*, Cary, NC.,
http://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/statug_logistic_sect035.htm, retrieved 18 March 2011.
- Schreiner, Mark. (2010) “A Simple Poverty Scorecard for Honduras”,
http://www.microfinance.com/English/Papers/Scoring_Poverty_Honduras_EN_2007.pdf, retrieved 18 March 2011.
- (2009a) “A Simple Poverty Scorecard for Peru (revised)”,
http://www.microfinance.com/English/Papers/Scoring_Poverty_Peru.pdf,
retrieved 18 March 2011.
- (2009b) “A Simple Poverty Scorecard for the Philippines”,
http://www.microfinance.com/English/Papers/Scoring_Poverty_Philippines.pdf,
retrieved 18 March 2011.
- (2009c) “A Simple Poverty Scorecard for Pakistan”,
http://www.microfinance.com/English/Papers/Scoring_Poverty_Pakistan_2005.pdf,
retrieved 18 March 2011.
- (2009d) “A Simple Poverty Scorecard for Bolivia”,
http://www.microfinance.com/English/Papers/Scoring_Poverty_Bolivia_EN_2007.pdf,
retrieved 18 March 2011.
- (2009e) “A Simple Poverty Scorecard for Mexico”,
http://www.microfinance.com/English/Papers/Scoring_Poverty_Mexico_2008_EN.pdf,
retrieved 18 March 2011.
- (2008a) “A Simple Poverty Scorecard for Peru”,
http://www.microfinance.com/English/Papers/Scoring_Poverty_Peru_May_2008.pdf,
retrieved 18 March 2011.
- (2008b) “A Simple Poverty Scorecard for India”,
http://www.microfinance.com/English/Papers/Scoring_Poverty_India.pdf,
retrieved 18 March 2011.
- (2006) “Is One Simple Poverty Scorecard Enough for India?”, memo for Grameen Foundation U.S.A.,
http://www.microfinance.com/English/Papers/Scoring_Poverty_India_Segments.pdf,
retrieved 18 March 2011.

- (2005a) “Un Índice de Pobreza para México”, report to Grameen Foundation, http://www.microfinance.com/Castellano/Documentos/Scoring_Pobreza_Mexico_2002.pdf, retrieved 18 March 2011.
- (2005b) “IRIS Questions on Poverty Scorecards”, memo for Grameen Foundation, http://www.microfinance.com/English/Papers/Scoring_Poverty_Response_to_IRIS.pdf, retrieved 18 March 2011.
- (2002) *Scoring: The Next Breakthrough in Microfinance?* Occasional Paper No. 7, Consultative Group to Assist the Poor, Washington, D.C., http://collab2.cgap.org//gm/document-1.9.29797/3276_076.pdf, retrieved 18 March 2011.
- ; Matul, Michal; Pawlak, Ewa; and Sean Kline. (2004) “Poverty Scorecards: Lessons from a Microlender in Bosnia-Herzegovina”, http://www.microfinance.com/English/Papers/Scoring_Poverty_in_BiH_Short.pdf, retrieved 18 March 2011.
- ; and Gary Woller. (2010a) “A Simple Poverty Scorecard for Ghana”, http://www.microfinance.com/English/Papers/Scoring_Poverty_Ghana_EN_2005.pdf, retrieved 18 March 2011.
- ; and Gary Woller. (2010b) “A Simple Poverty Scorecard for Guatemala”, http://www.microfinance.com/English/Papers/Scoring_Poverty_Guatemala_EN_2006.pdf, retrieved 18 March 2011.
- Sillers, Don. (2006) “National and International Poverty Lines: An Overview”, Washington, D.C.: U.S. Agency for International Development, http://www.microlinks.org/ev02.php?ID=12353_201&ID2=DO_TOPIC, retrieved 18 March 2011.
- Statistics Sierra Leone. (2007) *Final Statistical Report, Sierra Leone Integrated Household Survey 2003/4*, Freetown, http://www.statistics.sl/slihs_report.pdf, retrieved 8 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, pp. 773–792.
- Toohig, Jeff. (2008) “PPI Pilot Training Participant Guide”, Grameen Foundation, <http://www.progressoutofpoverty.org/toolkit>, retrieved 18 March 2011.
- United States Congress. (2004) “Microenterprise Results and Accountability Act of 2004 (HR 3818 RDS)”, November 20, http://www.povertytools.org/other_documents/HR%203818%20RDS.pdf, retrieved 18 March 2011.
- Wainer, Howard. (1976) “Estimating Coefficients in Linear Models: It Don’t Make No Nevermind”, *Psychological Bulletin*, Vol. 83, pp. 223–227.
- World Bank. (2009) “Sierra Leone: Joint IDA-IMF Staff Advisory Note on the Second Poverty Reduction Strategy Paper”, Washington, D.C., <http://go.worldbank.org/NVRFCZOGF0>, retrieved 18 March 2011.
- (2008) “International Comparison Project: Tables of Results”, Washington, D.C., <http://siteresources.worldbank.org/ICPINT/Resources/icp-final-tables.pdf>, retrieved 18 March 2011.
- Zeller, Manfred. (2004) “Review of Poverty Assessment Tools”, http://pdf.usaid.gov/pdf_docs/PNADH120.pdf, retrieved 18 March 2011.

Appendix A: Guide to Interpretation of Scorecard Indicators

The following comes from:

Statistics Sierra Leone. (2002) *Supervisor's and Enumerator's Manual, 2002/3 Sierra Leone Integrated Household Survey*, Freetown.

1. How many members does the household have?

According to p. 9 of the *Manual*, “*Household members* are defined to include all people who normally live and eat their meals together in the house and from the same pot, and have done so for nine or more months preceding the survey. Please note: the interviewer should be extremely careful to determine accurately the number of months absent from the household because that is the main criteria for determining who is a household member.

“Specific cases of members or non-members are:

Members:

- The household head is always a member of the household, even if that person does not eat and sleep in the household more than nine months
- Relatives of the household head who have been absent three months or less in the past 12 months and who are not permanent residents of other households
- Infants less than 3-months-old
- New permanent residents of the household who have been there less than six months because they were newly demobilized, married, or had a job transfer
- Students living outside the household but still only supported by their family
- Relatives of the household head whose work requires them to be outside the household for more than nine months of the year but who are considered in the household budget
- Guests living with the households for nine months or more
- Children who go to study and live away from the household, but for whom the household still provides living expenses and who are not members of any other household

Non-members:

- Individuals who died during the past 12 months, even if they lived in the household for more than nine months
- People who have lived in the household for nine months but left due to marriage etc.
- Lodgers, hired workers, and servants who are heads of their own households in another location, even though they sleep under the same roof and contribute to eat together with the household for nine months or more
- Guests and lodgers and all other people not listed in the definition of household members
- People who have joined the army but plan to return to the household”

According to p. 10 of the *Manual*, “For the purposes of this survey, a *household* will be defined as a person, or group of people, who usually slept in the same dwelling and took their meals together or ate from the same pot for at least nine months preceding the interview.

“The following are the different types of households:

- A household consisting of a single person
- A household consisting of a couple or several couples with or without their children
- All listed persons who have been away from the household for more than three months are not considered to be household members except:
 - The person identified as the head of the household, even if he has not been with the household for nine months or more
 - Newly born children not yet 6-months-old
 - Students and seasonal workers who have been living in or as part of another household
- Two-generation families, including parents and children
- Multi-generation families made up of the household head, wife and children of the household head, father of household head, grandchildren, and other people with or without a blood relationship, who sleep in the same place and eat together during nine months out of the 12 months before the interview”

2. Are all household members ages 6 to 13 in school now?

According to p. 14 of the *Manual*, “*formal schooling* refers to attendance of a primary, middle/JSS or SSS, vocational/technical or professional school, or training or an apprenticeship course. Attendance to a Koranic school is included only if the respondent attended no other school.”

3. What was the activity of the female head/spouse in her main occupation in the past 12 months?

According to p. 23 of the *Manual*, “*activities outside of agriculture (cultivation, animal husbandry), forestry, aquaculture* include production (industry, handicraft, and fishing), business (trade in various services and goods), services (doctor, lawyer, other services . . .).

“*In the past 12 months* is calculated from the time of the interview back to the past. For example, if the interview takes place on January 1, 2000, then the first day of the last 12 months is January 1, 1999.

“*Main employment or occupation* is the work for which the person spends most of his/her time in comparison with other work in a certain time period. Attending school is not considered main work. The main occupation of a respondent who farms mostly but often goes fishing during the raining season is farming.”

The *female head/spouse* is:

- The head of the household, if the head is female
- The spouse of the head of the household, if the head is male and has a spouse
- Non-existent, if the head is male and does not have a spouse

According to p. 11 of the *Manual*, “The *head of household* is usually the person (man or woman) who provides most of the needs of the household and is familiar with all the activities and occupations of the household members. He will be the person named when you ask the question “Who is the head of the household?” Usually (but not necessarily), the household head is the person with the highest income in the household, who holds a decision-making role in household work, and who at the same time should know the most information about economic activities and occupations of other household members in detail.”

According to pp. 2 and 4 of the *Manual*, “The interviewer should try to find the appropriate response code for the actual response of the person being interviewed. If there is no appropriate code, the interviewer can use the code ‘Other’. . . . If the response is not in the list, the interviewer must record the information in the “Other” category.”

4. How many rooms does the household occupy (exclude bathrooms, toilets, kitchen, pantry, hall, and storage)?

According to p. 11 of the *Manual*, the *dwelling* “is the structure or group of structures (rooms or buildings), separate or contiguous, occupied by members of the household. This can be:

- A single-family house/hut
- A flat/apartment (self-contained)
- Rooms (compound house)
- Several huts/buildings (same compound)
- Several huts/buildings (different compounds)”

According to p. 29 of the *Manual*, the *dwelling* “includes all types of structures occupied by members of a household. These may consist of a room inside a house, a group of houses, a multi-storied house, or a hut or group of huts.”

5. What is the main flooring material?

According to pp. 2 and 4 of the *Manual*, “The interviewer should try to find the appropriate response code for the actual response of the person being interviewed. If there is no appropriate code, the interviewer can use the code ‘Other’. . . . If the response is not in the list, the interviewer must record the information in the “Other” category.”

6. What is the main construction material of the outside walls?

According to p. 30 of the *Manual*, “If the exterior walls of the dwelling are composed of several materials, for instance, one part of the wall is of bamboo, another part of earth and yet another part of concrete, choose the predominant material.”

According to pp. 2 and 4 of the *Manual*, “The interviewer should try to find the appropriate response code for the actual response of the person being interviewed. If there is no appropriate code, the interviewer can use the code ‘Other’. . . . If the response is not in the list, the interviewer must record the information in the “Other” category.”

7. What type of toilet is used by the household?

According to pp. 2 and 4 of the *Manual*, “The interviewer should try to find the appropriate response code for the actual response of the person being interviewed. If there is no appropriate code, the interviewer can use the code ‘Other’. . . . If the response is not in the list, the interviewer must record the information in the “Other” category.”

8. What is the main source of lighting for your dwelling?

According to pp. 2 and 4 of the *Manual*, “The interviewer should try to find the appropriate response code for the actual response of the person being interviewed. If there is no appropriate code, the interviewer can use the code ‘Other’. . . . If the response is not in the list, the interviewer must record the information in the “Other” category.”

9. What is the main fuel used by the household for cooking?

According to pp. 2 and 4 of the *Manual*, “The interviewer should try to find the appropriate response code for the actual response of the person being interviewed. If there is no appropriate code, the interviewer can use the code ‘Other’. . . . If the response is not in the list, the interviewer must record the information in the “Other” category.”

10. How many radios, radio cassettes, record players, or 3-in-1 radio cassettes do members of the household own?

The *Manual* provides no additional information about this indicator.

Figure 2: Sample sizes, poverty lines, and poverty rates for all of Sierra Leone and by region, for the construction/calibration sample, and for the validation sample, by poverty line, and by household-level/person-level

	Level	Sample size	Poverty rates (% with expenditure below a poverty line) and poverty lines (SLL per adult equivalent or per person per day)						
			Food	National lines			USAID	Intl. 2005 PPP	
				75%	100%	150%	'Extreme'	\$1.25/day	\$2.50/day
Poverty lines:									
All Sierra Leone	N/A	3,702	1,133	1,772	2,363	3,545	1,084	1,374	2,749
Southern (Bo)	N/A	995	1,014	1,586	2,115	3,173	987	1,230	2,460
Eastern (Kenema)	N/A	981	1,097	1,715	2,286	3,430	860	1,330	2,660
Northern (Makeni)	N/A	1,261	1,022	1,598	2,130	3,196	934	1,239	2,478
Western (Freetown)	N/A	465	1,512	2,365	3,153	4,730	1,721	1,834	3,668
Poverty Rates:									
All Sierra Leone	Households	3,702	19.1	44.6	61.9	82.3	28.7	46.3	82.7
	People	N/A	22.4	49.7	66.3	84.9	33.1	51.7	86.5
Southern (Bo)	Households	995	13.1	37.2	57.1	79.4	26.4	38.4	81.1
	People	N/A	15.6	42.0	61.4	81.3	30.5	43.6	82.8
Eastern (Kenema)	Households	981	32.5	64.9	81.8	96.0	36.4	68.2	96.0
	People	N/A	37.5	69.3	84.3	95.8	42.3	72.8	96.0
Northern (Makeni)	Households	1,261	23.2	54.1	73.4	92.5	35.1	56.7	92.9
	People	N/A	26.8	60.5	78.9	94.5	39.5	63.4	95.0
Western (Freetown)	Households	465	3.7	14.6	26.2	53.7	12.3	13.5	53.3
	People	N/A	4.9	16.6	28.6	59.0	14.1	15.5	59.4
Construction and calibration sample									
Selecting indicators and points, and associating scores with likelihoods	Households	1,832	19.0	44.8	62.1	82.3	28.5	46.3	82.9
	People	N/A	22.6	50.3	67.3	85.3	33.2	51.9	86.1
Validation sample									
Measuring accuracy	Households	1,870	19.1	44.4	61.6	82.3	28.9	46.3	82.6
	People	N/A	22.2	49.1	65.3	84.5	33.1	51.4	84.9
Change in poverty rate (percentage points)									
Construction/calibration to validation	Households		-0.1	+0.4	+0.5	+0.0	-0.4	-0.0	+0.2

Source: 2003/4 IHS. National lines are per adult equivalent per day, all other lines are per person per day.

Figure 3: Poverty indicators by uncertainty coefficient

<u>Uncertainty coefficient</u>	<u>Indicator (Answers ordered starting with those most strongly linked with higher poverty likelihoods)</u>
13,139	What is the main source of drinking water for your household? (River, lake, spring, pond, rainwater, or other; Well without pump; Public standpipe; Private outside standpipe, or well with pump; Water vendor, water tanker, or neighboring household; Indoor plumbing, or inside standpipe)
12,265	Was any member of the household allocated any land (including land outside this area) or own any land in the past 12 months, and if so, did any member of the household own any work oxen, cattle, sheep, goats, pigs, rabbits, chickens, other poultry, or other livestock in the past 12 months? (Had some land and some animals; Did not have any land, but did have some animals; Had some land, but did not have any animals; Did not have any land nor animals)
12,260	How many radios, radio cassettes, record players, or 3-in-1 radio cassettes do members of the household own? (None; One; Two or more)
11,895	What is the main construction material of the outside walls? (Stone/burnt bricks, or other; Mud/mud bricks, or wood; Cement/sandcrete, or corrugated iron sheets)
11,809	What type of worker is the female head/spouse? (Does not work; Agricultural non-paid family worker; Non-agricultural non-paid family worker, or agricultural self-employed (with or without employees); No female head/spouse; Non-agricultural self-employed with employees, or non-agricultural self-employed without employees; Paid employee)
11,742	In the past seven days, what was the main occupation of the female head/spouse? (None; Agriculture, forestry, and fisheries; Other professionals; No female head/spouse; Armed forces, science and technology professional, health professional, education professional, personal services, protection, and sales, skilled manual labor, assemblers and machine operators, or unskilled workers)
11,741	What was the activity of the female head/spouse in her main occupation in the past 12 months? (No female head/spouse; Agriculture, forestry, mining, and quarrying; Any other, or does not work)
11,571	How many cutlasses are owned now by members of the household? (One or more; None)
11,486	Was any member of the household allocated any land (including land outside this area) or own any land in the past 12 months? (Yes; No)

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

<u>Uncertainty coefficient</u>	<u>Indicator (Answers ordered starting with those most strongly linked with higher poverty likelihoods)</u>
10,685	How many radio cassettes, record players, or 3-in-1 radio cassettes do members of the household own? (None; One; Two or more)
10,558	How many hand hoes are owned now by members of the household? (One or more; None)
9,948	What type of worker is the male head/spouse? (Agricultural non-paid family worker; Agricultural self-employed (with or without employees); No male head/spouse; Does not work, or non-agricultural non-paid family worker; Paid employee; Non-agricultural self-employed with employees, or non-agricultural self-employed without employees)
9,838	What is the main source of lighting for your dwelling? (Generator, kerosene, gas lamp, candles/torch light, or other; Electricity (mains))
9,768	What was the activity of the male head/spouse in his main occupation in the past 12 months? (Agriculture, forestry, mining and quarrying; No male head/spouse; Manufacturing/processing, transport, storage, and communication, or finance, insurance, real estate, and business services; Did not work; Wholesale/retail trade; Community/social and personal; Electricity, water, and gas, construction, major repair, or maintenance, restaurants, hotels, and food sellers, or community/social and personal services)
9,733	What is the main flooring material? (Earth/mud, stone/brick, or other; Wood, or cement/concrete)
9,684	In what industry was the main occupation of the male head/spouse in the past 12 months? (Agriculture, forestry, and fisheries; No male head/spouse; Did not work; Mining and extracting, manufacturing, electricity and water, commerce, transport and communication, finance, government administration and national defense, education and training, health and social services, health and social services (culture and sports, or personal services)
8,978	In the past seven days, what was the main occupation of the male head/spouse? (Agriculture, forestry, and fisheries; None; No male head/spouse; Personal services, protection, and sales, skilled manual labor, unskilled workers; Armed forces, science and technology professional, health professionals, education professionals, other professionals, or assemblers and machine operators)

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

<u>Uncertainty coefficient</u>	<u>Indicator (Answers ordered starting with those most strongly linked with higher poverty likelihoods)</u>
8,734	In the past seven days, what was the main occupation of the female head/spouse? (Agriculture, forestry, and fisheries; None; Other professionals; No female head/spouse; Armed forces, science and technology professional, health professional, education professional, personal services, protection, and sales, skilled manual labor, assemblers and machine operators, or unskilled workers)
8,327	What is the religious denomination of the female head/spouse? (Catholic, spiritualist, other Christian, Muslim, no religion, or other; Presbyterian; Anglican; No female head/spouse; Methodist, or Pentecostal)
6,576	Did any member of the household own any work oxen, cattle, sheep, goats, pigs, rabbits, chickens, other poultry, or other livestock in the past 12 months? (Yes; No)
6,564	What is the highest class or form completed by the male head/spouse? (Koranic, other, or none; Kindergarten, or Primary 1–4; No male head/spouse; Primary 5–6, or JSS 1-3; SSS 1-3; Voc/comm., tech/training (TC or HTC), technical, nursing, polytechnic, or university)
6,444	What is your present occupancy status? (Owning; House provided rent-free; Renting, or subsidized rent)
6,382	What type of toilet is used by the household? (Bush/river, none, or other; Bucket, common pit, or VIP; Private pit, common flush, or flush toilet)
6,299	How many household members did any work for pay, profit, family gain, or to produce anything for barter or home use during the last seven days? (Four or more; Three; Two; One; None)
6,021	How many household members are 18-years-old or younger? (Six or more; Five; Four; Three; Two; One; None)
5,964	How many household members are 17-years-old or younger? (Five or more; Four; Three; Two; One; None)
5,936	How many household members are 15-years-old or younger? (Five or more; Four; Three; Two; One; None)
5,879	How many household members are 16-years-old or younger? (Five or more; Four; Three; Two; One; None)
5,833	What is the main fuel used by the household for cooking? (Wood, or other; Charcoal; Gas, kerosene, or electricity)

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

<u>Uncertainty coefficient</u>	<u>Indicator (Answers ordered starting with those most strongly linked with higher poverty likelihoods)</u>
5,799	How many household members are 12-years-old or younger? (Five or more; Four; Three; Two; One; None)
5,721	What is the highest class or form completed by the female head/spouse? (Koranic, kindergarten, none, or other; Primary 1–6; JSS 1-3; No female head/spouse; SSS 1-3, voc./comm, tech/training (TC or HTC), technical, nursing, polytechnic, or university)
5,611	How many household members are 14-years-old or younger? (Five or more; Four; Three; Two; One; None)
5,441	How many radios do members of the household own? (None; One; Two or more)
5,421	Can the male head/spouse read a simple letter in English or in a language of Sierra Leone (state the one in which the male head/spouse is the most proficient)? (None; Mende; No male head/spouse; English; Temne, Krio, Limba, or other)
5,390	How many household members are 13-years-old or younger? (Five or more; Four; Three; Two; One; None)
5,366	What is the present marital status of the female head/spouse? (Married polygamously; Widowed; Married monogamously; Divorced/separated; No female head/spouse; Informal/loose union, or never-married)
5,322	How many household members are 11-years-old or younger? (Four or more; Three; Two; One; None)
5,296	How many members does the household have? (Ten or more; Seven, eight, or nine; Six; Five; Four; One, two, or three)
5,199	Can the female head/spouse do written calculations? (No; No female head/spouse; Yes)
5,069	Can the female head/spouse read a simple letter in English or in a language of Sierra Leone (state the one in which the female head/spouse is the most proficient)? (None; English, Mende, Temne, Krio, Limba, or other; No female head/spouse)
4,858	What is the religious denomination of the male head/spouse? (Catholic, spiritualist, other Christian, Muslim, no religion, or other; Presbyterian; Anglican; No male head/spouse; Pentecostal; Methodist)
4,245	Are all household members ages 6 to 18 in school now? (No; Yes, or no one aged 6 to 18)
4,189	Are all household members ages 6 to 17 in school now? (No; Yes, or no one aged 6 to 17)

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

<u>Uncertainty coefficient</u>	<u>Indicator (Answers ordered starting with those most strongly linked with higher poverty likelihoods)</u>
4,161	What is the present marital status of the male head/spouse? (Married polygamously; Married monogamously; No male head/spouse; Informal/loose union, divorced/separated, widowed, or never-married)
4,095	Can the male head/spouse do written calculations? (No; No male head/spouse; Yes)
4,087	Are all household members ages 6 to 16 in school now? (No; Yes, or no one aged 6 to 16)
4,077	Are all household members ages 6 to 15 in school now? (No; Yes, or no one aged 6 to 15)
4,060	What type of dwelling does the household live in? (Room(s) (Other type), or several huts/buildings (Different compound); Single-family house; Room(s) (Compound house); Apartment/flat, several huts/buildings (Same compound), or other)
4,020	Are all household members ages 6 to 13 in school now? (No; Yes, or no one aged 6 to 13)
3,997	To which ethnic group does the female head/spouse belong? (Kissi; Loko; Sherbro; Mende; Mandingo; Other; Limba; Temne; No female head/spouse; Vie; Krio)
3,881	How many pieces of land/plots do members of the household own? (One; Two or more; None)
3,845	Are all household members ages 6 to 12 in school now? (No; Yes, or no one aged 6 to 12)
3,769	Are all household members ages 6 to 14 in school now? (No; Yes, or no one aged 6 to 14)
3,520	How many household members read a simple letter in any language? (None; One; Two; Three; Four or more)
3,359	What is the main roofing material? (Thatch (grass, straw), or plastic sheeting; Wood, corrugated iron, cement/concrete, roofing tiles, or other)
3,322	How many houses do members of the household own? (One; None; Two or more)
3,032	Are all household members ages 6 to 11 in school now? (No; Yes, or no one aged 6 to 11)
2,816	How many household members are 6-years-old or younger? (Three or more; Two; One; None)
2,530	To which ethnic group does the male head/spouse belong? (Kissi, or Vie; Loko; Mende; Sherbro; Temne; Limba; Other; No male head/spouse; Mandingo; Krio)

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

<u>Uncertainty coefficient</u>	<u>Indicator (Answers ordered starting with those most strongly linked with higher poverty likelihoods)</u>
2,369	How many pieces of furniture do members of the household own? (None; One; Two or more)
2,290	How old is the female head/spouse? (35 to 39; 40 to 44; 50 to 54; 25 to 34; 45 to 49; 24 or younger; 55 or older; No female head/spouse)
1,902	Did the female head/spouse do any work for pay, profit, or family gain, or did she produce anything for barter or home use during the last seven days? (Yes; No; No female head/spouse)
1,713	What is the structure of household headship? (Both male and female heads/spouses; Female head/spouse only; Male head/spouse only)
1,590	How does your household get rid of rubbish? (Other; Dumped by household; Burned by household; Collected, or buried by household)
1,378	How many bicycles, motorcycles, or cars do members of the household own? (None; One or more)
1,371	How many rooms does the household occupy (exclude bathrooms, toilets, kitchen, pantry, hall, and storage), and do other households share the dwelling? (Shared, with three rooms; Shared, with four or rooms; Shared, with two rooms; Not shared, with three or more rooms; Shared, with one room; Not shared, with two rooms; Not shared, with one room)
1,148	Do any household members attend a school owned by a mission/religious body, a non-government organization, a private entity, or by some other non-governmental entity? (Yes; No)
989	How old is the male head/spouse? (60 or older; 40 to 59; 39 or younger; No male head/spouse)
757	How many years has your household been living in this dwelling? (Two to seven; Less than two; Seven or more)
713	How many rooms does the household occupy (exclude bathrooms, toilets, kitchen, pantry, hall, and storage)? (One; Two; Three or more)
622	Did any member of the household own any work oxen, cattle, sheep, goats, pigs, or other livestock in the past 12 months? (Yes; No)
366	Do other households share this dwelling with you? (Yes; No)

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

<u>Uncertainty coefficient</u>	<u>Indicator (Answers ordered starting with those most strongly linked with higher poverty likelihoods)</u>
300	Did the male head/spouse do any work for pay, profit, or family gain, or did he produce anything for barter or home use during the last seven days? (No; Yes; No male head/spouse)
203	Does any member of the household own furniture, radio, radio cassette, record player, or 3-in-1 radio cassette, bicycle, motorcycle, car, house, or land/plot? (No; Yes)
40	What is the area of the dwelling in square feet? (Less than 500; 500 or more)

Source: 2003/4 IHS and the national poverty line

Tables for the National Poverty Line
(and Tables Pertaining to All Six 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	100.0
5-9	100.0
10-14	100.0
15-19	97.3
20-24	97.6
25-29	98.6
30-34	90.5
35-39	85.2
40-44	81.3
45-49	74.1
50-54	57.3
55-59	45.0
60-64	28.2
65-69	21.3
70-74	15.4
75-79	10.5
80-84	4.1
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	0	÷	0	=	100.0
5-9	0	÷	0	=	100.0
10-14	0	÷	0	=	100.0
15-19	622	÷	639	=	97.3
20-24	2,262	÷	2,318	=	97.6
25-29	3,644	÷	3,697	=	98.6
30-34	7,667	÷	8,471	=	90.5
35-39	10,701	÷	12,560	=	85.2
40-44	12,361	÷	15,213	=	81.3
45-49	10,053	÷	13,567	=	74.1
50-54	6,216	÷	10,858	=	57.3
55-59	3,893	÷	8,659	=	45.0
60-64	1,523	÷	5,404	=	28.2
65-69	1,509	÷	7,086	=	21.3
70-74	805	÷	5,237	=	15.4
75-79	344	÷	3,270	=	10.5
80-84	54	÷	1,308	=	4.1
85-89	0	÷	736	=	0.0
90-94	0	÷	977	=	0.0
95-100	0	÷	0	=	0.0

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

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

Score	Likelihood of having expenditure in range demarcated by poverty lines per day per adult equivalent				
	Food	=>Food	=>75% Natl.	=>100% Natl.	=>150% Natl.
		and	and	and	
	<75% Natl.	<100% Natl.	<150% Natl.		
	=>SLL1,133	=>SLL1,772	=>SLL2,363		
<SLL1,133	and	and	and	=>SLL3,545	
	<SLL1,772	<SLL2,363	<SLL3,545		
0-4	100.0	0.0	0.0	0.0	0.0
5-9	100.0	0.0	0.0	0.0	0.0
10-14	100.0	0.0	0.0	0.0	0.0
15-19	47.3	41.2	8.8	0.0	2.7
20-24	45.2	37.7	14.7	1.8	0.6
25-29	48.0	42.3	8.3	0.7	0.7
30-34	33.1	36.1	21.3	8.4	1.1
35-39	35.1	35.1	15.0	8.6	6.2
40-44	25.3	32.9	23.1	14.8	4.0
45-49	16.9	30.7	26.5	22.3	3.6
50-54	12.8	23.7	20.8	31.7	11.0
55-59	5.8	22.5	16.7	26.6	28.5
60-64	1.4	12.4	14.4	42.5	29.3
65-69	1.5	8.4	11.4	30.3	48.5
70-74	0.8	2.3	12.3	35.2	49.4
75-79	0.0	8.9	1.6	19.9	69.6
80-84	0.0	1.9	2.3	32.0	63.9
85-89	0.0	0.0	0.0	10.4	89.6
90-94	0.0	0.0	0.0	18.3	81.7
95-100	0.0	0.0	0.0	0.0	100.0

All poverty likelihoods in percentage units.

Figure 6b: Distribution of household poverty likelihoods across expenditure ranges demarcated by per-person poverty lines

Score	Likelihood of having expenditure in range demarcated by poverty lines per day per person			
		=>USAID	=>\$1.25/day	
	<USAID	and	and	=>\$2.50/day
		<\$1.25/day	<\$2.50/day	
	=>SLL1,084	=>SLL1,374		
	<SLL1,084	and	and	=>SLL2,749
		<SLL1,374	<SLL2,749	
0-4	100.0	0.0	0.0	0.0
5-9	100.0	0.0	0.0	0.0
10-14	100.0	0.0	0.0	0.0
15-19	80.5	8.0	8.8	2.7
20-24	61.6	24.2	13.6	0.6
25-29	62.3	28.9	8.1	0.7
30-34	45.1	29.3	24.5	1.1
35-39	48.4	20.6	25.3	5.7
40-44	36.8	25.6	33.8	3.8
45-49	27.2	24.3	45.4	3.1
50-54	17.7	19.3	51.3	11.8
55-59	14.9	12.6	50.2	22.3
60-64	7.5	6.3	67.0	19.3
65-69	5.6	4.0	48.0	42.4
70-74	2.6	0.6	40.3	56.6
75-79	8.9	0.0	20.3	70.8
80-84	1.9	0.0	17.9	80.3
85-89	0.0	0.0	10.4	89.6
90-94	0.0	0.0	16.1	83.9
95-100	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

Score	Difference between estimate and true value			
	Diff.	Confidence interval (+/- percentage points)		
		90-percent	95-percent	99-percent
0-4	+0.0	0.0	0.0	0.0
5-9	+0.0	0.0	0.0	0.0
10-14	+0.0	0.0	0.0	0.0
15-19	+2.5	2.8	3.4	4.1
20-24	+11.9	3.0	3.5	4.6
25-29	+1.2	1.2	1.4	1.9
30-34	-0.6	1.3	1.5	2.1
35-39	+2.4	1.5	1.8	2.2
40-44	+12.6	1.8	2.2	3.1
45-49	+2.8	1.8	2.2	3.0
50-54	-8.6	5.4	5.6	5.9
55-59	-3.4	2.8	3.1	3.7
60-64	+2.5	2.8	3.2	4.2
65-69	+3.8	1.8	2.1	2.8
70-74	-5.1	4.0	4.2	4.7
75-79	+9.1	0.5	0.6	0.8
80-84	+2.9	0.6	0.7	0.9
85-89	-7.0	5.3	5.8	6.5
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 n	Difference between estimate and true value			
	Diff.	Confidence interval (+/- percentage points)		
		90-percent	95-percent	99-percent
1	+1.6	76.5	82.0	87.6
4	+1.8	37.4	46.8	62.0
8	+1.5	25.6	30.8	44.1
16	+1.1	18.6	22.9	32.0
32	+1.3	13.5	16.5	20.4
64	+1.4	9.3	11.2	13.5
128	+1.5	6.5	7.3	9.8
256	+1.7	4.6	5.5	7.2
512	+1.7	3.3	3.9	5.0
1,024	+1.8	2.3	2.7	3.6
2,048	+1.8	1.7	2.0	2.6
4,096	+1.8	1.2	1.5	1.9
8,192	+1.8	0.8	1.0	1.4
16,384	+1.8	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				USAID	Intl. 2005 PPP		
	Food	75%	100%	150%	'Extreme'	\$1.25/day	\$2.50/day	
<u>Estimate minus true value</u>								
Scorecard applied to validation sample	+0.3	+1.4	+1.8	+1.5	+0.6	+1.2	+1.2	
<u>Precision of difference</u>								
Scorecard applied to validation sample	0.4	0.5	0.6	0.8	0.5	0.5	0.7	
<u>α factor</u>								
Scorecard applied to validation sample	0.73	0.76	0.94	1.54	0.79	0.73	1.54	
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 $n = 256, 512, 1,024, 2,048, 4,096, 8,192, \text{ and } 16,384$.								
The national lines are in per-adult-equivalent units. The USAID "extreme" line and the 2005 PPP lines are in per-person units.								

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

		<u>Targeting segment</u>	
		<u>Targeted</u>	<u>Non-targeted</u>
<u>True poverty status</u>	<u>Below poverty line</u>	<u>Inclusion</u> Under poverty line Correctly Targeted	<u>Undercoverage</u> Under poverty line Mistakenly Non-targeted
	<u>Above poverty line</u>	<u>Leakage</u> Above poverty line Mistakenly Targeted	<u>Exclusion</u> 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	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Total Accuracy</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	=> poverty line mistakenly targeted	=> poverty line non-targeted	Inclusion + Exclusion	See text
0–4	0.0	61.6	0.0	38.4	38.4	–100.0
5–9	0.0	61.6	0.0	38.4	38.4	–100.0
10–14	0.0	61.6	0.0	38.4	38.4	–100.0
15–19	0.6	61.1	0.1	38.3	38.9	–98.0
20–24	2.5	59.1	0.4	37.9	40.5	–91.1
25–29	6.1	55.5	0.5	37.8	44.0	–79.3
30–34	13.7	47.9	1.4	36.9	50.6	–53.2
35–39	24.0	37.6	3.7	34.7	58.7	–16.2
40–44	34.9	26.7	8.0	30.4	65.3	+26.2
45–49	44.5	17.1	11.9	26.4	71.0	+63.9
50–54	51.8	9.8	15.5	22.8	74.6	+74.8
55–59	56.1	5.5	19.9	18.5	74.6	+67.8
60–64	58.1	3.6	23.3	15.0	73.1	+62.1
65–69	60.2	1.4	28.2	10.1	70.3	+54.2
70–74	61.3	0.3	32.4	6.0	67.3	+47.5
75–79	61.5	0.1	35.5	2.9	64.4	+42.4
80–84	61.6	0.1	36.7	1.6	63.2	+40.4
85–89	61.6	0.0	37.4	1.0	62.6	+39.3
90–94	61.6	0.0	38.4	0.0	61.6	+37.7
95–100	61.6	0.0	38.4	0.0	61.6	+37.7

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 successfully 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	0.0	100.0	0.0	Only poor targeted
5-9	0.0	100.0	0.0	Only poor targeted
10-14	0.0	100.0	0.0	Only poor targeted
15-19	0.6	89.9	0.9	8.9:1
20-24	3.0	85.3	4.1	5.8:1
25-29	6.7	92.0	9.9	11.5:1
30-34	15.1	90.6	22.2	9.6:1
35-39	27.7	86.7	38.9	6.5:1
40-44	42.9	81.4	56.6	4.4:1
45-49	56.5	78.9	72.3	3.7:1
50-54	67.3	76.9	84.0	3.3:1
55-59	76.0	73.9	91.1	2.8:1
60-64	81.4	71.3	94.2	2.5:1
65-69	88.5	68.1	97.7	2.1:1
70-74	93.7	65.5	99.5	1.9:1
75-79	97.0	63.4	99.8	1.7:1
80-84	98.3	62.6	99.9	1.7:1
85-89	99.0	62.2	100.0	1.6:1
90-94	100.0	61.6	100.0	1.6:1
95-100	100.0	61.6	100.0	1.6:1

Tables for the Food Poverty Line

Figure 4 (Food line): Estimated poverty likelihoods associated with scores

If a household's score is then the likelihood (%) of being below the poverty line is:
0-4	100.0
5-9	100.0
10-14	100.0
15-19	47.3
20-24	45.2
25-29	48.0
30-34	33.1
35-39	35.1
40-44	25.3
45-49	16.9
50-54	12.8
55-59	5.8
60-64	1.4
65-69	1.5
70-74	0.8
75-79	0.0
80-84	0.0
85-89	0.0
90-94	0.0
95-100	0.0

Figure 7 (Food 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

Score	Difference between estimate and true value			
	Diff.	Confidence interval (+/- percentage points)		
		90-percent	95-percent	99-percent
0-4	+0.0	0.0	0.0	0.0
5-9	+0.0	0.0	0.0	0.0
10-14	+0.0	0.0	0.0	0.0
15-19	+5.1	8.8	10.2	13.2
20-24	-2.5	4.6	5.5	7.1
25-29	-3.9	3.8	4.2	5.7
30-34	-11.0	6.8	7.1	7.5
35-39	+0.4	2.0	2.4	3.1
40-44	+3.2	1.5	1.9	2.6
45-49	+1.5	1.3	1.5	2.1
50-54	+2.5	1.2	1.5	2.0
55-59	+2.6	0.7	0.8	1.0
60-64	-0.6	0.6	0.7	1.0
65-69	+1.1	0.2	0.3	0.3
70-74	+0.8	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 (Food line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, scorecard applied to the validation sample

Sample Size n	Difference between estimate and true value			
	Diff.	Confidence interval (+/- percentage points)		
		90-percent	95-percent	99-percent
1	-0.6	54.9	64.2	71.1
4	+0.5	29.5	35.6	45.5
8	+0.2	19.4	23.3	29.6
16	+0.1	12.8	15.5	21.4
32	+0.0	8.9	10.5	14.2
64	+0.1	6.2	7.4	9.0
128	+0.2	4.0	4.8	6.4
256	+0.2	3.0	3.4	4.4
512	+0.3	2.1	2.5	3.3
1,024	+0.3	1.4	1.7	2.4
2,048	+0.3	1.1	1.2	1.6
4,096	+0.3	0.7	0.9	1.1
8,192	+0.3	0.5	0.6	0.8
16,384	+0.3	0.4	0.4	0.6

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

Score	Inclusion:	Undercoverage:	Leakage:	Exclusion:	Total Accuracy	BPAC
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	=> poverty line mistakenly targeted	=> poverty line non-targeted	Inclusion + Exclusion	See text
0–4	0.0	19.1	0.0	80.9	80.9	–100.0
5–9	0.0	19.1	0.0	80.9	80.9	–100.0
10–14	0.0	19.1	0.0	80.9	80.9	–100.0
15–19	0.2	18.9	0.4	80.4	80.7	–95.5
20–24	1.3	17.8	1.6	79.2	80.5	–77.7
25–29	3.3	15.9	3.4	77.5	80.7	–48.2
30–34	6.7	12.5	8.4	72.4	79.1	+13.9
35–39	11.0	8.2	16.7	64.1	75.1	+12.6
40–44	14.5	4.6	28.4	52.5	67.0	–48.1
45–49	17.1	2.1	39.4	41.5	58.6	–105.7
50–54	18.4	0.7	48.9	32.0	50.4	–155.4
55–59	18.9	0.3	57.1	23.7	42.6	–198.4
60–64	19.1	0.1	62.3	18.5	37.6	–225.5
65–69	19.1	0.0	69.3	11.5	30.7	–262.1
70–74	19.1	0.0	74.6	6.3	25.4	–289.5
75–79	19.1	0.0	77.8	3.0	22.2	–306.5
80–84	19.1	0.0	79.1	1.7	20.9	–313.4
85–89	19.1	0.0	79.9	1.0	20.1	–317.2
90–94	19.1	0.0	80.9	0.0	19.1	–322.3
95–100	19.1	0.0	80.9	0.0	19.1	–322.3

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

Figure 12 (Food 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 successfully 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	0.0	100.0	0.0	Only poor targeted
5-9	0.0	100.0	0.0	Only poor targeted
10-14	0.0	100.0	0.0	Only poor targeted
15-19	0.6	35.3	1.2	0.5:1
20-24	3.0	44.3	6.8	0.8:1
25-29	6.7	48.9	17.0	1.0:1
30-34	15.1	44.2	34.9	0.8:1
35-39	27.7	39.6	57.2	0.7:1
40-44	42.9	33.9	75.9	0.5:1
45-49	56.5	30.3	89.2	0.4:1
50-54	67.3	27.4	96.2	0.4:1
55-59	76.0	24.8	98.5	0.3:1
60-64	81.4	23.4	99.6	0.3:1
65-69	88.5	21.6	100.0	0.3:1
70-74	93.7	20.4	100.0	0.3:1
75-79	97.0	19.7	100.0	0.2:1
80-84	98.3	19.5	100.0	0.2:1
85-89	99.0	19.3	100.0	0.2:1
90-94	100.0	19.1	100.0	0.2:1
95-100	100.0	19.1	100.0	0.2:1

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	100.0
5-9	100.0
10-14	100.0
15-19	88.5
20-24	82.9
25-29	90.3
30-34	69.2
35-39	70.3
40-44	58.2
45-49	47.6
50-54	36.5
55-59	28.2
60-64	13.8
65-69	9.9
70-74	3.1
75-79	8.9
80-84	1.9
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

Score	Difference between estimate and true value			
	Diff.	Confidence interval (+/- percentage points)		
		90-percent	95-percent	99-percent
0-4	+0.0	0.0	0.0	0.0
5-9	+0.0	0.0	0.0	0.0
10-14	+0.0	0.0	0.0	0.0
15-19	+25.7	8.4	10.3	13.0
20-24	+0.3	3.1	3.8	4.9
25-29	+3.0	2.3	2.7	3.5
30-34	-2.1	2.3	2.8	3.4
35-39	+2.8	1.9	2.1	3.0
40-44	+6.5	1.9	2.3	2.8
45-49	-4.5	3.3	3.4	3.9
50-54	-2.4	2.2	2.5	3.2
55-59	+4.8	2.0	2.3	3.1
60-64	-3.5	2.9	3.2	3.6
65-69	+0.7	1.3	1.5	2.1
70-74	+1.3	0.5	0.6	0.8
75-79	+8.5	0.2	0.3	0.3
80-84	+0.6	0.6	0.7	0.9
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 n	Difference between estimate and true value			
	Diff.	Confidence interval (+/- percentage points)		
		90-percent	95-percent	99-percent
1	+1.0	66.4	71.0	89.3
4	+1.2	37.9	43.9	57.4
8	+0.6	25.8	30.5	41.9
16	+0.6	17.0	21.3	29.8
32	+0.8	11.4	13.8	17.7
64	+0.7	8.5	10.0	13.4
128	+1.1	5.6	6.4	7.9
256	+1.2	4.0	4.8	6.0
512	+1.3	2.7	3.3	4.3
1,024	+1.3	1.9	2.2	3.0
2,048	+1.3	1.4	1.7	2.1
4,096	+1.4	0.9	1.2	1.6
8,192	+1.4	0.7	0.8	1.1
16,384	+1.4	0.5	0.6	0.7

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:	Undercoverage:	Leakage:	Exclusion:	Total Accuracy	BPAC
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	=> poverty line mistakenly targeted	=> poverty line correctly non-targeted	Inclusion + Exclusion	See text
0-4	0.0	44.4	0.0	55.6	55.6	-100.0
5-9	0.0	44.4	0.0	55.6	55.6	-100.0
10-14	0.0	44.4	0.0	55.6	55.6	-100.0
15-19	0.4	44.1	0.3	55.3	55.6	-97.7
20-24	2.2	42.2	0.7	54.8	57.0	-88.3
25-29	5.5	39.0	1.2	54.4	59.8	-72.7
30-34	11.6	32.8	3.5	52.0	63.6	-39.9
35-39	19.9	24.6	7.8	47.8	67.7	+7.0
40-44	28.2	16.3	14.7	40.8	69.0	+59.9
45-49	35.3	9.2	21.2	34.4	69.7	+52.4
50-54	39.8	4.7	27.5	28.0	67.8	+38.1
55-59	41.9	2.5	34.1	21.5	63.4	+23.3
60-64	43.0	1.4	38.4	17.2	60.2	+13.7
65-69	44.1	0.4	44.4	11.2	55.3	+0.1
70-74	44.3	0.1	49.4	6.2	50.5	-11.1
75-79	44.4	0.1	52.6	3.0	47.3	-18.3
80-84	44.4	0.0	53.8	1.7	46.2	-21.1
85-89	44.4	0.0	54.6	1.0	45.4	-22.8
90-94	44.4	0.0	55.6	0.0	44.4	-25.0
95-100	44.4	0.0	55.6	0.0	44.4	-25.0

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 successfully 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	0.0	100.0	0.0	Only poor targeted
5-9	0.0	100.0	0.0	Only poor targeted
10-14	0.0	100.0	0.0	Only poor targeted
15-19	0.6	57.4	0.8	1.3:1
20-24	3.0	75.2	5.0	3.0:1
25-29	6.7	82.1	12.3	4.6:1
30-34	15.1	76.7	26.1	3.3:1
35-39	27.7	71.9	44.8	2.6:1
40-44	42.9	65.7	63.4	1.9:1
45-49	56.5	62.5	79.4	1.7:1
50-54	67.3	59.1	89.5	1.4:1
55-59	76.0	55.1	94.3	1.2:1
60-64	81.4	52.8	96.8	1.1:1
65-69	88.5	49.8	99.2	1.0:1
70-74	93.7	47.3	99.7	0.9:1
75-79	97.0	45.8	99.8	0.8:1
80-84	98.3	45.2	100.0	0.8:1
85-89	99.0	44.9	100.0	0.8:1
90-94	100.0	44.4	100.0	0.8:1
95-100	100.0	44.4	100.0	0.8:1

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	100.0
5-9	100.0
10-14	100.0
15-19	97.3
20-24	99.4
25-29	99.3
30-34	98.9
35-39	93.8
40-44	96.0
45-49	96.4
50-54	89.0
55-59	71.5
60-64	70.7
65-69	51.5
70-74	50.6
75-79	30.5
80-84	36.1
85-89	10.4
90-94	18.3
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

Score	Difference between estimate and true value			
	Diff.	Confidence interval (+/- percentage points)		
		90-percent	95-percent	99-percent
0-4	+0.0	0.0	0.0	0.0
5-9	+0.0	0.0	0.0	0.0
10-14	+0.0	0.0	0.0	0.0
15-19	-2.7	1.3	1.3	1.3
20-24	+0.5	0.7	0.9	1.1
25-29	+1.7	1.2	1.4	1.7
30-34	+2.8	0.8	1.0	1.2
35-39	-2.2	1.4	1.5	1.6
40-44	+1.6	0.8	1.0	1.2
45-49	+8.4	1.4	1.7	2.4
50-54	+0.3	1.5	1.7	2.2
55-59	-12.3	7.0	7.2	7.4
60-64	-7.2	4.9	5.1	5.4
65-69	+8.2	3.1	3.7	4.7
70-74	-10.2	6.7	6.9	7.4
75-79	+21.5	1.7	2.0	2.5
80-84	+34.9	0.6	0.7	0.9
85-89	+3.4	3.5	4.2	5.5
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 n	Difference between estimate and true value			
	Diff.	Confidence interval (+/- percentage points)		
		90-percent	95-percent	99-percent
1	+1.9	68.7	72.7	90.3
4	+1.0	36.7	41.9	57.2
8	+1.1	27.7	32.3	42.5
16	+1.3	22.4	25.8	32.2
32	+1.5	16.4	19.8	25.3
64	+1.7	11.8	13.5	17.3
128	+1.7	8.1	10.1	12.8
256	+1.6	5.8	7.0	9.7
512	+1.4	4.3	5.1	7.1
1,024	+1.4	3.0	3.7	4.8
2,048	+1.4	2.2	2.6	3.6
4,096	+1.4	1.6	1.9	2.5
8,192	+1.5	1.1	1.2	1.7
16,384	+1.5	0.8	0.9	1.2

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:	Undercoverage:	Leakage:	Exclusion:	Total Accuracy	BPAC
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	=> poverty line mistakenly targeted	=> poverty line correctly non-targeted	Inclusion + Exclusion	See text
0–4	0.0	82.3	0.0	17.7	17.7	–100.0
5–9	0.0	82.3	0.0	17.7	17.7	–100.0
10–14	0.0	82.3	0.0	17.7	17.7	–100.0
15–19	0.6	81.7	0.0	17.7	18.3	–98.4
20–24	2.9	79.4	0.0	17.7	20.6	–92.9
25–29	6.5	75.8	0.1	17.6	24.1	–84.0
30–34	14.7	67.6	0.4	17.3	32.0	–63.8
35–39	26.6	55.7	1.1	16.6	43.2	–34.1
40–44	40.8	41.5	2.0	15.6	56.5	+1.7
45–49	53.0	29.3	3.4	14.3	67.3	+33.1
50–54	62.7	19.6	4.6	13.1	75.7	+58.0
55–59	69.9	12.4	6.1	11.6	81.5	+77.2
60–64	73.8	8.6	7.6	10.1	83.8	+88.5
65–69	77.8	4.5	10.7	7.0	84.8	+87.0
70–74	81.0	1.3	12.7	5.0	85.9	+84.5
75–79	81.7	0.6	15.3	2.4	84.1	+81.4
80–84	81.8	0.5	16.5	1.2	82.9	+79.9
85–89	81.8	0.5	17.2	0.5	82.3	+79.1
90–94	82.3	0.0	17.7	0.0	82.3	+78.5
95–100	82.3	0.0	17.7	0.0	82.3	+78.5

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 successfully 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	0.0	100.0	0.0	Only poor targeted
5-9	0.0	100.0	0.0	Only poor targeted
10-14	0.0	100.0	0.0	Only poor targeted
15-19	0.6	100.0	0.8	Only poor targeted
20-24	3.0	98.7	3.5	76.4:1
25-29	6.7	98.3	8.0	59.2:1
30-34	15.1	97.1	17.9	34.0:1
35-39	27.7	96.0	32.3	24.2:1
40-44	42.9	95.2	49.6	19.9:1
45-49	56.5	93.9	64.4	15.5:1
50-54	67.3	93.1	76.2	13.5:1
55-59	76.0	92.0	84.9	11.5:1
60-64	81.4	90.6	89.6	9.7:1
65-69	88.5	87.9	94.5	7.3:1
70-74	93.7	86.4	98.4	6.4:1
75-79	97.0	84.2	99.3	5.3:1
80-84	98.3	83.2	99.3	5.0:1
85-89	99.0	82.6	99.4	4.8:1
90-94	100.0	82.3	100.0	4.7:1
95-100	100.0	82.3	100.0	4.7:1

Tables for the 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	100.0
5–9	100.0
10–14	100.0
15–19	80.5
20–24	61.6
25–29	62.3
30–34	45.1
35–39	48.4
40–44	36.8
45–49	27.2
50–54	17.7
55–59	14.9
60–64	7.5
65–69	5.6
70–74	2.6
75–79	8.9
80–84	1.9
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

Score	Difference between estimate and true value			
	Diff.	Confidence interval (+/- percentage points)		
		90-percent	95-percent	99-percent
0-4	+0.0	0.0	0.0	0.0
5-9	+0.0	0.0	0.0	0.0
10-14	+0.0	0.0	0.0	0.0
15-19	+38.3	8.8	10.2	13.2
20-24	+13.4	4.7	5.6	7.4
25-29	-5.9	4.7	5.0	5.7
30-34	-8.8	5.7	6.0	6.5
35-39	+2.1	2.0	2.4	3.2
40-44	+3.1	1.8	2.1	2.8
45-49	+0.4	1.7	2.1	2.7
50-54	-9.2	5.5	5.7	6.1
55-59	+0.1	1.8	2.1	2.8
60-64	+4.9	0.7	0.8	1.1
65-69	-1.4	1.3	1.4	1.8
70-74	+2.3	0.1	0.2	0.2
75-79	+8.5	0.2	0.3	0.3
80-84	+0.6	0.6	0.7	0.9
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 n	Difference between estimate and true value			
	Diff.	Confidence interval (+/- percentage points)		
		90-percent	95-percent	99-percent
1	-0.3	60.6	71.9	82.4
4	+1.2	37.0	43.0	51.0
8	+0.2	23.2	27.7	39.8
16	+0.2	15.3	18.4	25.1
32	+0.3	10.7	12.5	16.6
64	+0.1	7.6	9.2	11.8
128	+0.4	5.1	6.0	8.0
256	+0.4	3.8	4.5	5.8
512	+0.5	2.6	3.3	4.3
1,024	+0.5	1.9	2.2	2.9
2,048	+0.5	1.3	1.6	2.0
4,096	+0.5	0.9	1.1	1.4
8,192	+0.6	0.6	0.8	1.0
16,384	+0.6	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	0.0	28.9	0.0	71.1	71.1
5-9	0.0	28.9	0.0	71.1	71.1	-100.0
10-14	0.0	28.9	0.0	71.1	71.1	-100.0
15-19	0.2	28.7	0.4	70.7	70.9	-97.0
20-24	1.3	27.5	1.6	69.5	70.8	-85.1
25-29	3.9	25.0	2.8	68.3	72.2	-63.5
30-34	8.2	20.7	6.9	64.2	72.5	-19.1
35-39	14.2	14.7	13.5	57.6	71.7	+44.8
40-44	19.6	9.3	23.3	47.8	67.5	+19.5
45-49	23.6	5.3	32.9	38.2	61.8	-13.9
50-54	26.5	2.4	40.8	30.3	56.8	-41.3
55-59	27.7	1.2	48.3	22.8	50.5	-67.2
60-64	28.0	0.9	53.4	17.7	45.7	-84.9
65-69	28.7	0.2	59.8	11.3	40.1	-106.8
70-74	28.8	0.1	64.9	6.2	34.9	-124.8
75-79	28.8	0.1	68.2	3.0	31.8	-135.9
80-84	28.9	0.0	69.4	1.7	30.6	-140.2
85-89	28.9	0.0	70.1	1.0	29.9	-142.7
90-94	28.9	0.0	71.1	0.0	28.9	-146.1
95-100	28.9	0.0	71.1	0.0	28.9	-146.1

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 successfully 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	0.0	100.0	0.0	Only poor targeted
5–9	0.0	100.0	0.0	Only poor targeted
10–14	0.0	100.0	0.0	Only poor targeted
15–19	0.6	35.3	0.8	0.5:1
20–24	3.0	45.6	4.7	0.8:1
25–29	6.7	58.5	13.5	1.4:1
30–34	15.1	54.4	28.5	1.2:1
35–39	27.7	51.1	49.0	1.0:1
40–44	42.9	45.8	68.0	0.8:1
45–49	56.5	41.7	81.6	0.7:1
50–54	67.3	39.4	91.7	0.6:1
55–59	76.0	36.4	95.8	0.6:1
60–64	81.4	34.4	96.8	0.5:1
65–69	88.5	32.5	99.4	0.5:1
70–74	93.7	30.7	99.6	0.4:1
75–79	97.0	29.7	99.8	0.4:1
80–84	98.3	29.4	100.0	0.4:1
85–89	99.0	29.2	100.0	0.4:1
90–94	100.0	28.9	100.0	0.4:1
95–100	100.0	28.9	100.0	0.4:1

Tables for the \$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	100.0
5-9	100.0
10-14	100.0
15-19	88.5
20-24	85.8
25-29	91.2
30-34	74.4
35-39	69.0
40-44	62.4
45-49	51.5
50-54	36.9
55-59	27.5
60-64	13.8
65-69	9.6
70-74	3.1
75-79	8.9
80-84	1.9
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

Score	Difference between estimate and true value			
	Diff.	Confidence interval (+/- percentage points)		
		90-percent	95-percent	99-percent
0-4	+0.0	0.0	0.0	0.0
5-9	+0.0	0.0	0.0	0.0
10-14	+0.0	0.0	0.0	0.0
15-19	+9.6	5.9	6.8	8.8
20-24	+2.9	3.2	3.8	4.9
25-29	+0.4	2.1	2.5	3.1
30-34	-0.4	2.3	2.7	3.5
35-39	-3.3	2.5	2.7	3.0
40-44	+8.3	1.9	2.2	2.8
45-49	-0.2	2.1	2.4	3.4
50-54	-7.4	4.8	5.1	5.4
55-59	+4.7	2.0	2.4	3.0
60-64	+0.9	2.0	2.5	3.4
65-69	+0.1	1.3	1.5	2.1
70-74	+1.3	0.5	0.6	0.8
75-79	+8.5	0.2	0.3	0.3
80-84	+0.6	0.6	0.7	0.9
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 n	Difference between estimate and true value			
	Diff.	Confidence interval (+/- percentage points)		
		90-percent	95-percent	99-percent
1	+1.3	66.0	73.4	89.5
4	+1.4	36.2	43.1	56.6
8	+0.7	25.3	29.9	40.4
16	+0.7	17.0	20.9	28.3
32	+0.8	11.3	13.7	18.1
64	+0.7	8.1	9.5	13.2
128	+1.0	5.3	6.1	7.8
256	+1.0	3.9	4.4	6.1
512	+1.1	2.7	3.1	4.4
1,024	+1.2	1.8	2.2	2.8
2,048	+1.1	1.3	1.6	2.0
4,096	+1.2	1.0	1.1	1.5
8,192	+1.2	0.7	0.8	1.1
16,384	+1.2	0.5	0.6	0.7

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:	Undercoverage:	Leakage:	Exclusion:	Total Accuracy	BPAC
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	=> poverty line mistakenly targeted	=> poverty line correctly non-targeted	Inclusion + Exclusion	See text
0–4	0.0	46.3	0.0	53.7	53.7	–100.0
5–9	0.0	46.3	0.0	53.7	53.7	–100.0
10–14	0.0	46.3	0.0	53.7	53.7	–100.0
15–19	0.4	45.9	0.2	53.4	53.9	–97.7
20–24	2.3	44.0	0.7	53.0	55.3	–88.7
25–29	5.7	40.7	1.0	52.7	58.3	–73.4
30–34	12.1	34.2	3.0	50.7	62.8	–41.2
35–39	21.0	25.4	6.7	47.0	67.9	+5.0
40–44	29.7	16.6	13.2	40.5	70.2	+56.7
45–49	36.9	9.4	19.6	34.1	71.0	+57.8
50–54	41.9	4.4	25.4	28.2	70.2	+45.1
55–59	43.9	2.4	32.0	21.6	65.5	+30.8
60–64	44.9	1.5	36.5	17.1	62.0	+21.2
65–69	46.0	0.4	42.5	11.2	57.1	+8.3
70–74	46.2	0.1	47.5	6.2	52.4	–2.5
75–79	46.3	0.1	50.7	3.0	49.2	–9.4
80–84	46.3	0.0	51.9	1.7	48.1	–12.1
85–89	46.3	0.0	52.7	1.0	47.3	–13.7
90–94	46.3	0.0	53.7	0.0	46.3	–15.8
95–100	46.3	0.0	53.7	0.0	46.3	–15.8

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

Figure 12 (\$1.25/day 2005 PPP line): For a given score cut-off, the percentage of all households who are targeted (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 successfully 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	0.0	100.0	0.0	Only poor targeted
5-9	0.0	100.0	0.0	Only poor targeted
10-14	0.0	100.0	0.0	Only poor targeted
15-19	0.6	66.3	0.9	2.0:1
20-24	3.0	77.7	5.0	3.5:1
25-29	6.7	85.1	12.2	5.7:1
30-34	15.1	80.1	26.2	4.0:1
35-39	27.7	75.8	45.3	3.1:1
40-44	42.9	69.2	64.1	2.3:1
45-49	56.5	65.3	79.6	1.9:1
50-54	67.3	62.2	90.4	1.6:1
55-59	76.0	57.8	94.8	1.4:1
60-64	81.4	55.1	96.8	1.2:1
65-69	88.5	52.0	99.2	1.1:1
70-74	93.7	49.3	99.7	1.0:1
75-79	97.0	47.7	99.9	0.9:1
80-84	98.3	47.1	100.0	0.9:1
85-89	99.0	46.8	100.0	0.9:1
90-94	100.0	46.3	100.0	0.9:1
95-100	100.0	46.3	100.0	0.9:1

Tables for the \$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	100.0
10-14	100.0
15-19	97.3
20-24	99.4
25-29	99.3
30-34	98.9
35-39	94.3
40-44	96.2
45-49	96.9
50-54	88.2
55-59	77.7
60-64	80.8
65-69	57.6
70-74	43.4
75-79	29.2
80-84	19.7
85-89	10.4
90-94	16.1
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

Score	Difference between estimate and true value			
	Diff.	Confidence interval (+/- percentage points)		
		90-percent	95-percent	99-percent
0-4	+0.0	0.0	0.0	0.0
5-9	+0.0	0.0	0.0	0.0
10-14	+0.0	0.0	0.0	0.0
15-19	-2.7	1.3	1.3	1.3
20-24	+0.5	0.7	0.9	1.1
25-29	+1.7	1.2	1.4	1.7
30-34	+2.8	0.8	1.0	1.2
35-39	-2.2	1.4	1.5	1.6
40-44	-0.0	0.6	0.7	0.9
45-49	+5.1	1.1	1.3	1.7
50-54	-0.9	1.5	1.7	2.1
55-59	-3.8	2.8	3.1	3.3
60-64	+0.4	2.4	2.8	3.7
65-69	+14.7	3.1	3.7	4.7
70-74	-16.0	9.6	9.8	10.3
75-79	+22.1	1.5	1.8	2.3
80-84	+18.5	0.6	0.7	0.9
85-89	-0.2	4.0	4.8	6.0
90-94	-32.1	19.5	20.2	21.1
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 n	Difference between estimate and true value			
	Diff.	Confidence interval (+/- percentage points)		
		90-percent	95-percent	99-percent
1	+1.9	61.6	76.4	91.4
4	+1.0	36.1	43.5	58.4
8	+1.1	28.5	33.0	42.1
16	+1.2	21.6	26.0	32.0
32	+1.3	15.9	19.4	25.6
64	+1.5	11.6	13.7	17.5
128	+1.4	8.1	9.7	12.8
256	+1.3	5.8	7.2	9.6
512	+1.1	4.3	5.1	6.8
1,024	+1.2	3.0	3.7	4.9
2,048	+1.2	2.1	2.6	3.6
4,096	+1.2	1.6	1.9	2.6
8,192	+1.2	1.0	1.2	1.7
16,384	+1.2	0.7	0.9	1.2

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:	Undercoverage:	Leakage:	Exclusion:	Total Accuracy	BPAC
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	=> poverty line mistakenly targeted	=> poverty line correctly non-targeted	Inclusion + Exclusion	See text
0-4	0.0	82.6	0.0	17.4	17.4	-100.0
5-9	0.0	82.6	0.0	17.4	17.4	-100.0
10-14	0.0	82.6	0.0	17.4	17.4	-100.0
15-19	0.6	82.0	0.0	17.4	18.0	-98.5
20-24	2.9	79.7	0.0	17.3	20.3	-92.9
25-29	6.5	76.1	0.1	17.3	23.8	-84.0
30-34	14.7	67.9	0.4	16.9	31.6	-63.9
35-39	26.7	56.0	1.0	16.4	43.0	-34.2
40-44	41.2	41.5	1.7	15.6	56.8	+1.7
45-49	53.7	28.9	2.8	14.6	68.3	+33.3
50-54	63.4	19.2	3.9	13.5	76.9	+58.2
55-59	70.5	12.1	5.4	11.9	82.5	+77.3
60-64	74.5	8.1	6.8	10.5	85.1	+88.7
65-69	78.5	4.1	9.9	7.4	86.0	+88.0
70-74	81.5	1.1	12.2	5.2	86.7	+85.3
75-79	82.0	0.6	14.9	2.4	84.5	+81.9
80-84	82.1	0.5	16.2	1.2	83.3	+80.4
85-89	82.2	0.4	16.8	0.6	82.8	+79.7
90-94	82.6	0.0	17.4	0.0	82.6	+79.0
95-100	82.6	0.0	17.4	0.0	82.6	+79.0

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 successfully 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	0.0	100.0	0.0	Only poor targeted
5-9	0.0	100.0	0.0	Only poor targeted
10-14	0.0	100.0	0.0	Only poor targeted
15-19	0.6	100.0	0.8	Only poor targeted
20-24	3.0	98.7	3.5	76.4:1
25-29	6.7	98.3	7.9	59.2:1
30-34	15.1	97.1	17.8	34.0:1
35-39	27.7	96.4	32.3	26.5:1
40-44	42.9	95.9	49.8	23.6:1
45-49	56.5	95.1	65.0	19.4:1
50-54	67.3	94.2	76.8	16.2:1
55-59	76.0	92.8	85.4	13.0:1
60-64	81.4	91.6	90.2	10.9:1
65-69	88.5	88.8	95.1	7.9:1
70-74	93.7	87.0	98.7	6.7:1
75-79	97.0	84.6	99.3	5.5:1
80-84	98.3	83.5	99.4	5.1:1
85-89	99.0	83.0	99.5	4.9:1
90-94	100.0	82.6	100.0	4.8:1
95-100	100.0	82.6	100.0	4.8:1