Simple Poverty Scorecard®
Malawi

Mark Schreiner

revised 1 February 2011

This document and related tools are at `SimplePovertyScorecard.com`.

Abstract
The Simple Poverty Scorecard® uses ten low-cost indicators from Malawi's 2004/5 Integrated Household Survey to estimate the likelihood that a household has consumption below a given poverty line. Field workers can collect responses in about ten minutes. The scorecard's accuracy is reported for a range of poverty lines. The scorecard is a practical way for pro-poor programs in Malawi to measure poverty rates, to track changes in poverty rates over time, and to segment clients for targeted services.

Acknowledgements
This redone-from-scratch paper replaces an earlier one that had an error. It was funded by the Consultative Group to Assist the Poorest under the CGAP/Ford Social Indicators Project. The 2004/5 IHS data is from Malawi's National Statistical Office. Thanks go to Malika Anand, Todd Benson, Shaohua Chen, Frank DeGiovanni, Syed Hashemi, Anthony Leegwater, Antonio Nucifora, Tony Sheldon, and Don Sillers. The Simple Poverty Scorecard® is the same as what Grameen Foundation calls the Progress out of Poverty Index®. The PPI® is a performance-management tool that Grameen Foundation promotes to help institutions achieve their social objectives more effectively.

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

Simple Poverty Scorecard®

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

Indicator	Value	Points	Score
1. How many household members are 14-years-old or younger?	A. Five or more	0	
	B. Four	4	
	C. Three	6	
	D. Two	12	
	E. One	19	
	F. None	30	
2. How many household members worked in their main activity in the past seven days as a farmer (*mlimi*)?	A. Four or more	0	
	B. Three	2	
	C. Two	7	
	D. One	8	
	E. None	10	
3. Can the female head/spouse read a one-page letter in any language?	A. No	0	
	B. Yes	5	
	C. No female head/spouse	9	
4. The roof of the main dwelling is predominantly made of what material?	A. Grass	0	
	B. Anything besides grass	4	
5. What is your main source of cooking fuel?	A. Collected firewood from forest reserve, crop residue, sawdust, animal waste, or other	0	
	B. Collected firewood from unfarmed areas of community	1	
	C. Collected firewood from own woodlot, community woodlot, or other places	5	
	D. Purchased firewood	7	
	E. Paraffin, charcoal, gas, or electricity	9	
6. What is your main source of lighting fuel?	A. Collected firewood, grass, or other	0	
	B. Paraffin	4	
	C. Purchased firewood, electricity, gas, battery/dry cell (torch), or candles	13	
7. Does the household own any lanterns (paraffin)?	A. No	0	
	B. Yes	5	
8. Does the household own any bicycles, motorcycles/scooters, cars, mini-buses, or lorries?	A. No	0	
	B. Yes	5	
9. Does the household own any irons (for pressing clothes)?	A. No	0	
	B. Yes	8	
10. How many sickles does the household own?	A. None	0	
	B. One	3	
	C. Two or more	7	

Score:

Simple Poverty Scorecard®
Malawi

1. Introduction

This paper presents the Simple Poverty Scorecard®, easy-to-use too that pro-poor programs in Malawi can use to estimate the likelihood that a household has expenditure below a given poverty line, to measure groups' poverty rates at a point in time, to track changes in groups' poverty rates between two points in time, and to target services to households.

The direct approach to poverty measurement via surveys is difficult and costly. As a case in point, the 2004/5 Malawi Integrated Household Survey (IHS) runs 53 pages. The expenditure module asks households a battery of questions about more than 300 expenditure items. An example set of questions is: "Over the past one week (7 days), did you or others in your household consume any maize *ufa mgaiwa* (normal flour)? How much *ufa mgaiwa* (normal flour) in total did your household consume in the past week? How much came from purchases? How much did you spend? How much came from own-production? How much came from gifts and other sources? Now then, Over the past one week (7 days), did you or others in your household consume any maize *ufa* refined (fine flour)? . . .".

In contrast, the indirect approach via poverty scoring is simple, quick, and inexpensive. It uses 10 verifiable indicators (such as "What is your main source of cooking fuel?" or "Does your household own any irons (for pressing clothes)?") to get a score that is highly correlated with poverty status as measured by the exhaustive survey.

The 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 workers) or blunt (such as rules based on land-ownership or housing quality). Results from these approaches are not comparable across organizations, they may be costly, and their accuracy is unknown.

Pro-poor organizations can use the scorecard to measure the share of their participants who are below a given poverty line, such as the Millennium Development Goals' $1.25/day poverty line 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 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, not because they do not work, but because they are presented (when they are presented at all) as tables of regression coefficients incomprehensible to lay people (with cryptic indicator names such as “`HHSIZE_2`”, negative values, and many decimal places). Thanks to the predictive-modeling phenomenon known as the “flat maximum” (discussed later), simple scorecards can be 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-assessment tools.

The scorecard (Figure 1) is based on the 2004/5 IHS conducted by the National Statistical Office of Malawi (NSO) from March 2004 to March 2005. Indicators are selected to be:

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

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

Poverty scoring can be used to estimate three basic quantities. First, it can estimate a particular household's "poverty likelihood", that is, the probability that the household has per-capita 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 estimate is 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 managers choose an appropriate targeting cut-off for their purposes, this paper reports several measures of targeting accuracy for a range of possible cut-offs.

This paper presents a single scorecard whose indicators and points are derived from household expenditure data and Malawi's national poverty line. Scores from this one scorecard are calibrated to poverty likelihoods for seven poverty lines.

The scorecard is constructed and calibrated using half of the data from the 2004/5 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. (The survey approach is unbiased by assumption.) There is bias because scoring must assume that the future relationships between indicators and poverty will be the same as in the data used to build the scorecard. Of course, this assumption—ubiquitous and inevitable in predictive modeling—holds only partly.

When applied to the validation sample with bootstrap samples of n = 16,384, the difference between scorecard estimates of groups' poverty rates and the true rates at a point in time for the national line is +0.1 percentage points, and the average absolute difference across all seven lines is 0.2 percentage points. These differences are due to sampling variation and not bias; the average of each difference would be zero if the whole 2004/5 IHS were to be repeatedly redrawn and divided into sub-samples before repeating the entire process of building and calibrating scorecards.

The 90-percent confidence intervals for these estimates are ±0.6 percentage points or less. For n = 1,024, the 90-percent intervals are ±2.2 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 the field. Sections 5 and 6 detail

[1] Important examples include nationally representative samples at a different point in time or non-nationally representative sub-groups (Tarozzi and Deaton, 2007).

the estimation of households' poverty likelihoods and of groups' poverty rates at a point in time. Section 7 discusses estimating changes in poverty rates through time, and Section 8 covers targeting. Section 9 places the new scorecard here in the context of existing exercises for Malawi, and Section 10 is a summary.

2. Data and poverty lines

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

2.1 Data

The scorecard is based on data from the 2004/5 IHS. Households are randomly divided into two sub-samples (Figure 2):

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

2.2 Poverty rates and poverty lines

2.2.1 Rates

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

Beyond this general definition, there two special cases, *household-level poverty rates* and *person-level poverty rates*. With household-level rates, each household is counted as if it had only one person, regardless of true household size, so all households are counted equally. With person-level rates (the “head-count index”), each household is weighted by the number of people in it, so larger households count more.

For example, consider a group of two households, the first with one member and the second with two members. Suppose further that the first household has per-capita expenditure above a poverty line (it is "non-poor") and that the second household has per-capita 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 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 rates and poverty lines for Malawi at both the household-level and the person-level for its regions (Urban, Northern Rural, Central Rural, and Southern Rural) and for Malawi as a whole. The scorecard is constructed using the 2004/5 IHS and household-level lines, scores are calibrated to household-level poverty likelihoods, and accuracy is measured for household-level rates.

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

2.2.2 Poverty lines

The national poverty line of 43.92 Kwacha (MWK) per person per day is defined as the food (ultra) poverty line (the cost of 2,400 calories, or MWK27.25) plus the average non-food expenditure for households whose food expenditure per capita is within five percent of the food poverty line (World Bank, 2005).

The scorecard here is constructed using the national poverty line. For Malawi as a whole, the national line implies a household-level poverty rate of 43.6 percent and a person-level poverty rate of 52.4 percent. For the food poverty line, the household-level poverty rate for Malawi as a whole is 16.6 percent, and the person-level rate is 22.2 percent.

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:

- National
- Food
- USAID "extreme"
- $1.08/day 1993 PPP
- $2.16/day 1993 PPP
- $1.25/day 2005 PPP
- $2.50/day 2005 PPP

The USAID "extreme" line (U.S. Congress, 2002) is defined as the median expenditure of people (not households) below the national line.

The $1.08/day 1993 PPP is from World Bank (2005), correcting for their use of $1.00/day. The $2.16/day 1993 PPP line is twice the $1.08/day line.

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

- 2005 PPP exchange rate for "individual consumption expenditure by households" of MWK56.92 per $1 (World Bank, 2008)
- Average all-Malawi Consumer Price Index (CPI) for March 2004 to March 2005 of 178.915[2]
- Average all-Malawi CPI for 2005 of 198.475

[2] http://www.rbm.mw/inflation_rates_detailed.aspx, retrieved 4 January 2010.

Thus, the \$1.25/day 2005 PPP line for Malawi applied to the 2004/5 IHS is (Sillers, 2006):

$$(\text{2005 PPP exchange rate}) \cdot \$1.25 \cdot \left(\frac{\text{CPI}_{\text{Ave. March 2004 to March 2005}}}{\text{CPI}_{\text{Ave. 2005}}} \right) =$$

$$\left(\frac{\text{MWK56.92}}{\$1.00} \right) \cdot \$1.25 \cdot \left(\frac{178.915}{198.475} \right) = \text{MWK64.14}.$$

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

The \$1.25/day line of MWK64.14 does not exactly match that in Figure 2 because the national figure is divided by each household's price deflator to account for regional differences in cost-of-living. Aggregating the results back up to the national level with personal-level weights produces the number in Figure 2.

3. Scorecard construction

About 90 potential indicators are initially prepared in the areas of:

- Family composition (such as household size)
- Education (such as literacy of the female household head)
- Housing (such as the main source of cooking fuel)
- Ownership of durable goods (such as irons and sickles)

Each indicator is first screened with the entropy-based “uncertainty coefficient” (Goodman and Kruskal, 1979) that measures how well it predicts poverty on its own. Figure 3 lists all potential indicators, ranked by uncertainty coefficient.

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

The scorecard itself is built using the national poverty line and Logit regression on the construction sub-sample (Figure 2). Indicator selection uses both judgment and statistics (forward stepwise). 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 scorecard here applies to all of Malawi. 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 segmentsation 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 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)

A field worker using the paper scorecard would:

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

Of course, field workers must be trained. The quality of outputs depends on the quality of inputs. If organizations or field workers 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 (2007b) and Toohig (2008) are useful nuts-and-bolts guides for budgeting, training field workers 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 of terms and

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

concepts in the scorecard is essential (Appendix A). 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 the suggested procedure for poverty scoring in Malawi.

In terms of sampling design, an organization must make choices about:

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

In general, the sampling design should follow from the organization's goals for the exercise.

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

The subjects to be scored can be:

- All participants (or all new participants)
- A representative sample of all participants (or of all new participants)
- All participants (or all new participants) in a representative sample of branches
- A representative sample of all participants (or of all new participants) in a representative sample of branches
- A representative sample of participants relevant for a given business question

If not determined by other factors, the number of participants to be scored can be derived from sample-size formulas (presented later) for a desired 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 worker 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 are illustrated by BRAC and ASA, two microlenders in Bangladesh who each have more than 7 million participants and who are applying the Simple Poverty Scorecard® (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.

5. Estimates of household poverty likelihoods

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

To get absolute units, scores must be converted to *poverty likelihoods*, that is, probabilities of being below a poverty line. This is done via simple look-up tables. For the example of the national line, scores of 35–39 have a poverty likelihood of 47.8 percent, and scores of 40–44 have a poverty likelihood of 36.1 percent (Figure 4).

The poverty likelihood associated with a score varies by poverty line. For example, scores of 35–39 are associated with a poverty likelihood of 47.8 percent for the national line but 12.0 percent for the food line.[4]

5.1 Calibrating scores with poverty likelihoods

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

[4] Starting with Figure 4, most figures have seven versions, one for each poverty line. To keep them straight, they are grouped by poverty line. Single tables that pertain to all poverty lines are placed with the tables for the national line.

For the example of the national line (Figure 5), there are 12,401 households in the calibration sub-sample with a score of 35–39, of whom 5,924 are below the poverty line. The estimated poverty likelihood associated with a score of 35–39 is then 47.8 percent, because 5,924 ÷ 12,401 = 47.8 percent.

To illustrate with the national line and a score of 40–44, there are 11,846 households in the calibration sample, of whom 4,272 are below the line (Figure 5). Thus, the poverty likelihood for this score is 4,272 ÷ 11,846 = 36.1 percent.

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

Figure 6 shows, for all scores, the likelihood that expenditure falls in a range demarcated by two adjacent poverty lines. For example, the daily expenditure of someone with a score of 35–39 falls in the following ranges with probability:

- 12.0 percent below the food line
- 2.8 percent between the food and USAID "extreme" lines
- 33.0 percent between the USAID "extreme" lines and the national line
- 29.8 percent between the national and $1.25/day 2005 PPP lines
- 21.2 percent between the $1.25/day 2005 PPP and $2.50/day 2005 PPP lines
- 1.3 percent above the $2.50/day 2005 PPP line

Even though the scorecard is constructed partly based on judgment, the calibration process produces poverty likelihoods that are objective, that is, derived from survey data on 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 based only on judgment (Fuller, 2006; Caire, 2004; Schreiner *et al.*, 2004). Of course, the scorecard here is constructed with both data and judgment. The fact that this paper acknowledges that some choices in scorecard construction—as in any statistical analysis—are informed by judgment in no way impugns the objectivity of the poverty likelihoods, as this depends on using data in score calibration, not on using data (and nothing else) in scorecard construction.

Although the points in Malawi's 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, converting scores to poverty likelihoods requires no arithmetic at all, just a look-up table. This non-parametric calibration can also improve accuracy, especially with large calibration samples.

5.2 Accuracy of estimates of poverty likelihoods

As long as the relationship between indicators and poverty does not change, 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 and of changes in poverty rates between two points in time.[5]

Of course, the relationship between indicators and poverty does change with time, so the scorecard applied after March 2005 (as it must be in practice) will generally be biased.

How accurate are estimates of poverty likelihoods? To measure, the scorecard is applied to 1,000 bootstrap samples of size n = 16,384 from the validation sub-sample (Figure 2). 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, 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 35–39 in the validation sample is too high by 0.5 percentage points (Figure 7). For scores of 30–34, the estimate is too low by 1.8 percentage points.[6]

---

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

For the validation sample, the 90-percent confidence interval for the differences for scores of 35–39 is ±1.9 percentage points (Figure 7). This means that in 900 of 1,000 bootstraps, the difference between the estimate and the true value is between –1.4 and +2.4 percentage points (because 0.5 – 1.9 = –1.4, and 0.5 + 1.9 = +2.4). In 950 of 1,000 bootstraps (95 percent), the difference is 0.5 ±2.3 percentage points, and in 990 of 1,000 bootstraps (99 percent), the difference is 0.5 ±3.2 percentage points.

For almost all scores below 85, Figure 7 shows some differences between estimated poverty likelihoods and true values. This is because the validation sub-sample is a single sample that—thanks to sampling variation—differs in distribution from the construction/calibration sub-samples and from Malawi'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.

Of course, if estimates of groups' poverty rates are to be usefully accurate, then errors for individual households must largely cancel out. As discussed later, this is generally the case.

---

[6] There are differences, despite the estimator's unbiasedness, because the estimates come from a single sample. Their average difference would be zero if samples were repeatedly drawn from the population and split into sub-samples before repeating the entire scorecard-building process.

By construction, the scorecard here is unbiased. It may still, however, be *overfit* when applied after the end of the IHS fieldwork in March 2005. That is, it may fit the IHS data so closely that it captures not only some timeless patterns but also some random patterns that, due to sampling variation, show up only in the IHS. Or the scorecard may be overfit in the sense that it becomes biased as the relationships between indicators and poverty change over time.

Overfitting can be mitigated by simplifying the scorecard and by not relying only on data but rather also considering experience, judgment, and theory. Of course, the scorecard here does this. Bootstrapping can also mitigate overfitting by reducing (but not eliminating) dependence on a single sampling instance. 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, much of the differences may come from non-scorecard sources such as changes in the relationships between indicators and poverty, sampling variation, inconsistencies in data quality, and inconsistencies/imperfections in cost-of-living adjustments. 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 82.5, 59.3, and 36.1 percent (national line, Figure 4). The group's estimated poverty rate is the households' average poverty likelihood of (82.5 + 59.3 + 36.1) ÷ 3 = 59.3 percent.[7]

6.1 Accuracy of estimated poverty rates at a point in time

For the Malawi scorecard applied to the validation sample with n = 16,384, the absolute differences between the estimated poverty rate at a point in time and the true rate are 0.5 percentage points or less (Figure 8, summarizing Figure 9 across poverty lines). The average absolute difference across the seven poverty lines is 0.2 percentage points. At least part of these differences is due to sampling variation in the validation sample and in the division of the 2004/5 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 ±0.6 percentage points or less (Figure 8). This means that in 900 of 1,000 bootstraps of this size, the difference between the

[7] In general, the group's poverty rate is *not* the poverty likelihood associated with the average score. Here, it is pure coincidence that the poverty likelihood associated with the average score of 30 is 59.3 percent, which is also the average of the three poverty likelihoods associated with each of the three scores.

estimate and the true value is within 0.6 percentage points of the average difference. In the specific case of the national line and the validation sample, 90 percent of all samples of n = 16,384 produce estimates that differ from the true value in the range of +0.1 – 0.6 = –0.5 to +0.1 + 0.6 = +0.7 percentage points. This is because +0.1 is the average difference, and ±0.6 is its 90-percent confidence interval. The average difference is +0.1 because the average scorecard estimate is too high by 0.1 percentage points; it tends to estimate a poverty rate of 43.7 percent for the validation sample, but the true value is 43.6 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) values, 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-assessment tools (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, $\sqrt{\dfrac{p \cdot (1-p)}{n}}$,

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

n is the sample size.

For example, this implies that for a sample n of 16,384 with 90-percent confidence ($z = 1.64$) and a poverty rate p of 43.6 percent (the average poverty rate in the construction and calibration samples in Figure 2 for the national line), the confidence interval c is $+/- z \cdot \sqrt{\dfrac{p \cdot (1-p)}{n}} = +/- 1.64 \cdot \sqrt{\dfrac{0.436 \cdot (1-0.436)}{16{,}384}} = \pm 0.635$ percentage points.

Scorecards, however, do not measure poverty directly, so this formula is not immediately applicable. To derive a formula for the Malawi scorecard, consider Figure 9, which reports empirical confidence intervals c for the differences for the scorecard applied to 1,000 bootstrap samples of various sample sizes from the validation sample.

For n = 16,384 and the national line, the 90-percent confidence interval is 0.550 percentage points.[8]

Thus, the 90-percent confidence interval with n = 16,384 is ±0.550 percentage points for the Malawi scorecard and ±0.635 percentage points for direct measurement. The ratio of the two intervals is 0.550 ÷ 0.635 = 0.87.

Now consider the same case, but with n = 8,192. The confidence interval under direct measurement is $+/-1.64 \cdot \sqrt{\dfrac{0.436 \cdot (1-0.436)}{8{,}192}}$ = ±0.899 percentage points. The empirical confidence interval with the Malawi scorecard (Figure 9) is 0.760 percentage points. Thus for n = 8,192, the ratio of the two intervals is 0.760 ÷ 0.899 = 0.85.

This ratio of 0.85 for n = 8,182 is not far from the ratio of 0.87 for n = 16,384. Across all sample sizes of 256 or more in Figure 9, the average ratio turns out to be 0.86, implying that confidence intervals for indirect estimates of poverty rates via the Malawi scorecard and this poverty line are about 14 percent narrower than confidence intervals for direct estimates via the 2004/5 IHS. This 0.86 appears in Figure 8 as the "α factor" because if α = 0.86, then the formula relating confidence intervals c and standard errors σ for the Malawi scorecard is $c = +/- z \cdot \alpha \cdot \sigma$. That is, formula for the standard error σ for point-in-time estimates of poverty rates via scoring is

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

[8] Due to rounding, Figure 9 displays 0.6, not 0.550.

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. This is the case for all seven poverty lines in Figure 8.

The formula relating confidence intervals with standard errors for poverty scoring can be rearranged to give a formula for determining sample size before measurement.[9] 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 $c = 0.04425$ and $z = 1.64$ (90-percent confidence). Then the formula gives $n = \left(\frac{0.86 \cdot 1.64}{0.04425}\right)^2 \cdot 0.436 \cdot (1 - 0.436) = 250$, close to the sample size of 256 observed for these parameters in Figure 9 for the national line.

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

In practice after the end of fieldwork for the IHS in March 2005, an organization would select a poverty line (say, the national line), select a desired confidence level

---

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

(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 43.6 percent national average in the 2004/5 IHS in Figure 2), look up α (here, 0.86), assume that the scorecard will still work in the future and/or for non-nationally representative sub-groups,[10] and then compute the required sample size. In this illustration, $n = \left(\frac{0.86 \cdot 1.64}{0.02}\right)^2 \cdot 0.436 \cdot (1 - 0.436) = 1{,}223$.

[10] 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 March 2005 will resemble that in the 2004/5 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 2004/5 IHS, this paper cannot test estimates of change over time for Malawi, 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 82.5, 59.3, and 36.1 percent (national line, Figure 4). The group's baseline estimated poverty rate is the households' average poverty likelihood of (82.5 + 59.3 + 36.1) ÷ 3 = 59.3 percent.

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

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

By way of illustration, suppose that a year later on Jan. 1, 2012, the program samples three additional households who are in the same cohort as the three households originally sampled (or suppose that the program scores the same three original households a second time) and finds that their scores are 25, 35, and 45 (poverty likelihoods of 70.0, 47.8, and 25.5 percent, national line, Figure 4). Their average poverty likelihood at follow-up is now (70.0 + 47.8 + 25.5) ÷ 3 = 47.8 percent, an improvement of 59.3 – 47.8 = 11.5 percentage points.[11]

This suggests that about one in nine participants in this hypothetical example crossed the poverty line in 2011.[12] Among those who started below the line, about one in five (11.5 ÷ 59.3 = 19.4 percent) on net ended up above the line.[13]

---

[11] 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.
[12] 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 2004/5 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 Malawi scorecard to estimate change. The rest of this section suggests approximate formulas for standard errors and sample sizes that may be used until there is additional data.

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

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

z, c, and p are defined as above, n is the sample size at both baseline and follow-up,[14] and α is the average (across a range of bootstrapped sample sizes) of the ratio of the observed confidence interval from a scorecard and the theoretical confidence interval under direct measurement.

[13] Poverty scoring does not reveal the reasons for this change.

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

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

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

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

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 $\hat{p} = 0.436$ (from Figure 2). Then the baseline sample size is $n = 2 \cdot \left(\frac{1.19 \cdot 1.64}{0.02} \right)^2 \cdot 0.436 \cdot (1 - 0.436) =$ 4,683, and the follow-up sample size is also 4,683.

7.4 Accuracy for estimated change for one sample, scored twice

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

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

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:

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

[15] 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 Malawi scorecard is applied twice (once after March 2005 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 43.6 percent ($p_{2004/5} = 0.436$, Figure 2), and suppose $\alpha = 1.30$. Then the baseline sample size is

$n = 2 \cdot \left(\frac{1.30 \cdot 1.64}{0.02}\right)^2 \cdot \{-0.02 + 0.016 \cdot 3 + 0.47 \cdot [0.436 \cdot (1 - 0.436)]\} = 3{,}264$. The same group of 3,264 households is scored at follow-up as well.

8. Targeting

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

There is a distinction between *targeting status* (scoring at or below a targeting cut-off) and *poverty status* (having 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 35–39, outcomes for the national line in the validation sample are:

- Inclusion: 35.1 percent are below the line and correctly targeted
- Undercoverage: 8.5 percent are below the line and mistakenly not targeted
- Leakage: 17.4 percent are above the line and mistakenly targeted
- Exclusion: 39.0 percent are above the line and correctly not targeted

Increasing the cut-off to 40–44 improves inclusion and undercoverage but worsens leakage and exclusion:

- Inclusion: 39.2 percent are below the line and correctly targeted
- Undercoverage: 4.4 percent are below the line and mistakenly not targeted
- Leakage: 25.2 percent are above the line and mistakenly targeted
- Exclusion: 31.2 percent are above the line and correctly not targeted

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

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

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

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

The most difficult step is assigning benefits and costs to targeting outcomes. A program that uses targeting—with or without scoring—should thoughtfully consider how it values successful inclusion and 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:

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

Figure 11 shows "Total Accuracy" for all cut-offs for the Malawi scorecard. For the national line in the validation sample, total net benefit is greatest (74.8) for a cut-off of 30–34, with about three in four households in Malawi correctly classified.

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

As an alternative to assigning benefits and costs to targeting outcomes and then choosing a cut-off to maximize total net benefit, a program could set a cut-off to achieve a desired poverty rate among targeted households. The third column of Figure 12 ("% targeted who are poor") shows, for the Malawi scorecard applied to the

[16] Figure 11 also reports "BPAC", discussed in Section 9.

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 39 or less would target 52.6 percent of all households (second column) and produce a poverty rate among those targeted of 66.8 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 39 or less, 80.6 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 39 or less, covering 2.0 poor households means leaking to 1 non-poor household

9. Context of poverty-assessment tools for Malawi

This section discusses several existing poverty-assessment tools for Malawi in terms of goals, methods, poverty lines, indicators, accuracy, standard-error formula, and costs.

In general, all the tools use the national poverty line and adjust for regional differences in cost-of-living (unless they estimate a different definition of poverty using Principal Components Analysis). None of the tools report standard-error formulas. All of the tools except one use the 1997/8 IHS or the 2004/5 IHS.

9.1 Benson

Rather than build a poverty-assessment tool using indicators at the level of the household to predict household poverty likelihoods, Benson (2002) uses indicators at the level of Malawi's 308 Traditional Authorities (TAs) from the 1997/8 IHS to build tools to rank TAs for purposes of geographic poverty targeting. Much like this paper tests whether a simple, low-cost household-level scorecard can substitute for high-cost expenditure surveys, Benson tests whether a simple, low-cost TA-level tool based on the 1997/8 IHS is accurate enough to substitute for more-costly poverty mapping using the 1997/8 IHS and the 1998 census.[17]

---

[17] *Poverty mapping* uses national expenditure survey data to build a household-level poverty-assessment tool using only indicators in both the national expenditure survey and in census data. This tool is then applied to census data to estimate poverty rates at

Benson segments by urban/rural and builds tools using backward stepwise least-squares regression on the logarithm of aggregate per-capita household expenditure. The TA-level indicators include:

- Average household size
- Average age
- Birth rate in the past twelve months
- Average highest education attained by households
- Percentage of households with a female head
- Percentage of children aged 6 to 13 who are enrolled in primary school
- Percentage of the economically active population who work in the tertiary sector
- Percentage of people aged 10 or above who are economically active
- Percentage of people aged 20 or younger who have lost one or both parents
- Percentage of people living in households that own a bicycle
- Percentage of people living in housing made of traditional materials
- Percentage of people using paraffin for lighting
- Percentage of people getting water from a protected source in the dry season

Benson tests targeting accuracy by comparing how well the TA-indicator tool assigns TAs to the same quintile rank as does poverty mapping. Not surprisingly, the TA-indicator tool does not perfectly reproduce poverty-mapping's rankings, especially in the middle three quintiles. Of course, if the goal is targeting, then it is simpler and clearer to measure accuracy via an ROC curve (Wodon, 1997) or, as this paper does, the ROC curve's tabular equivalent (columns 2 and 4 in Figure 12).

Overall, Benson resembles this paper in that it seeks a simple, low-cost way to measure poverty for targeting. It differs in its focus on the TA instead of the household.

---

finer levels of spatial disaggregation than would be possible with only data from the expenditure survey (Demombynes *et al.*, 2002).

9.2 Doctor

Doctor (2004) uses Malawi's 1987 and 1998 censuses with Principal Components Analysis (PCA) to check how a "living standards index" (segmented by urban/rural) to is associated with child mortality. Doctor finds that poorer households have higher child mortality in 1987, but that richer households have higher mortality in 1998, possibly due to higher HIV prevalence among richer households.

Derived from census data, Doctor's indicators—like those here—are simple, quick-to-collect, and verifiable:

- Characteristics of the residence:
    - Source of water
    - Type of roof
    - Type of wall
    - Type of floor
    - Type of toilet arrangement
- Ownership of consumer durables:
    - Radio
    - Bicycle
    - Motorbike
    - Motor vehicle
- Sector of occupation of household head
- Education of household head
- Source of energy for:
    - Cooking
    - Lighting

Doctor's PCA-based index is close kin to the scorecard here except that, because the censuses do not measure expenditure, the index may not be closely correlated with expenditure-based poverty status (Howe *et al.*, 2009).[18] Examples of the PCA-index approach include Stifel and Christiaensen (2007), Zeller *et al.* (2006), Sahn and Stifle (2003 and 2000), and Filmer and Pritchett (2001).

The advantage of PCA-based indices is that, because they do not require expenditure data, they can be applied to a wide array of "light" surveys such as censuses, Demographic and Health Surveys, Welfare Monitoring Surveys, and Core Welfare Indicator Questionnaires. Of course, the flip side is that, without expenditure data, they are forced to use a different definition of poverty, a definition based on the indicators in the PCA index. Thus, while PCA-based indices can be used for targeting, they cannot estimate expenditure-based poverty likelihoods for households or poverty rates for groups. Doctor does not report targeting accuracy for his PCA-based index because his goal is to relate living standards with child mortality, not to provide a way to help target services to households with high risk of child mortality.

9.3 Morris *et al.*

Morris *et al.* (1999) use 1998 data on 707 rural households in central Malawi to test an approach to poverty assessment that measures "socioeconomic position" inexpensively enough to be included in health surveys and epidemiological studies.

[18] Still, because their indicators are so similar, the PCA-based index and expenditure-based poverty-assessment tools may pick up the same underlying construct (such as "permanent income", see Bollen, Glanville, and Stecklov, 2007) and rank households much the same. Filmer and Pritchett (2001) and Montgomery *et al.* (2000) test how well PCA-based indices predict expenditure.

They report that their indicators cover 22 assets and nine types of livestock. Each indicator's value is defined as the number of the item that the household owns. Each indicator's points are defined as the reciprocal of the share of households that own the item, so rarer items get more points. (For example, if one-third of households own a bicycle, then each bicycle owned gets $1 \div (1 \div 3) = 3$ points.) The total index is the logarithm of the sum of each indicator multiplied by its points.

Morris *et al.* define *socioeconomic status* as the logarithm of the total value of household assets. They then measure accuracy as the correlation coefficient between the total index and their measure of socioeconomic status.

The scorecard here differs from Morris *et al.* in several ways. First, the scorecard here has a directly practical purpose: to help local, pro-poor programs in Malawi improve their service quality and outreach to the poor. In contrast, Morris *et al.* have purely methodological aims; indeed, they do not report indicators or points.

Second, the new scorecard here is based on a nationally representative database that is newer and larger.

Third, the new scorecard defines socioeconomic status as whether per-capita household expenditure is below a given poverty line. This is more common than the logarithm of the value of household assets.

Fourth, the new scorecard produces poverty likelihoods that have absolute units, whereas scores from Morris *et al.* have relative units. Furthermore, poverty likelihoods can be used not only as controls in epidemiological regressions but also for targeting and for estimating groups' poverty rates and their changes over time.

Fifth, the new scorecard is tested on data that is not used in its construction. In contrast, Morris *et al.* build and test their tool with the same data, which overstates accuracy. Beyond correlation coefficients, this paper reports differences between estimates and true values, precision, and standard-error formulas.

Sixth, the new scorecard is less costly than Morris *et al.* (10 indicators versus about 30) and simpler for non-specialists to understand (no reciprocals or logarithms).

9.4 Gwatkin *et al.*

Gwatkin *et al.* (2007) apply to Malawi an approach used by USAID in 56 countries with Demographic and Health Surveys (Rutstein and Johnson, 2004). Like Doctor, Gwatkin *et al.* use PCA to make a "wealth index" from simple, low-cost indicators available for the 62,398 households in Malawi's 2000 DHS. The PCA index is like the scorecard here except that, because the DHS does not collect data on income or expenditure, it is based on a different conception of poverty, its accuracy vis-à-vis expenditure-based poverty is unknown, and it may not be a good proxy for long-term wealth/economic status.

The 12 indicators in Gwatkin *et al.* are similar in spirit to those here:

- Characteristics of the residence:
    — Presence of electricity
    — Source of drinking water
    — Type of fuel for cooking
    — Type of toilet arrangement
    — Type of floor
- Presence of domestic servants
- Whether household members work their own or their family's agricultural land
- Ownership of consumer durables:
    — Radio
    — Television
    — Bicycle
    — Motorcycle or scooter
    — Car or truck

Gwatkin *et al.* have three basic goals for their wealth index:[19]

- Segment people by quintiles in order to see how health, population, and nutrition vary with socio-economic status
- Monitor (via exit surveys) how well health-service points reach the poor
- Measure coverage of services via small-scale local surveys

These last two goals are the same as the monitoring and targeting goals here, and the first goal of ranking household be quintiles is akin to targeting. As here, Gwatkin *et al.* present the index in a format that could be photocopied and taken to the field, although theirs is more difficult to use because there are 76 point values, all of which have five decimal places and half of which are negative.

[19] NSO (2008) also uses a PCA-based index with similar goals, but it is documented only in a footnote (p. 14), with no discussion of accuracy. Its indicators are "persons per sleeping room, type of floor, type of roof, type of wall, type of cooking fuel, and other types of assets".

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

9.5 Howe, Hargreaves, and Huttly

Howe, Hargreaves, and Huttly ("HHH", 2008) focus on methods, asking whether PCA-based indices (like those in Doctor, Morris *et al.*, and Gwatkin *et al.*) are the best approach to ranking (targeting) households by socio-economic status. HHH use Malawi's 2004/5 IHS to build and test six types of poverty-assessment tools:

- PCA-based index using the indicators in Gwatkin *et al.*
- PCA-based index using dichotomized versions of the indicators in Gwatkin *et al.*
- Equal (0/1) points using dichotomized indicators
- Points as the inverse proportion of ownership rates, as in Morris *et al.*
- Index based on Multiple Correspondence Analysis (MCA is like PCA, but explicitly accounts for the categorical nature of the indicators)

In theory, MCA should produce a better index than PCA, although in practice, MCA is almost never used. The appeal of the indices with dichotomized indicators is that they are simpler than the (already simple) indicators in Gwatkin *et al.* and here.

For each of the five approaches, HHH rank households by quintiles and then compare the extent of agreement with quintile ranks based on expenditure. The five approaches all have about the same targeting capacity. HHH conclude that "PCA appears to offer little advantage over the simpler, more easily understood methods, nor over the more statistically appropriate MCA". Still, considering factors beyond targeting accuracy, they say "there seems to be little reason to adopt any of the alternatives."

Indeed, accuracy's being more or less constant across approaches is such a common result in the predictive-modeling literature that is has a name, the "flat maximum". This is why this paper can present a new method whose strengths are transparency and simplicity, confident that the cost in terms of accuracy is low.

9.6 Mukherjee and Benson

Mukherjee and Benson (2003) use the 1997/8 IHS to construct a poverty-assessment tool "to assess the likely impact on poverty of a number of poverty-reduction policy interventions" (p. 339). A distinguishing feature is that they use only indicators that affect current poverty but that are not affected by current poverty:

- Household demographics:
  - — Age of household head
  - — Sex of household head
  - — Number of members ages 9 or younger
  - — Number of members ages 10 to 17
  - — Number of women aged 18 to 59
  - — Number of men aged 18 to 59
  - — Number of members aged 60 or older
  - — Number of members of all ages squared
- Education:
  - — Highest education for any member aged 20 or older
  - — Number of men ages 20 to 59 with Junior Secondary School Qualification
  - — Number of women ages 20 to 59 with Junior Secondary School Qualification
  - — Number of men ages 20 to 59 with Senior Secondary School Qualification
  - — Number of women ages 20 to 59 with Senior Secondary School Qualification

- Sector of economic activity (number of members):
  - — Primary
  - — Secondary
  - — Tertiary
  - — Formal employment
- Agricultural assets:
  - — Logarithm of the per-capita value of livestock
  - — Acres cultivated per capita
  - — Whether the household cultivates tobacco
  - — Number of non-tobacco, non-maize crops cultivated
- Community characteristics:
  - — Average maize yield
  - — District
  - — Interaction of district with average maize yield
  - — Availability of agricultural inputs
  - — Availability of electricity
  - — Availability of a public-works program
  - — Mean time to travel to health center, bus stage, Agricultural Development Marketing Corporation depot, bank, and post office

Mukherjee and Benson use least-squares regression on the logarithm of per-capita household expenditure to build four regional tools. Because their indicators are causes of current poverty but are not caused by current poverty, Mukherjee and Benson argue that their tools can be used to simulate the effects of policies that could change the indicators. For example, removing one child from all households with children 9-years-old or younger (simulating the possible effects of a family-planning policy) would reduce the poverty rate (national line) by 23.1 percent in urban areas and by 12.5 to 15.0 percent in rural areas.

Likewise, Mukherjee and Benson find it “very encouraging” (p. 353) that increasing by one the number of women with the Senior Secondary School Qualification (in households with at least one adult woman) would reduce poverty rates by 28.1 percent in urban areas and by 5.3 to 10.1 percent in rural areas.

Of course, Malawi's government probably already knows that smaller families and greater education for girls would mean large reductions in poverty. The contribution of Mukherjee and Benson is to quantify the magnitude of the reductions. Still, poverty reduction in practice is usually constrained not by technical knowledge of poverty drivers but rather by political, financial, and organizational factors. Why would the people in Malawi's government prioritize poverty reduction? How could they fund a family-planning campaign or secondary education for more girls? How could they design effective policies and then implement them effectively?

Thus, the scorecard here differs from Mukherjee and Benson chiefly in focus. Rather than seek to identify poverty drivers, this paper seeks to identify poor households, both for targeting and for monitoring. Rather than identify promising policies, it aims to help implement a given pro-poor policy effectively. Because the scorecard here aims to be applied thousands of times by low-level field agents rather than once by high-level researchers, it tries to keep costs low.

For these reasons, the scorecard here has fewer indicators (10 versus 28), only household indicators (excluding community indicators), and only simple indicators (omitting complex indicators such as the logarithm of the per-capita value of livestock or acres cultivated per capita). Finally, because the scorecard here is not concerned with counterfactual cause-and-effect, its accuracy can be measured.

9.7 World Bank

World Bank (2007) closely follows Mukherjee and Benson in aims and methods, but it uses the 2004/5 IHS instead of the 1997/8 IHS. Indicators are again selected for being causes of poverty while not being short-term effects of poverty:

- Demographics:
  - — Sex of the household head
  - — Age of the household head
  - — Widowhood of household head
  - — Household size and its square
  - — Number of children:
    - 0 to 4
    - 5 to 10
    - 11 to 14
- Highest education attained by a household member
- Whether the household head has formal wage employment
- Whether the household has a non-farm enterprise
- Agriculture:
  - — Presence of rain-fed plots
  - — Logarithm of hectares of rain-fed agricultural land
  - — Ownership of a *dimba* plot
  - — Whether the household head grew tobacco in the past season
- Community characteristics:
  - — Presence of regular bus service
  - — Presence of a health clinic
  - — Presence of a bank branch
  - — Presence of a daily market
  - — Presence of an ADMARC market
  - — Residence in a *boma* or trading center
  - — Distance to the nearest *boma*
  - — Presence of a tarmac/asphalt road
  - — Region (North, Center, South)

Like Mukherjee and Benson, World Bank focuses on estimating coefficients, not on which indicators might be malleable by policy, nor how regression estimates could be combined with estimates of costs to help prioritize policies.

Because the 1997/8 IHS and the 2004/5 IHS measure expenditure differently, World Bank also constructs a second poverty-assessment tool to estimate comparable poverty rates for the two surveys. This matches one of the purposes of the scorecard here and resembles Mathiassen's (2006) use of the 1997/8 IHS and 2005 Welfare Measurement Survey to estimate changes in poverty rates (see below).

To do this, World Bank first selects indicators available in both the 1997/8 IHS and 2004/5 IHS. A tool is then built using least-squares on 2004/5 HIS data for the logarithm of per-capita household expenditure. (World Bank does not report this tool's indicators or points.) This tool is then applied to the 1997/8 IHS to estimate Malawi's poverty rate in 1997/8. Given that the two surveys measure expenditure differently, it is not possible to check the accuracy of this, although it could have been tested on the 2004/5 IHS (as here) by using construction and validation samples.

9.8 IRIS Center

USAID commissioned IRIS Center ("IRIS", 2007a) to build a poverty-assessment tool so that USAID's microenterprise partners in Malawi can report on their participants' poverty rates. The tool is based on the $1.25/day 2005 PPP poverty line. Overall, the IRIS tool is like the one here, except it is less transparent, it uses more indicators (19 versus 10), and some aspects of accuracy are not reported.

Like this paper, IRIS uses the 2004/5 IHS. After comparing several statistical approaches, IRIS settles on a quantile regression (Koenker and Hallock, 2001) that estimates not poverty likelihoods but rather the 41st percentile of the logarithm of per-capita household expenditure. Unlike the non-parametric, poverty-likelihood approach here, IRIS' estimator of poverty rates is non-linear in estimated expenditure and so is biased. Its 19 indicators are:

- Demographics:
  - Household size
  - Age of head
  - Marital status of the head
- Education of members 5-years-old and older:
  - Number who can read
  - Number who never attended school or report "no education"
- Characteristics of the residence:
  - Type of floor
  - Number of rooms
  - Presence of electricity
- Ownership of consumer durables:
  - Bed
  - Iron
  - Tape player, CD player, or HiFi
  - Refrigerator
  - Bicycle
  - Car
- Purchases by household members in the past month:
  - Bar soap, for body or clothes
  - Powdered soap for clothes
- Whether the household cultivated a *dimba* garden in the last completed dry season
- Location:
  - Region
  - Urban/rural

Except for past purchases of soap, these indicators are simple, inexpensive, and verifiable. IRIS reports only the questionnaire used to collect data and not the actual tool or its points, so actual indicators may differ slightly from those listed here.

IRIS' accuracy tests focus on the difference between the estimated poverty rate and its true value. Unlike this paper, they do not report confidence levels and confidence intervals for this difference,[20] nor do they report standard-error formula.

IRIS doubts that their tool is useful for measuring changes in poverty rates, noting that "it is unclear that the tools will be able to identify real changes in poverty over time due to their inherent measurement errors. Unless the changes in the poverty rate are exceptionally large and the tools exceptionally accurate, the changes identified are likely to be contained within the margin of error."[21] Of course, this statement would be easier to evaluate if IRIS reported margins of error for estimates of change.

IRIS also states that its tool should not be used for targeting.[22] Nevertheless, IRIS reports measures of targeting accuracy.

IRIS' preferred measure of accuracy is the "Balanced Poverty Accuracy Criterion" (BPAC), and USAID adopted BPAC as its criterion for certifying poverty-assessment tools. BPAC is designed to consider accuracy both in terms of the estimated poverty rate and in terms of inclusion, that is, successful classification of households below the poverty line (IRIS Center, 2005). The BPAC formula is:

BPAC = (Inclusion – |Undercoverage – Leakage|) x [100 ÷ (Inclusion+Undercoverage)].

BPAC for IRIS for the $1.25/day 2005 PPP line is 85.5 (IRIS, 2008), while for the scorecard here, $1.25/day 2005 PPP, and a cut-off of 40–44, BPAC is 80.0 (Figure

[20] Anthony Leegwater did provide them in personal communication, November 26, 2008
[21] `http://www.povertytools.org/faq/faq.html#12`, retrieved 19 February 2009.
[22] `http://www.povertytools.org/faq/faq.html#11`, retrieved 19 February 2009.

12). If scores are not grouped, then the scorecard here has a BPAC of 84.1 for a cut-off of 44 or less. In terms of BPAC, the scorecard here is about accurate as IRIS.

9.9 Matthiassen

The approach here and in Mathiassen (2006) are similar. Both use the 2004/5 IHS to build simple, inexpensive poverty-assessment tools with the explicit goal of measuring poverty rates at a point in time.[23] Both estimate poverty rates as the average of the individual households' poverty likelihoods.[24] And both divide the 2004/5 IHS into two sub-samples, one for tool construction and one for testing accuracy.[25]

Before Malawi's 2005 Welfare Measurement Survey (WMS) was designed, Matthiassen built a 29-indicator tool based on the 2004/5 IHS. These 29 indicators were then put in the WMS with the express purpose of using a poverty-assessment tool to estimate Malawi's poverty rate in 2005 without incurring the cost of measuring expenditure. Thus, Matthiassen is a rare example of poverty-assessment tools' being used for one of their most-commonly proposed purposes: to update poverty estimates

[23] This paper also seeks to measure changes in poverty rates and to provide a tool for targeting. Matthiassen's tool could also be used for these purposes, although she does not discuss them.

[24] Matthiassen estimates poverty likelihoods even though she builds her tool not with Logit regression on poverty status (as here) but rather with least-squares on the log of per-capita household expenditure. Matthiassen's approach is a correct, valid, and sometimes preferred alternative to Logit (Coudouel, Hentschel, and Wodon, 2002).

[25] Both papers report, across many simulated or bootstrapped samples, average differences between estimated and true poverty rates. Unlike Matthiassen, this paper also reports confidence intervals.

between expenditure surveys using "light" surveys. Matthiassen estimates that the share of Malawi's population with expenditure under the national line fell from 52 percent in 2004 to 50 percent in 2005.

The main difference between Matthiassen and this paper is the indicators. This paper uses 10, all of them verifiable, whereas Matthiassen uses 29, 14 of which deal with consumption or past spending and thus are not verifiable, and two more of which involve ratios and thus would be difficult to calculate on paper in the field:

- Demographics:
  - Household size
  - Dependency ratio (number younger than 15 or older than 60 divided by household size)
  - Crowding ratio (Household size divided by number of rooms)
  - Number of members younger than 15
  - Age of household head
- Highest education for a household member
- Characteristics of the residence:
  - Type of roof
  - Type of floor
- Ownership of consumer durables:
  - Whether the household head sleeps under sheets
  - Number of changes of clothes for the household head
  - Number of radios
  - Bed
  - Iron
  - Refrigerator
  - Mobile telephone
- Whether the household used/consumed an item in an undocumented past period:
  - Transport
  - Eggs
  - Meat
  - Rice
  - Bread
  - Fresh milk
  - Cooking oil
  - Sugar

    — Toothpaste

- Whether the household had expenses for an item in an undocumented past period:
    - — Cooking oil
    - — Sugar
    - — Bar soap
- Whether the household purchased clothing items in the past three months:
    - — Men's clothing
    - — Shoes

In sum, the approach here and in Matthiassen are similar in construction and testing. The scorecard here, however, uses fewer indicators and only verifiable indicators, so it is less expensive to use and more difficult to game.

9.10 Benson *et al.*

Benson *et al.* (2006) resembles this paper even more than Mathiassen does. In particular, Benson *et al.* seek "simple and efficient assessment methods . . . for identifying the poor for targeting services and for the timely monitoring of poverty levels" (p. 1). They also discuss using their tool to measure change over time, thus matching all three goals of this paper. Also like this paper, they divide their data into construction and validation samples, thus obtaining accurate measures of accuracy. Finally, Benson *et al.* also report targeting accuracy and compare true versus estimated poverty rates, as well as standard errors for the differences.

The main contrasts between this paper and Benson *et al.* is that they:

- Use the 1997/8 IHS (versus the 2004/5 IHS)
- Segment their tools by urban/rural (versus no segmenting)
- Use 17 indicators for rural and 7 for urban (versus 10 for all-Malawi)
- Estimate per-capita daily consumption directly via a stepwise least-squares regression (versus poverty likelihoods from a Logit regression based on accuracy as well as non-statistical practicality criteria)
- Report results at the person-level (versus the household-level)
- Judge the tools as effective or not (versus simply reporting accuracy)
- Use two indicators that are complex or not verifiable (versus all simple and verifiable)

As noted earlier, the scorecard here is not segmented because tests elsewhere find that segmentation does not improve targeting accuracy much. Segmentation may affect, however, the accuracy of estimated poverty rates.

This paper estimates poverty likelihoods, rather than expenditure, because the poverty-likelihood approach makes explicit and transparent the error inherent in any estimate. Furthermore, the non-parametric poverty likelihood approach here produces unbiased estimates of poverty rates, whereas poverty rates derived from estimates of expenditure are biased.[26] Benson *et al.* note that converting their estimates of expenditure to poverty likelihoods (as recommended by Coudouel, Hentschel, and Wodon, 2002) might reduce this bias.[27]

This paper reports household-level results because, in practice, most targeting is at the household level. Like most other poverty-assessment tools for all other countries, Benson *et al.* report person-level results. Of course, estimates of poverty rates at either level can be converted to the other by weighting (or unweighting) by household size.

This paper reports accuracy without judging the scorecard effective or ineffective. This is because such judgments—if they are to be informative—require an explicit benchmark or objective function. In any case, for estimating poverty rates, the absolute differences between estimated and true values are smaller here (0.2 percentage points, averaged across a single scorecard applied to seven lines, with a maximum absolute difference of 0.5 percentage points) than in Benson *et al.* (4.9 percentage points, averaged across two segments and two poverty lines, with a maximum of 12.3 percentage points and a minimum of 0.7 percentage points). Of course, these differences may be due to variation in data quality across the 1997/8 IHS and 2004/5 IHS. In

---

[26] This is because the function converting the estimate of expenditure to poor/non-poor status is non-linear.

[27] Still, because this conversion assumes a parametric distribution for poverty likelihoods and because estimates of poverty rates are again non-linear functions of estimated expenditure, there will still be bias and maybe even more bias.

terms of targeting accuracy, direct comparisons are not possible because of differences in data sets, poverty rates, and reporting levels.

Finally, Benson *et al.* use indicators that resemble those here in terms of simplicity and verifiability, save for two that are non-verifiable or difficult to report:

- Demographics:
    - — Household size
    - — Household size squared
- Education of the household head
- Number of salaried household members
- Fuel use:
    - — Collected firewood for cooking
    - — Gas or electricity for lighting
- Ownership of consumer durables:
    - — Bicycle
    - — Motorcycle or car
    - — Refrigerator
    - — Bed
- Agricultural assets:
    - — Acres cultivated
    - — Whether cultivates tobacco
    - — Whether cultivates hybrid maize
    - — Number of cattle
- District of residence
- Reported having purchased sugar in the past two weeks

In particular, it is difficult to verify the recent purchase of sugar, and households sometimes cannot easily report acres cultivated. Apart from that, Benson *et al.* believes that (in contrast to this paper) none of their indicators can be corroborated through direct observation. Experience with similar tools for other countries, however, suggests that these sorts of indicators are indeed verifiable through home visits.

Overall, Benson *et al.*, like Matthiassen, is quite similar to the scorecard here.

10. Conclusion

This paper presents the Simple Poverty Scorecard®, an easy-to-use tool that pro-poor organizations in Malawi can use 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 in order to speed up their participants' progress out of poverty.

The scorecard is built with a sub-sample of data from the 2004/5 IHS, tested with a different sub-sample, and calibrated to seven poverty lines.

Accuracy and formulas for standard errors 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 in poverty rates are not the same as estimates of program impact. Targeting accuracy is also reported.

When the scorecard is applied to the validation sample, the absolute difference between estimates versus true poverty rates for groups of households at a point in time is 0.5 percentage points or less and averages—across the seven poverty lines—about 0.2 percentage points. For $n = 16{,}384$ and 90-percent confidence, the precision of these differences is ±0.6 percentage points or better, and for $n = 1{,}024$, precision is ±2.2 percentage points or better.

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

Although the statistical technique is innovative, and although technical accuracy is important, the design of the scorecard here focuses on transparency and ease-of-use. After all, a perfectly accurate scorecard is worthless if programs feel so daunted by its complexity or its cost that they do not even try to use it. For this reason, the scorecard is kept simple, using 10 indicators that are inexpensive to collect and that are straightforward to verify. Points are all zeros or positive integers, and scores range from 0 to 100. Scores are related to poverty likelihoods via simple look-up tables, and targeting cut-offs are likewise simple to apply. The design attempts to facilitate adoption by helping managers understand and trust scoring and by allowing non-specialists to generate scores quickly in the field.

In sum, the Simple Poverty Scorecard® is a practical, objective way for pro-poor programs in Malawi to estimate poverty rates, track changes in poverty rates over time, and target services. The same approach can be applied to any country with similar data 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.

Benson, Todd. (2002) "Correlates of Poverty Measures in Malawi: An Application of Poverty Mapping at the Sub-District Scale", `povertymap.net/publications/doc/2002CSAEconf_Benson_Malawi_paper.pdf`, retrieved 19 February 2009.

_____; Payongayong, Ellen; Ahmed, Akhter U.; and Manohar Sharma. (2006) "Simple Household Poverty Assessment Models for Malawi", `t.benson@cgiar.org`.

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

Caire, Dean. (2004) "Building Credit Scorecards for Small-Business Lending in Developing Markets", `microfinance.com/English/Papers/Scoring_SMEs_Hybrid.pdf`, retrieved 1 February 2011.

Chen, Shiyuan; and Mark Schreiner. (2009a) "Simple Poverty Scorecard®: Bangladesh", `SimplePovertyScorecard.com/BGD_2005_ENG.pdf`, accessed 31 January 2011.

______. (2009b) "Simple Poverty Scorecard®: Vietnam", `SimplePovertyScorecard.com/VNM_2006_ENG.pdf`, accessed 31 January 2011.

Coady, David; Grosh, Margaret; and John Hoddinott. (2004) *Targeting of Transfers in Developing Countries*, `hdl.handle.net/10986/14902`, retrieved 13 May 2016.

Cochran, William G. (1977) *Sampling Techniques, Third Edition.*

Coudouel, Aline; Hentschel, Jesko S.; and Quentin T. Wodon. (2002) "Poverty Measurement and Analysis", pp. 27–74 in *The PRSP Sourcebook.*

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

Demombynes, Gabriel; Elbers, Chris; Lanjouw, Jenny; Lanjouw, Peter; Mistiaen, Johan; and Berk Ozler. (2002) "Producing an Improved Geographic Profile of Poverty: Methodology and Evidence from Three Developing Countries", World Institute for Development Economics Research Discussion Paper No. 2002/39, siteresources.worldbank.org/INTPGI/Resources/342674-1092157888460/21888_Wider_Poverty_(DP39,_March_2002).pdf, retrieved 1 February 2011.

Doctor, Henry V. (2004) "The Effect of Living Standards on Childhood Mortality in Malawi", *African Population Studies*, Vol. 19, Supplement A, pp. 249–263, tspace.library.utoronto.ca/bitstream/1807/5820/1/ep04029.pdf, retrieved 1 February 2011.

Efron, Bradley; and Robert J. Tibshirani. (1993) *An Introduction to the Bootstrap*.

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

Falkenstein, Eric. (2008) "DefProb™: A Corporate Probability-of-Default Model", papers.ssrn.com/sol3/papers.cfm?abstract_id=1103404, retrieved 1 February 2011.

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

Fuller, Rob. (2006) "Measuring the Poverty of Microfinance Clients in Haiti", microfinance.com/English/Papers/Scoring_Poverty_Haiti_Fuller.pdf, retrieved 1 February 2011.

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

Grootaert, Christiaan; and Jeanine Braithwaite. (1998) "Poverty Correlates and Indicator-Based Targeting in Eastern Europe and the Former Soviet Union", World Bank Policy Research Working Paper No. 1942, go.worldbank.org/VPMWVLU8E0, retrieved 1 February 2011.

Grosh, Margaret; and Judy L. Baker. (1995) "Proxy Means Tests for Targeting Social Programs: Simulations and Speculation", LSMS Working Paper No. 118, go.worldbank.org/W9OWN57PD0, retrieved 1 February 2011.

Gwatkin, Davidson R.; Rutstein, Shea; Johnson, Kiersten; Suliman, Eldaw; Wagstaff, Adam; and Agbessi Amouzou. (2007) "Socio-Economic Differences in Health, Nutrition, and Population: Malawi", Country Reports on HNP and Poverty, go.worldbank.org/T6LCN5A340, retrieved 1 February 2011.

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

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

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

_____.; Hargreaves, James R.; and Sharon R.A. Huttly. (2008) "Issues in the Construction of Wealth Indices for the Measurement of Socio-Economic Position in Low-Income Countries", *Emerging Themes in Epidemiology*, Vol. 5, No. 3.

IRIS Center. (2008) "Accuracy Results for 20 Poverty Assessment Tool Countries", povertytools.org/other_documents/PAT_20_country_accuracy_results_Aug2008.pdf, retrieved 1 February 2011.

_____. (2007a) "Client Assessment Survey—Malawi", povertytools.org/USAID_documents/Tools/Current_Tools/USAID_PAT_Mwi_08_30_2010.xls, retrieved 1 February 2011.

_____. (2007b) "Manual for the Implementation of USAID Poverty Assessment Tools", povertytools.org/training_documents/Manuals/USAID_PAT_Manual_Eng.pdf, retrieved 1 February 2011.

_____. (2007c) "Introduction to Sampling for the Implementation of PATs", povertytools.org/training_documents/Sampling/Introduction_Sampling.ppt, retrieved 1 February 2011.

_____. (2005) "Notes on Assessment and Improvement of Tool Accuracy", povertytools.org/other_documents/AssessingImproving_Accuracy.pdf, retrieved 1 February 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, retrieved 1 February 2011.

Koenker, Roger; and Kevin F. Hallock. (2001) "Quantile Regression", *Journal of Economic Perspectives*, Vol. 15, No. 4, pp. 143–156.

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", `ciep.itam.mx/~martinel/lies4.pdf`, retrieved 1 February 2011.

Mathiassen, Astrid. (2006) "Predicting the Poverty Headcount Ratio Based on IHS2 and WMS Data", pp. 106–108 in National Statistical Office, *Welfare Monitoring Survey 2005*, `www.nso.malawi.net/data_on_line/agriculture/wms_2006/WMS2006_Final%20Report.pdf`, retrieved 1 February 2011.

Matul, Michal; and Sean Kline. (2003) "Scoring Change: Prizma's Approach to Assessing Poverty", Microfinance Centre for Central and Eastern Europe and the New Independent States Spotlight Note No. 4, `www.mfc.org.pl/doc/Research/ImpAct/SN/MFC_SN04_eng.pdf`, retrieved 1 February 2011.

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

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

Morris, Saul; Carletto, Calogero; Hoddinott, John; and Luc J.M. Christiaensen. (2000) "Validity of Rapid Estimates of Household Wealth and Income for Health Surveys in Rural Africa", *Journal of Epidemiology and Community Health*, Vol. 54, pp. 381–387.

Mukherjee, Sanjukta; and Todd Benson. (2002) "The Determinants of Poverty in Malawi, 1998", *World Development*, Vol. 31, No. 2, pp. 339–358.

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

Narayan, Ambar; and Nobuo Yoshida. (2005) "Proxy Means Tests for Targeting Welfare Benefits in Sri Lanka", World Bank Report No. SASPR–7, documents.worldbank.org/curated/en/2005/07/6209268/proxy-means-test-targeting-welfare-benefits-sri-lanka, retrieved 5 May 2016.

National Statistical Office. (2008) "Multiple Indicator Cluster Survey 2006: Malawi", www.nso.malawi.net/data_on_line/demography/mics/MICS%20Report.pdf, retrieved 1 February 2011.

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

Rutstein, Shea Oscar; and Kiersten Johnson. (2004) "The DHS Wealth Index", DHS Comparative Reports No. 6, measuredhs.com/pubs/pdf/CR6/CR6.pdf, retrieved 1 February, 2011.

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

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

SAS Institute Inc. (2004) "The LOGISTIC Procedure: Rank Correlation of Observed Responses and Predicted Probabilities", in *SAS/STAT User's Guide, Version 9*, support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_logistic_sect035.htm, retrieved 1 February 2011.

Schreiner, Mark. (2010) "Simple Poverty Scorecard®: Honduras", SimplePovertyScorecard.com/HND_2007_ENG.pdf, retrieved 1 February 2011.

_____. (2009a) "Simple Poverty Scorecard®: Peru", SimplePovertyScorecard.com/PER_2007_ENG.pdf, retrieved 1 February 2011.

_____. (2009b) "Simple Poverty Scorecard®: Philippines", SimplePovertyScorecard.com/PHL_2002_ENG.pdf, retrieved 1 February 2011.

_____. (2009c) "Simple Poverty Scorecard®: Pakistan", SimplePovertyScorecard.com/PAK_2005_ENG.pdf, retrieved 1 February 2011.

_____. (2009d) "Simple Poverty Scorecard®: Bolivia", SimplePovertyScorecard.com/BOL_2007_ENG.pdf, retrieved 1 February 2011.

_____. (2009e) "Simple Poverty Scorecard®: Mexico", SimplePovertyScorecard.com/MEX_2008_ENG.pdf, retrieved 1 February 2011.

_____. (2008a) "Simple Poverty Scorecard®: Peru", SimplePovertyScorecard.com/PER_2003_ENG.pdf, retrieved 1 February 2011.

_____. (2008b) "Simple Poverty Scorecard®: India", SimplePovertyScorecard.com/IND_2005_ENG.pdf, retrieved 1 February 2011.

_____. (2006) "Is One Simple Poverty Scorecard® Enough for India?", microfinance.com/English/Papers/Scoring_Poverty_India_Segments.pdf, retrieved 1 February 2011.

_____. (2005a) "Índice de Calificación de la Pobreza™: México", SimplePovertyScorecard.com/MEX_2002_SPA.pdf, retrieved 1 February 2011.

_____. (2005b) "IRIS Questions on the Simple Poverty Scorecard®", microfinance.com/English/Papers/Scoring_Poverty_Response_to_IRIS.pdf, retrieved 1 February 2011.

_____. (2002) *Scoring: The Next Breakthrough in Microfinance?* CGAP Occasional Paper No. 7, microfinance.com/English/Papers/Scoring_Breakthrough_CGAP.pdf, retrieved 13 May 2016.

_____; Matul, Michal; Pawlak, Ewa; and Sean Kline. (2004) "The Simple Poverty Scorecard®: Lessons from a Microlender in Bosnia-Herzegovina", microfinance.com/English/Papers/Scoring_Poverty_in_BiH_Short.pdf, retrieved 1 February 2011.

_____; and Gary Woller. (2010a) "Simple Poverty Scorecard®: Ghana", SimplePovertyScorecard.com/GHA_2005_ENG.pdf, retrieved 1 February 2011.

_____; and Gary Woller. (2010b) "Simple Poverty Scorecard®: Guatemala", SimplePovertyScorecard.com/GTM_2006_ENG.pdf, retrieved 1 February 2011.

Sillers, Don. (2006) "National and International Poverty Lines: An Overview", pdf.usaid.gov/pdf_docs/Pnadh069.pdf, retrieved 13 May 2016.

Stifel, David; and Luc Christiaensen. (2007) “Tracking Poverty over Time in the Absence of Comparable Consumption Data”, *World Bank Economic Review*, Vol. 21, No. 2, pp. 317–341.

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) “Progress Out of Poverty Index: Training Guide”, `progressoutofpoverty.org/toolkit`, retrieved 1 February 2011.

United States Congress. (2004) “Microenterprise Results and Accountability Act of 2004 (HR 3818 RDS)”, November 20, `smith4nj.com/laws/108-484.pdf`, retrieved 13 May 2016.

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

Wodon, Quentin T. (1997) “Targeting the Poor Using ROC Curves”, *World Development*, Vol. 25, No. 12, pp. 2083–2092.

World Bank. (2008) “International Comparison Project: Tables of Results”, `siteresources.worldbank.org/ICPINT/Resources/icp-final-tables.pdf`, retrieved 1 February 2011.

_____. (2007) “Malawi Poverty and Vulnerability Assessment Investing in Our Future, Volume I: Synthesis of Main Findings and Recommendations” and “Volume II: Detailed Background Papers” and “Appendices”, World Bank Report No. 36546-MW, `www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2008/10/03/000333038_20081003000558/Rendered/PDF/365460vol020ES1disclosed09130120081.pdf`, retrieved 1 February 2011.

_____. (2005) “Note on Construction of Expenditure Aggregate and Poverty Lines for IHS2”, `worldbank.org/html/prdph/lsms/country/malawi04/docs/ihs2agg.pdf`, retrieved 19 February 2009.

Zeller, Manfred. (2004) "Review of Poverty Assessment Tools", `pdf.usaid.gov/pdf_docs/PNADH120.pdf`, retrieved 1 February 2011.

_____; Sharma, Manohar; Henry, Carla; and Cécile Lapenu. (2006) "An Operational Method for Assessing the Poverty Outreach Performance of Development Policies and Projects: Results of Case Studies in Africa, Asia, and Latin America", *World Development*, Vol. 34, No. 3, pp. 446–464.

Appendix A:
Guide to Interpretation of Scorecard Indicators

The following information comes from:

National Statistical Office. (2004)*Enumerator Manual for Household Characteristics, Income, and Expenditure Questionnaire*, Government of Malawi.

1. <u>How many household members are 14-years-old or younger?</u>

According to pp. 7–9 of the *Enumerator's Manual*: "A household may be either a person living alone or a group of people, either related or unrelated, who live together as a single unit in the sense that they have common housekeeping arrangements (that is, share or are supported by a common budget). A standard definition of a household is "a group of people who live together, pool their money, and eat at least one meal together each day". It is possible that individuals who are not members of the household may be residing with the household at the time of the survey. In most cases, but not all, someone who does not live with the household during the survey period is not a current member of the household. The definition of who is and who is not a household member is given below.

"It is important to recognize that members of a household need not necessarily be related by blood or by marriage. On the other hand, not all those who are related and are living in the same compound or dwelling are necessarily members of the same household. Two brothers who live in the same dwelling with their own wives and children may or may not form a common housekeeping arrangement. If they do not, they should be considered separate households.

"One should make a distinction between family and household. The first reflects social relationships, blood descent, and marriage. The second is used here to identify an economic unit. While families and households are often the same, this is not necessarily the case. You must be cautious and use the criteria provided on household membership to determine which individuals make up a particular household.

"In the case of polygamous men and extended family systems, household members are distributed over two or more dwellings. If these dwelling units are in the same compound or nearby (but necessarily within the same EA) and they have a common housekeeping arrangement with a common household budget, the residents of these separate dwelling units should be treated as one household. . . .

"Those individuals who have been absent from the household for more than 9 months during the past 12 months—that is, have been resident in the household for less than 3 of the past 12 months—should not be considered household members. However, there are several exceptions to this rule:

- The individual whom household members commonly regard as the head of household should always be included as a household member, even if he or she has been absent from the household for more than 9 of the past 12 months
- Young infants less than 3 months old
- New spouses who have recently come into the household and are now residing with the household
- Household members residing in an institution elsewhere, but still dependent on the household. This principally includes boarding-school students. However, it does not include military personnel, prisoners, or other individuals who are not primarily dependent on the household for their welfare.

"It is important to highlight that non-relatives who are resident in the household for more than three months and are included in a common household keeping arrangement under the head of household are to be considered household members. However, servants, other hired workers, and lodgers (individuals who pay to reside in the dwelling of the household) should not be considered to be household members if they have their own household elsewhere which they head or upon which they are dependent.

"You should be very careful when dealing with this rather complex task of determining who should be included and who should not be included as a member of a survey household."

2. <u>How many household members worked in their main activity in the past seven days as a farmer (*mlimi*)?</u>

According to p. 33 of the *Enumerator's Manual*: "You should categorize an individual according to his or her dominant activity. In cases where this cannot be done, you should assign the individual to the activity classification category that is of most economic significance for the household."

3. <u>Can the female head/spouse read a one-page letter in any language?</u>

The *Enumerator's Manual* provides no additional information about this indicator.

4. <u>The roof of the main dwelling is predominantly made of what material?</u>

According to p. 40 of the *Enumerator's Manual*: "If two or more different types of materials are used for the [roof], report the material that is used in the majority."

5. <u>What is your main source of cooking fuel?</u>

The *Enumerator's Manual* provides no additional information about this indicator.

6. <u>What is your main source of lighting fuel?</u>

The *Enumerator's Manual* provides no additional information about this indicator.

7. <u>Does the household own any lanterns (paraffin)?</u>

The *Enumerator's Manual* provides no additional information about this indicator.

8. <u>Does the household own any bicycles, motorcycles/scooters, cars, mini-buses, or lorries?</u>

The *Enumerator's Manual* provides no additional information about this indicator.

9. <u>Does the household own any irons (for pressing clothes)?</u>

The *Enumerator's Manual* provides no additional information about this indicator.

10. <u>How many sickles does the household own?</u>

The *Enumerator's Manual* provides no additional information about this indicator.

Figure 2: Sample sizes and poverty rates at the household level and the person level for all Malawi and by region, sub-sample, and poverty line

	Level	Sample size	Poverty rates (% with expenditure below a poverty line) and poverty lines (MWK/person/day)						
					USAID	Intl. 1993 PPP		Intl. 2005 PPP	
			National	Food	'Extreme'	$1.08/day	$2.16/day	$1.25/day	$2.50/day
Poverty lines:									
All Malawi	N/A	11,280	43.92	27.25	29.75	30.12	60.25	63.60	127.20
Urban	N/A	1,440	50.04	31.05	37.30	34.33	68.65	72.47	144.95
Northern Rural	N/A	1,440	48.04	29.81	31.34	32.95	65.91	69.57	139.15
Central Rural	N/A	3,840	41.96	26.03	30.25	28.78	57.57	60.77	121.54
Southern Rural	N/A	4,560	43.00	26.68	26.77	29.50	58.99	62.27	124.55
Poverty Rates:									
All Malawi	Households	11,280	43.6	16.6	19.6	21.6	62.7	66.1	90.8
	People	N/A	52.4	22.2	26.2	28.2	71.1	74.2	94.0
Urban	Households	1,440	19.9	5.0	8.9	6.8	34.5	38.3	70.7
	People	N/A	25.4	7.5	12.6	9.9	41.9	45.9	76.9
Northern Rural	Households	1,440	46.3	18.5	20.4	23.5	65.8	68.9	93.8
	People	N/A	56.3	25.9	28.1	31.4	75.0	77.8	96.4
Central Rural	Households	3,840	38.7	12.1	18.1	16.5	60.0	64.3	92.4
	People	N/A	46.7	16.1	23.3	21.2	68.1	72.1	95.3
Southern Rural	Households	4,560	53.8	23.3	23.7	29.6	72.2	74.8	94.4
	People	N/A	64.4	31.5	32.2	39.1	81.0	83.3	96.9
Construction and calibration samples									
Selecting indicators and points, and	Households	5,645	43.6	16.6	19.5	21.6	62.8	66.0	90.7
associating scores with likelihoods	People	N/A	52.8	22.6	26.4	28.7	71.8	74.8	94.3
Validation sample									
Measuring accuracy	Households	5,635	43.6	16.6	19.6	21.5	62.5	66.2	90.9
	People	N/A	52.0	22.1	25.9	27.7	70.3	73.6	93.7
Change in poverty rate (percentage points)									
Construction/calibration to validation	Households		–0.0	–0.1	–0.1	+0.2	+0.3	–0.2	–0.3

Source: 2004/5 HIS.

Figure 3: Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (Responses ordered starting with those most strongly indicative of poverty)
1,210	How many household members are 16-years-old or younger? (Five or more; Four; Three; Two; One; None)
1,205	How many household members are 15-years-old or younger? (Five or more; Four; Three; Two; One; None)
1,197	How many household members are 17-years-old or younger? (Five or more; Four; Three; Two; One; None)
1,194	How many household members are 18-years-old or younger? (Five or more; Four; Three; Two; One; None)
1,190	How many household members are 14-years-old or younger? (Five or more; Four; Three; Two; One; None)
1,164	How many household members are 13-years-old or younger? (Four or more; Three; Two; One; None)
1,086	How many household members are 12-years-old or younger? (Four or more; Three; Two; One; None)
1,043	How many people live in the household? (Eight or more; Seven; Six; Five; Four; Three; Two; One)
1,016	How many household members are 11-years-old or younger? (Four or more; Three; Two; One; None)
902	Do all household members ages 6 to 14 currently attend school, or, if school is not now in session, did all attend school in the session just completed and plan to attend next session? (No; Yes; No one in age range)
881	Do all household members ages 6 to 13 currently attend school, or, if school is not now in session, did all attend school in the session just completed and plan to attend next session? (No; Yes; No one in age range)
878	Do all household members ages 6 to 12 currently attend school, or, if school is not now in session, did all attend school in the session just completed and plan to attend next session? (No; Yes; No one in age range)
848	Do all household members ages 6 to 15 currently attend school, or, if school is not now in session, did all attend school in the session just completed and plan to attend next session? (No; Yes; No one in age range)

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

Uncertainty coefficient	Indicator (Respsonses ordered starting with those most strongly associated with poverty)
835	Do all household members ages 6 to 11 currently attend school, or, if school is not now in session, did all attend school in the session just completed and plan to attend next session? (No; Yes; No one in age range)
828	Do all household members ages 6 to 16 currently attend school, or, if school is not now in session, did all attend school in the session just completed and plan to attend next session? (No; Yes; No one in age range)
814	What class is the female head/spouse in or what was the highest level she ever attended? (None to primary standard 2; Primary standards 3 to 7; Primary standard 8 to secondary form 1; Secondary forms 2 to 3; No female head/spouse; Secondary form 4 or higher)
736	Do all household members ages 6 to 17 currently attend school, or, if school is not now in session, did all attend school in the session just completed and plan to attend next session? (No; Yes; No one in age range)
625	Do all household members ages 6 to 18 currently attend school, or, if school is not now in session, did all attend school in the session just completed and plan to attend next session? (No; Yes; No one in age range)
609	Can the female head/spouse read a one-page letter in any language? (No; Yes; No female head/spouse)
584	What is your main source of cooking fuel? (A. Collected firewood from forest reserve, crop residue, sawdust, animal waste, or other; Collected firewood from unfarmed areas of community; Collected firewood from own woodlot, community woodlot, or other places; Purchased firewood; Paraffin, charcoal, gas, or electricity)
568	What has been the main activity of the female head/spouse during the last seven days? (Farmer (*mlimi*), or employer; Other; Self-employed; Family business worker; Employee; No female head/spouse)
524	Has the main activity of the female head/spouse during the last seven days been agriculture? (Yes; No; No female head/spouse)

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

Uncertainty coefficient	Indicator (Respsonses ordered starting with those most strongly associated with poverty)
498	The floor of the main dwelling is predominantly made of what material? (Sand, or smoothed mud; Smooth cement, wood, tile, or other)
495	How many household members are 6-years-old or younger? (Three or more; Two; One; None)
472	What does the household head sleep on? (Mat (grass) on floor, cloth/sack on floor, floor (nothing else), or other; Bed and mat (grass), or mattress on floor; Bed and mattress, or bed alone)
465	The roof of the main dwelling is predominantly made of what material? (Grass; Anything besides grass)
443	Has the main activity of the female head/spouse during the last seven days been non-agricultural as an employee, family business worker, self-employed, or employer? (No; Yes; No female head/spouse)
443	How many household members worked in their main activity in the past seven days as a farmer (*mlimi*)? (Four or more; Three; Two; One; None)
441	What general type of construction materials are used for the dwelling? (Traditional; Semi-permanent (mix of traditional (mud, grass) and modern materials (iron sheet, cement); Permanent)
436	Does the household own any irons (for pressing clothes)? (No; Yes)
435	What class is the male head/spouse in or what was the highest level he ever attended? (None to primary standard 3; No male head/spouse; Primary standards 4 to 5; Primary standards 6 to 8; Secondary forms 1 to 3; Secondary form 4 or higher)
422	Does the household own any clocks? (No; Yes)
418	What language does the female head/spouse speak at home? (Not Chewa; Chewa; No female head/spouse)
417	Does the household own any upholstered chairs, sofa sets, coffee tables (for sitting room), cupboards, drawers, or bureaus? (No; Yes)
404	What is your main source of lighting fuel? (Collected firewood, grass, or other; Paraffin; Purchased firewood, electricity, gas, battery/dry cell (torch), or candles)

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

Uncertainty coefficient	Indicator (Respsonses ordered starting with those most strongly associated with poverty)
403	How many beds does the household own? (None; One; Two or more)
374	Does the household own any coffee tables (for sitting room)? (No; Yes)
371	Do you own or are purchasing this house, is it provided to you by an employer, do you use it for free, or do you rent this house? (Owned; Being purchased, employer provides, free (authorized or not authorized); Rented)
365	How many chairs does the household own? (None; One; Two; Three; Four; Five or more)
354	What is the present marital status of the female head/spouse? (Divorced; Polygamous married or non-formal union; Widow; Separated; Monogamous married or non-formal union; Never-married; No female head/spouse)
349	What is the highest educational qualification that the female head/spouse has acquired? (None, or PSLC; JCE, MSCE, or non-university diploma; University diploma or degree, or post-graduate degree; No female head/spouse)
345	Do you have electricity working in your dwelling? (No; Yes)
333	Does the household own any tapes or CD players/hifis? (No; Yes)
326	Does the household own any upholstered chairs/sofa sets? (No; Yes)
322	What is the structure of household headship? (Female head/spouse only; Both male and female heads/spouses; Male head/spouse only)
315	What was your main source of drinking water over the past month? (River/spring, lake/reservoir, or other; Communal hand pump; Communal open, unprotected well; Communal standpipe, personal hand pump, protected spring, or personal open, unprotected well; Piped into dwelling, or piped outside dwelling, personal)
266	What is the present marital status of the male head/spouse? (Polygamous married or non-formal union; No male head/spouse; Monogamous married or non-formal union; Separated, divorced, widower, or never-married)

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

Uncertainty coefficient	Indicator (Respsonses ordered starting with those most strongly associated with poverty)
265	What is the highest educational qualification acquired by any member of the household? (None, PSLC, or JCE; MSCE, non-university diploma, or university diploma or degree; Post-graduate degree)
257	What is the area of the plots (in acres) cultivated by the household during rains? (None; >0 to 1; >1 to 1.5; >1.5 to 2.5; >2.5 to 5; More than 5)
253	How many years ago was this house built? How old is it? (2 to 20 years; 21 or more years; 0 or 1 year; Unknown)
252	Does the household own any cupboards, drawers, or bureaus? (No; Yes)
245	What kind of toilet facility does your household use? (None, traditional latrine without roof only for household members, or other; Traditional latrine without roof shared with other households; Traditional latrine with roof only for household members; Traditional latrine with roof shared with other households; Flush toilet, or VIP latrine)
238	How many tables does the household own? (None; One; Two or more)
231	Does the household own any beds, tables, or chairs? (No; Yes)
222	How many household members worked in their main activity in the past seven days as a farmer (*mlimi*), employee, family-business worker, self-employed, or employer? (Four or more; Three; Two; One; None)
203	How many hoes does the household own? (Five or more; Four; Three; Two; One; None)
201	How many household members worked in their main activity in the past seven days as an employee, family-business worker, self-employed, or employer? (None; One; Two or more)
174	Does any household member engage in any agricultural activities or own agricultural land of any sort? (Yes; No)
157	Can the male head/spouse read a one-page letter in any language? (No; No male head/spouse; Yes)
148	What has been the main activity of the male head/spouse during the last seven days (Farmer (*mlimi*); No male head/spouse; Other; Employee, family business worker, self-employed, or employer)

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

Uncertainty coefficient	Indicator (Respsonses ordered starting with those most strongly associated with poverty)
144	Have the bed nets ever been dipped in insecticide against mosquitoes in the past six months? (No bed nets are used; No, or all nets treated and less than six months old; Yes)
142	Has the main activity of the male head/spouse during the last seven days been non-agricultural as an employee, family business worker, self-employed, or employer? (No male head/spouse; No; Yes)
130	Do any members of your household sleep under a bed net to protect against mosquitos at some time during the year? (No; Yes)
126	Has the main activity of the male head/spouse during the last seven days been agriculture? (Yes; No male head/spouse; No)
125	Does the household own any lanterns (paraffin)? (No; Yes)
120	How many radios ('wireless') does the household own? (None; One; Two or more)
112	The outer walls of the main dwelling of the household are predominantly made of what material? (Grass, or mud (*yomata*); Compacted earth (*yamdindo*); Mud brick (unfired); Burnt bricks; Concrete, wood, iron sheets, or other)
96	What language does the male head/spouse speak at home? (Not Chewa; No male head/spouse; Chewa)
84	Does the household own any bicycles, motorcycles/scooters, cars, mini-buses, or lorries? (No; Yes)
37	What is the highest educational qualification that the male head/spouse has acquired (None, PSLC, JCE, MSCE, or non-university diploma; No male head/spouse; University diploma or degree, or post-graduate degree)
37	How many household members can read a one-page letter in any language? (None; Three or more; Two; One)
37	How many mortars/pestles (mtondo) does the household own? (None; One; Two or more)
34	What type of dwelling does the household live in? (Several separate structures, improvised housing, or other; Single house, flat, or room in a larger dwelling)

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

Uncertainty coefficient	Indicator (Responses ordered starting with those most strongly associated with poverty)
23	How many axes does the household own? (None; One; Two or more)
23	How many sickles does the household own? (None; One; Two or more)
17	How many pangas does the household own? (None; One; Two or more)
13	Does the household own at present any oxen? (No; Yes)
11	Does the household own at present any cattle or oxen? (No; Yes)
10	Does the household own at present any cattle? (No; Yes)
10	How many separate rooms do the members of your household occupy (do not count bathrooms, toilets, storerooms, or garage)? (None, or one; Two; Three; Four or more)
8	Do the children under 5 in the household sleep under a bed net at those times of year when there are mosquitoes present? (Yes, for some or no children under five, no children under five, or no bed nets are used; Yes, for all children under five)
6	Does the household own at present any cattle, oxen, goats, sheep, pigs, chickens, or other poultry? (No; Yes)
0	Does the household own at present any chickens or other poultry? (No; Yes)
0	Does the household own at present any goats, sheep, or pigs? (No; Yes)
0	Does the household own at present any cattle, oxen, goats, sheep, or pigs? (No; Yes)
0	Do any household members attend school as a boarder, or attend a private non-religious school, church/mission school, or Islamic school? (No; Yes)
0	Does the household own at present any sheep? (No; Yes)
0	Does the household own at present any goats? (No; Yes)
0	Does the household own at present any pigs? (No; Yes)

Source: 2004/5 IHS, national poverty line.

Tables for the National Poverty Line

(and tables pertaining to all seven poverty lines)

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

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	95.2
10–14	95.5
15–19	88.9
20–24	82.5
25–29	70.0
30–34	59.3
35–39	47.8
40–44	36.1
45–49	25.5
50–54	13.4
55–59	7.1
60–64	3.9
65–69	0.9
70–74	0.0
75–79	2.2
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

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

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	127	÷	127	=	100.0
5–9	364	÷	383	=	95.2
10–14	1,859	÷	1,946	=	95.5
15–19	4,534	÷	5,101	=	88.9
20–24	6,871	÷	8,326	=	82.5
25–29	8,051	÷	11,503	=	70.0
30–34	7,575	÷	12,783	=	59.3
35–39	5,924	÷	12,401	=	47.8
40–44	4,272	÷	11,846	=	36.1
45–49	2,352	÷	9,230	=	25.5
50–54	1,159	÷	8,671	=	13.4
55–59	425	÷	6,001	=	7.1
60–64	177	÷	4,518	=	3.9
65–69	28	÷	3,193	=	0.9
70–74	0	÷	1,948	=	0.0
75–79	22	÷	988	=	2.2
80–84	0	÷	704	=	0.0
85–89	0	÷	243	=	0.0
90–94	0	÷	64	=	0.0
95–100	0	÷	25	=	0.0

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

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

	Likelihood of having expenditure in range demarcated by adjacent poverty lines					
	<Food	=>Food and <USAID	=>USAID and <National	=>National and <$1.25/day	=>$1.25/day and <$2.50/day	=>$2.50/day
Score	<MWK27.25	=>MWK27.25 and <MWK29.75	=>MWK29.75 and <MWK43.92	=>MWK43.92 and <MWK63.60	=>MWK63.60 and <MWK127.20	=>MWK127.20
0–4	78.8	0.0	21.2	0.0	0.0	0.0
5–9	60.2	0.0	35.1	4.8	0.0	0.0
10–14	63.9	7.0	24.6	3.7	0.8	0.0
15–19	60.2	5.3	23.4	8.6	2.5	0.0
20–24	40.8	7.9	33.8	13.8	3.5	0.2
25–29	30.8	5.0	34.2	21.2	8.2	0.6
30–34	20.1	4.5	34.7	27.6	12.4	0.8
35–39	12.0	2.8	33.0	29.8	21.2	1.3
40–44	6.6	1.7	27.7	31.7	27.7	4.5
45–49	3.5	1.0	21.0	30.5	38.2	5.8
50–54	2.0	0.7	10.6	28.1	48.6	10.0
55–59	0.9	0.4	5.8	17.2	53.2	22.6
60–64	0.0	0.5	3.5	13.1	51.6	31.4
65–69	0.0	0.0	0.9	7.1	42.0	50.0
70–74	0.0	0.0	0.0	5.8	33.6	60.6
75–79	0.0	0.0	2.2	0.0	27.1	70.7
80–84	0.0	0.0	0.0	2.6	23.5	73.9
85–89	0.0	0.0	0.0	10.4	8.7	80.9
90–94	0.0	0.0	0.0	10.4	8.7	80.9
95–100	0.0	0.0	0.0	0.0	0.0	100.0

All poverty likelihoods in percentage units.
The $1.08/day 1993 PPP line (MWK30.12 per person) is about the same as the USAID "extreme" line.
The $2.16/day 1993 PPP line (MWK 60.25 per person) is about the same as the $1.25/day 2005 PPP line.

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

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+0.0	0.0	0.0	0.0
5–9	–4.8	2.4	2.4	2.4
10–14	+1.6	2.1	2.5	3.3
15–19	+0.8	1.8	2.1	2.7
20–24	+1.2	1.8	2.1	2.8
25–29	+0.3	1.7	2.0	2.9
30–34	–1.8	1.8	2.0	2.6
35–39	+0.5	1.9	2.3	3.2
40–44	+1.9	1.7	2.0	2.6
45–49	+1.9	1.8	2.1	2.7
50–54	–2.5	2.1	2.2	2.5
55–59	–1.2	1.7	1.9	2.6
60–64	–0.7	1.3	1.5	2.0
65–69	–1.4	1.3	1.4	1.7
70–74	–0.6	0.6	0.7	0.8
75–79	+2.2	0.0	0.0	0.0
80–84	–1.7	1.7	1.9	2.4
85–89	+0.0	0.0	0.0	0.0
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

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

	Poverty line						
			USAID	Intl. 1993 PPP		Intl. 2005 PPP	
	National	Food	'Extreme'	\$1.08/day	\$2.16/day	\$1.25/day	\$2.50/day
Estimate minus true value							
Scorecard applied to validation sample	+0.1	+0.4	+0.5	+0.1	–0.4	+0.0	+0.2
Precision of difference							
Scorecard applied to validation sample	0.6	0.4	0.4	0.5	0.5	0.5	0.3
α factor							
Scorecard applied to validation sample	0.86	0.92	0.88	0.90	0.83	0.83	0.87

Precision is measured as 90-percent confidence intervals in units of +/– percentage points.
Differences and precision estimated from 500 bootstraps of size n = 16,384.
α is estimated from 1,000 bootstrap samples of n = 256, 512, 1,024, 2,048, 4,096, 8,192, and 16,384.
The USAID "extreme" line is in per-person units.

Figure 9 (National line): Differences and precision of differences for bootstrapped estimates of households' poverty likelihoods, by sample size, scorecard applied to validation sample

Sample Size n	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
	Diff.	90-percent	95-percent	99-percent
1	+0.8	67.0	78.5	87.7
4	+0.2	35.6	42.0	54.3
8	–0.1	25.1	29.5	36.9
16	–0.2	17.9	20.9	26.3
32	–0.2	12.5	14.2	18.3
64	–0.1	8.4	10.2	13.7
128	+0.0	5.9	7.3	9.7
256	+0.0	4.4	5.3	6.7
512	+0.1	3.1	3.6	4.4
1,024	+0.0	2.2	2.5	3.3
2,048	+0.0	1.6	1.8	2.4
4,096	+0.1	1.1	1.3	1.8
8,192	+0.0	0.8	0.9	1.1
16,384	+0.1	0.6	0.6	0.8

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

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

Figure 11 (National line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to 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	0.1	43.5	0.0	56.4	56.5	–99.4
5–9	0.5	43.1	0.0	56.4	56.9	–97.7
10–14	2.3	41.3	0.1	56.3	58.6	–89.0
15–19	6.8	36.8	0.8	55.6	62.4	–67.1
20–24	13.5	30.1	2.4	54.1	67.6	–32.5
25–29	21.5	22.1	5.9	50.5	71.9	+12.0
30–34	29.3	14.3	10.9	45.5	74.8	+59.3
35–39	35.1	8.5	17.4	39.0	74.1	+60.0
40–44	39.2	4.4	25.2	31.2	70.4	+42.2
45–49	41.4	2.2	32.2	24.2	65.6	+26.0
50–54	42.8	0.8	39.5	16.9	59.7	+9.4
55–59	43.3	0.3	45.0	11.4	54.7	–3.3
60–64	43.5	0.1	49.3	7.1	50.6	–13.2
65–69	43.6	0.0	52.5	3.9	47.5	–20.3
70–74	43.6	0.0	54.4	2.0	45.6	–24.8
75–79	43.6	0.0	55.4	1.0	44.6	–27.0
80–84	43.6	0.0	56.1	0.3	43.9	–28.6
85–89	43.6	0.0	56.3	0.1	43.7	–29.2
90–94	43.6	0.0	56.4	0.0	43.6	–29.3
95–100	43.6	0.0	56.4	0.0	43.6	–29.4

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

Figure 12 (National line): Households below the poverty line and all households at a given score or at or below a given score cut-off, scorecard applied to 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	0.1	100.0	0.3	Only poor targeted
5–9	0.5	100.0	1.2	Only poor targeted
10–14	2.5	94.6	5.3	17.6:1
15–19	7.6	90.0	15.6	9.0:1
20–24	15.9	85.2	31.0	5.7:1
25–29	27.4	78.4	49.2	3.6:1
30–34	40.2	72.9	67.2	2.7:1
35–39	52.6	66.8	80.6	2.0:1
40–44	64.4	60.9	90.0	1.6:1
45–49	73.6	56.2	95.0	1.3:1
50–54	82.3	52.0	98.2	1.1:1
55–59	88.3	49.0	99.3	1.0:1
60–64	92.8	46.8	99.8	0.9:1
65–69	96.0	45.4	99.9	0.8:1
70–74	98.0	44.5	100.0	0.8:1
75–79	99.0	44.0	100.0	0.8:1
80–84	99.7	43.7	100.0	0.8:1
85–89	99.9	43.6	100.0	0.8:1
90–94	100.0	43.6	100.0	0.8:1
95–100	100.0	43.6	100.0	0.8:1

Tables for the Food (Ultra) Poverty Line

Figure 4 (Food (ultra) 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	78.8
5–9	60.2
10–14	63.9
15–19	60.2
20–24	40.8
25–29	30.8
30–34	20.1
35–39	12.0
40–44	6.6
45–49	3.5
50–54	2.0
55–59	0.9
60–64	0.0
65–69	0.0
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Figure 7 (Food (ultra) line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) from the validation sample, with confidence intervals

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	–7.6	14.0	15.2	18.1
5–9	–17.8	13.1	13.7	14.8
10–14	–7.4	5.9	6.2	7.0
15–19	+6.0	2.8	3.3	4.6
20–24	–0.9	2.4	2.8	3.6
25–29	+3.4	1.7	2.1	2.7
30–34	+2.3	1.4	1.7	2.3
35–39	+0.7	1.2	1.5	1.9
40–44	–2.8	1.9	2.1	2.3
45–49	+0.5	0.7	0.9	1.1
50–54	–0.3	0.7	0.8	1.0
55–59	–0.1	0.6	0.7	1.0
60–64	–1.2	0.9	1.0	1.2
65–69	+0.0	0.0	0.0	0.0
70–74	+0.0	0.0	0.0	0.0
75–79	+0.0	0.0	0.0	0.0
80–84	+0.0	0.0	0.0	0.0
85–89	+0.0	0.0	0.0	0.0
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 9 (Food (ultra) line): Differences and precision of differences for bootstrapped estimates of households' poverty likelihoods, by sample size, scorecard applied to validation sample

Sample Size n	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
	Diff.	90-percent	95-percent	99-percent
1	–0.6	60.4	64.4	80.2
4	+0.3	27.0	32.3	43.9
8	+0.1	18.4	21.7	30.7
16	+0.3	13.5	15.9	20.8
32	+0.3	9.6	11.8	14.7
64	+0.4	6.9	8.4	10.9
128	+0.4	4.9	5.9	7.5
256	+0.5	3.5	4.1	5.1
512	+0.4	2.5	3.0	4.1
1,024	+0.4	1.8	2.2	2.7
2,048	+0.4	1.2	1.4	1.9
4,096	+0.4	0.9	1.0	1.4
8,192	+0.4	0.6	0.7	1.0
16,384	+0.4	0.4	0.5	0.7

Figure 11 (Food (ultra) line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to 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	0.1	16.4	0.0	83.4	83.5	–98.6
5–9	0.4	16.1	0.1	83.3	83.7	–94.5
10–14	1.8	14.8	0.7	82.7	84.5	–74.6
15–19	4.5	12.0	3.0	80.4	84.9	–27.1
20–24	8.0	8.6	7.9	75.5	83.5	+44.1
25–29	11.1	5.4	16.3	67.2	78.3	+1.8
30–34	13.5	3.1	26.7	56.7	70.2	–61.4
35–39	14.9	1.7	37.7	45.7	60.6	–127.8
40–44	16.0	0.6	48.4	35.0	51.0	–192.7
45–49	16.2	0.3	57.4	26.0	42.3	–246.8
50–54	16.5	0.1	65.9	17.6	34.0	–297.9
55–59	16.5	0.1	71.8	11.6	28.1	–333.9
60–64	16.6	0.0	76.3	7.2	23.7	–360.8
65–69	16.6	0.0	79.5	4.0	20.5	–380.1
70–74	16.6	0.0	81.4	2.0	18.6	–391.9
75–79	16.6	0.0	82.4	1.0	17.6	–397.9
80–84	16.6	0.0	83.1	0.3	16.9	–402.1
85–89	16.6	0.0	83.4	0.1	16.6	–403.6
90–94	16.6	0.0	83.4	0.0	16.6	–404.0
95–100	16.6	0.0	83.4	0.0	16.6	–404.1

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

Figure 12 (Food (ultra) line): Households below the poverty line and all households at a given score or at or below a given score cut-off, scorecard applied to 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	0.1	85.6	0.7	5.9:1
5–9	0.5	79.4	2.4	3.8:1
10–14	2.5	71.5	10.6	2.5:1
15–19	7.6	59.7	27.3	1.5:1
20–24	15.9	50.2	48.1	1.0:1
25–29	27.4	40.6	67.2	0.7:1
30–34	40.2	33.5	81.3	0.5:1
35–39	52.6	28.3	89.8	0.4:1
40–44	64.4	24.8	96.5	0.3:1
45–49	73.6	22.0	98.1	0.3:1
50–54	82.3	20.0	99.4	0.2:1
55–59	88.3	18.7	99.7	0.2:1
60–64	92.8	17.8	100.0	0.2:1
65–69	96.0	17.2	100.0	0.2:1
70–74	98.0	16.9	100.0	0.2:1
75–79	99.0	16.7	100.0	0.2:1
80–84	99.7	16.6	100.0	0.2:1
85–89	99.9	16.6	100.0	0.2:1
90–94	100.0	16.6	100.0	0.2:1
95–100	100.0	16.6	100.0	0.2:1

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	78.8
5–9	60.2
10–14	70.9
15–19	65.5
20–24	48.7
25–29	35.8
30–34	24.5
35–39	14.8
40–44	8.4
45–49	4.4
50–54	2.7
55–59	1.3
60–64	0.5
65–69	0.0
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

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

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Score	**Diff.**	**90-percent**	**95-percent**	**99-percent**
0–4	–21.2	10.6	10.6	10.6
5–9	–14.7	11.6	12.3	14.0
10–14	–6.4	5.1	5.5	6.0
15–19	+3.0	2.7	3.2	4.6
20–24	+0.8	2.3	2.9	3.6
25–29	+2.1	1.8	2.2	2.9
30–34	+1.1	1.6	1.9	2.5
35–39	+1.9	1.2	1.4	1.9
40–44	–0.9	1.1	1.4	1.8
45–49	+0.9	0.8	1.0	1.2
50–54	–0.8	0.8	1.0	1.3
55–59	+0.2	0.6	0.7	1.0
60–64	–1.7	1.3	1.4	1.7
65–69	+0.0	0.0	0.0	0.0
70–74	+0.0	0.0	0.0	0.0
75–79	+0.0	0.0	0.0	0.0
80–84	+0.0	0.0	0.0	0.0
85–89	+0.0	0.0	0.0	0.0
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 9 (USAID “extreme” line): Differences and precision of differences for bootstrapped estimates of households’ poverty likelihoods, by sample size, scorecard applied to validation sample

Sample Size n	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
	Diff.	90-percent	95-percent	99-percent
1	–0.3	62.1	67.0	84.1
4	+0.3	27.3	33.2	42.9
8	+0.1	20.2	23.8	30.6
16	+0.2	13.7	16.4	21.6
32	+0.4	10.5	12.1	15.3
64	+0.5	7.4	8.5	10.9
128	+0.5	5.0	6.1	7.4
256	+0.5	3.4	4.3	6.1
512	+0.5	2.7	3.2	3.9
1,024	+0.4	1.8	2.2	2.9
2,048	+0.5	1.3	1.5	1.9
4,096	+0.5	0.9	1.1	1.4
8,192	+0.5	0.6	0.8	1.0
16,384	+0.5	0.4	0.6	0.7

Figure 11 (USAID "extreme" line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to 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	0.1	19.4	0.0	80.5	80.6	–98.7
5–9	0.4	19.1	0.1	80.4	80.8	–95.3
10–14	1.9	17.6	0.6	79.9	81.8	–77.8
15–19	5.0	14.5	2.5	78.0	83.0	–35.4
20–24	9.0	10.5	6.9	73.6	82.6	+27.7
25–29	12.9	6.6	14.5	66.0	78.9	+25.6
30–34	15.9	3.6	24.3	56.2	72.2	–24.4
35–39	17.5	2.0	35.0	45.5	63.0	–79.6
40–44	18.7	0.8	45.8	34.7	53.4	–134.6
45–49	19.0	0.5	54.6	25.8	44.8	–180.2
50–54	19.3	0.2	63.0	17.5	36.9	–222.9
55–59	19.4	0.1	68.9	11.6	31.0	–253.4
60–64	19.5	0.0	73.3	7.2	26.7	–276.0
65–69	19.5	0.0	76.5	4.0	23.5	–292.4
70–74	19.5	0.0	78.5	2.0	21.5	–302.4
75–79	19.5	0.0	79.5	1.0	20.5	–307.4
80–84	19.5	0.0	80.2	0.3	19.8	–311.1
85–89	19.5	0.0	80.4	0.1	19.6	–312.3
90–94	19.5	0.0	80.5	0.0	19.5	–312.6
95–100	19.5	0.0	80.5	0.0	19.5	–312.8

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

Figure 12 (USAID "extreme" line): Households below the poverty line and all households at a given score or at or below a given score cut-off, scorecard applied to 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	0.1	100.0	0.7	Only poor targeted
5–9	0.5	80.1	2.1	4.0:1
10–14	2.5	76.2	9.6	3.2:1
15–19	7.6	66.7	25.8	2.0:1
20–24	15.9	56.7	46.2	1.3:1
25–29	27.4	47.0	66.0	0.9:1
30–34	40.2	39.6	81.6	0.7:1
35–39	52.6	33.4	89.9	0.5:1
40–44	64.4	29.0	95.7	0.4:1
45–49	73.6	25.8	97.4	0.3:1
50–54	82.3	23.5	99.2	0.3:1
55–59	88.3	22.0	99.5	0.3:1
60–64	92.8	21.0	100.0	0.3:1
65–69	96.0	20.3	100.0	0.3:1
70–74	98.0	19.9	100.0	0.2:1
75–79	99.0	19.7	100.0	0.2:1
80–84	99.7	19.6	100.0	0.2:1
85–89	99.9	19.5	100.0	0.2:1
90–94	100.0	19.5	100.0	0.2:1
95–100	100.0	19.5	100.0	0.2:1

Tables for the $1.08/day 1993 PPP Poverty Line

Figure 4 ($1.08/day 1993 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	78.8
5–9	67.0
10–14	71.9
15–19	68.3
20–24	49.9
25–29	38.2
30–34	28.1
35–39	17.2
40–44	10.7
45–49	6.0
50–54	3.6
55–59	2.1
60–64	0.5
65–69	0.0
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Figure 7 ($1.08/day 1993 PPP line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample ($n = 16,384$) from the validation sample, with confidence intervals

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Score	**Diff.**	**90-percent**	**95-percent**	**99-percent**
0–4	–21.2	10.6	10.6	10.6
5–9	–11.0	9.7	10.3	12.8
10–14	–4.7	4.3	4.6	5.8
15–19	+3.0	2.7	3.2	4.7
20–24	+0.5	2.3	2.8	3.6
25–29	+0.4	1.9	2.3	3.0
30–34	+1.0	1.7	2.0	2.5
35–39	+0.1	1.4	1.7	2.2
40–44	–1.1	1.2	1.4	1.9
45–49	+1.0	0.9	1.1	1.5
50–54	–0.1	0.8	1.0	1.3
55–59	–0.3	0.9	1.0	1.4
60–64	–1.3	1.1	1.1	1.4
65–69	+0.0	0.0	0.0	0.0
70–74	+0.0	0.0	0.0	0.0
75–79	+0.0	0.0	0.0	0.0
80–84	+0.0	0.0	0.0	0.0
85–89	+0.0	0.0	0.0	0.0
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 9 ($1.08/day 1993 PPP line): Differences and precision of differences for bootstrapped estimates of households' poverty likelihoods, by sample size, scorecard applied to validation sample

Sample Size n	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
	Diff.	90-percent	95-percent	99-percent
1	–0.8	60.9	69.6	83.6
4	–0.4	29.4	34.5	43.7
8	–0.5	21.9	25.0	32.6
16	–0.3	15.2	18.0	23.3
32	–0.2	10.5	12.8	16.4
64	–0.1	7.6	9.0	11.8
128	+0.0	5.4	6.3	8.2
256	+0.1	3.8	4.7	6.1
512	+0.1	2.8	3.4	4.3
1,024	+0.0	1.9	2.3	3.2
2,048	+0.1	1.3	1.6	2.0
4,096	+0.1	1.0	1.2	1.5
8,192	+0.1	0.6	0.8	1.1
16,384	+0.1	0.5	0.6	0.7

Figure 11 ($1.08/day 1993 PPP line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to 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	0.1	21.5	0.0	78.4	78.5	–98.8
5–9	0.4	21.2	0.1	78.3	78.7	–95.7
10–14	1.9	19.7	0.6	77.8	79.7	–80.0
15–19	5.2	16.4	2.4	76.0	81.2	–41.1
20–24	9.3	12.4	6.6	71.7	81.0	+16.3
25–29	13.6	8.0	13.8	64.6	78.2	+36.2
30–34	17.1	4.5	23.1	55.3	72.4	–6.8
35–39	19.2	2.4	33.4	45.0	64.2	–54.2
40–44	20.6	1.0	43.8	34.6	55.2	–102.5
45–49	21.1	0.6	52.6	25.8	46.9	–143.1
50–54	21.4	0.2	60.9	17.5	38.9	–181.5
55–59	21.5	0.1	66.8	11.6	33.1	–208.7
60–64	21.6	0.0	71.2	7.2	28.8	–229.2
65–69	21.6	0.0	74.4	4.0	25.6	–244.0
70–74	21.6	0.0	76.3	2.0	23.7	–253.0
75–79	21.6	0.0	77.3	1.0	22.7	–257.6
80–84	21.6	0.0	78.0	0.3	22.0	–260.8
85–89	21.6	0.0	78.3	0.1	21.7	–261.9
90–94	21.6	0.0	78.3	0.0	21.7	–262.2
95–100	21.6	0.0	78.4	0.0	21.6	–262.4

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

Figure 12 ($1.08/day 1993 PPP line): Households below the poverty line and all households at a given score or at or below a given score cut-off, scorecard applied to 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	0.1	100.0	0.6	Only poor targeted
5–9	0.5	83.0	2.0	4.9:1
10–14	2.5	76.6	8.7	3.3:1
15–19	7.6	68.7	24.0	2.2:1
20–24	15.9	58.3	42.8	1.4:1
25–29	27.4	49.6	62.8	1.0:1
30–34	40.2	42.5	79.0	0.7:1
35–39	52.6	36.6	88.8	0.6:1
40–44	64.4	32.0	95.3	0.5:1
45–49	73.6	28.6	97.4	0.4:1
50–54	82.3	26.0	99.1	0.4:1
55–59	88.3	24.4	99.6	0.3:1
60–64	92.8	23.3	100.0	0.3:1
65–69	96.0	22.5	100.0	0.3:1
70–74	98.0	22.1	100.0	0.3:1
75–79	99.0	21.9	100.0	0.3:1
80–84	99.7	21.7	100.0	0.3:1
85–89	99.9	21.6	100.0	0.3:1
90–94	100.0	21.6	100.0	0.3:1
95–100	100.0	21.6	100.0	0.3:1

Tables for the \$2.16/day 1993 PPP Poverty Line

Figure 4 ($2.16/day 1993 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	97.8
15–19	96.7
20–24	94.3
25–29	89.4
30–34	83.2
35–39	74.0
40–44	61.2
45–49	48.9
50–54	36.5
55–59	21.4
60–64	13.0
65–69	5.2
70–74	5.8
75–79	2.2
80–84	2.6
85–89	10.4
90–94	10.4
95–100	0.0

Figure 7 ($2.16/day 1993 PPP line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) from the validation sample, with confidence intervals

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	**Diff.**	**90-percent**	**95-percent**	**99-percent**
0–4	+0.0	0.0	0.0	0.0
5–9	+0.0	0.0	0.0	0.0
10–14	+0.7	1.6	1.8	2.4
15–19	+0.2	1.0	1.3	1.6
20–24	–1.1	1.0	1.1	1.5
25–29	+0.3	1.2	1.5	1.8
30–34	–3.0	2.1	2.2	2.4
35–39	+1.1	1.7	2.0	2.6
40–44	+1.5	1.9	2.2	3.0
45–49	+0.0	2.2	2.6	3.4
50–54	–4.2	3.3	3.4	3.8
55–59	–1.3	2.4	2.9	3.7
60–64	+1.6	2.0	2.5	3.1
65–69	–2.6	2.3	2.4	2.9
70–74	+3.5	1.3	1.5	2.1
75–79	+2.0	0.3	0.4	0.5
80–84	+1.0	1.7	1.9	2.4
85–89	+10.4	0.0	0.0	0.0
90–94	+10.4	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 9 ($2.16/day 1993 PPP line): Differences and precision of differences for bootstrapped estimates of households' poverty likelihoods, by sample size, scorecard applied to validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+0.4	62.6	73.3	90.7
4	–0.3	30.4	37.3	48.2
8	–0.8	21.8	25.7	32.7
16	–0.8	16.0	19.5	26.4
32	–0.8	11.0	13.4	17.0
64	–0.6	8.2	9.6	12.2
128	–0.5	5.7	6.9	9.4
256	–0.5	4.2	4.9	6.2
512	–0.4	3.0	3.6	4.7
1,024	–0.4	2.1	2.5	3.2
2,048	–0.5	1.5	1.8	2.2
4,096	–0.5	1.1	1.2	1.5
8,192	–0.5	0.7	0.9	1.1
16,384	–0.4	0.5	0.6	0.8

Figure 11 ($2.16/day 1993 PPP line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to 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	0.1	62.7	0.0	37.2	37.3	–99.6
5–9	0.5	62.3	0.0	37.2	37.7	–98.4
10–14	2.4	60.4	0.1	37.1	39.5	–92.3
15–19	7.3	55.5	0.2	36.9	44.2	–76.3
20–24	15.3	47.6	0.6	36.6	51.8	–50.4
25–29	25.5	37.3	1.9	35.3	60.8	–15.8
30–34	36.5	26.3	3.7	33.5	70.0	+22.0
35–39	45.5	17.3	7.0	30.1	75.7	+56.1
40–44	52.6	10.2	11.8	25.4	78.0	+81.2
45–49	57.1	5.7	16.5	20.6	77.7	+73.7
50–54	60.6	2.2	21.7	15.5	76.1	+65.5
55–59	62.0	0.8	26.3	10.8	72.8	+58.1
60–64	62.5	0.3	30.3	6.8	69.3	+51.7
65–69	62.8	0.1	33.3	3.9	66.7	+47.0
70–74	62.8	0.0	35.2	2.0	64.8	+44.0
75–79	62.8	0.0	36.1	1.0	63.8	+42.5
80–84	62.8	0.0	36.8	0.3	63.2	+41.4
85–89	62.8	0.0	37.1	0.1	62.9	+41.0
90–94	62.8	0.0	37.1	0.0	62.9	+40.9
95–100	62.8	0.0	37.2	0.0	62.8	+40.8

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

Figure 12 ($2.16/day 1993 PPP line): Households below the poverty line and all households at a given score or at or below a given score cut-off, scorecard applied to 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	0.1	100.0	0.2	Only poor targeted
5–9	0.5	100.0	0.8	Only poor targeted
10–14	2.5	97.7	3.8	43.0:1
15–19	7.6	96.8	11.6	30.1:1
20–24	15.9	96.2	24.3	25.2:1
25–29	27.4	93.2	40.6	13.7:1
30–34	40.2	90.9	58.1	10.0:1
35–39	52.6	86.6	72.5	6.5:1
40–44	64.4	81.7	83.8	4.5:1
45–49	73.6	77.5	90.9	3.5:1
50–54	82.3	73.7	96.5	2.8:1
55–59	88.3	70.2	98.7	2.4:1
60–64	92.8	67.3	99.5	2.1:1
65–69	96.0	65.4	99.9	1.9:1
70–74	98.0	64.1	100.0	1.8:1
75–79	99.0	63.5	100.0	1.7:1
80–84	99.7	63.0	100.0	1.7:1
85–89	99.9	62.9	100.0	1.7:1
90–94	100.0	62.8	100.0	1.7:1
95–100	100.0	62.8	100.0	1.7:1

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	99.2
15–19	97.5
20–24	96.3
25–29	91.2
30–34	86.8
35–39	77.5
40–44	67.8
45–49	56.0
50–54	41.5
55–59	24.3
60–64	17.0
65–69	8.0
70–74	5.8
75–79	2.2
80–84	2.6
85–89	10.4
90–94	10.4
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) from the validation sample, with confidence intervals

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	**Diff.**	**90-percent**	**95-percent**	**99-percent**
0–4	+0.0	0.0	0.0	0.0
5–9	+0.0	0.0	0.0	0.0
10–14	+1.6	1.5	1.8	2.3
15–19	–0.5	0.8	1.0	1.3
20–24	+0.5	0.9	1.1	1.5
25–29	–0.5	1.1	1.2	1.6
30–34	–2.0	1.5	1.6	1.8
35–39	+1.1	1.6	1.9	2.4
40–44	+2.4	1.9	2.3	2.8
45–49	+2.2	2.3	2.7	3.5
50–54	–4.0	3.1	3.3	3.7
55–59	–2.2	2.5	3.0	4.0
60–64	+2.8	2.2	2.6	3.7
65–69	–3.1	2.7	2.9	3.6
70–74	+2.5	1.6	1.9	2.4
75–79	+2.0	0.3	0.4	0.5
80–84	+1.0	1.7	1.9	2.4
85–89	+10.4	0.0	0.0	0.0
90–94	+10.4	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 9 ($1.25/day 2005 PPP line): Differences and precision of differences for bootstrapped estimates of households' poverty likelihoods, by sample size, scorecard applied to validation sample

Sample Size n	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
	Diff.	90-percent	95-percent	99-percent
1	+1.2	68.0	77.0	88.4
4	–0.0	30.5	35.7	47.1
8	–0.6	21.0	25.0	31.1
16	–0.4	15.3	19.2	23.9
32	–0.4	10.7	12.7	16.8
64	–0.2	7.9	9.4	11.7
128	–0.0	5.8	6.6	9.1
256	–0.1	4.1	4.7	6.3
512	+0.0	2.9	3.3	4.4
1,024	+0.0	2.0	2.5	3.1
2,048	–0.0	1.4	1.7	2.2
4,096	+0.0	1.0	1.2	1.5
8,192	+0.0	0.7	0.8	1.1
16,384	+0.0	0.5	0.6	0.8

Figure 11 ($1.25/day 2005 PPP line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to 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	0.1	65.9	0.0	34.0	34.1	–99.6
5–9	0.5	65.5	0.0	34.0	34.5	–98.5
10–14	2.4	63.6	0.0	34.0	36.4	–92.6
15–19	7.4	58.6	0.2	33.8	41.2	–77.3
20–24	15.4	50.6	0.5	33.5	48.9	–52.6
25–29	26.0	40.1	1.4	32.6	58.5	–19.2
30–34	37.2	28.8	2.9	31.1	68.3	+17.3
35–39	46.7	19.3	5.9	28.1	74.9	+50.4
40–44	54.4	11.6	10.0	24.0	78.4	+80.0
45–49	59.4	6.6	14.3	19.7	79.1	+78.4
50–54	63.3	2.7	19.0	15.0	78.4	+71.2
55–59	64.9	1.1	23.4	10.6	75.5	+64.5
60–64	65.5	0.5	27.3	6.7	72.3	+58.7
65–69	65.9	0.1	30.1	3.9	69.8	+54.4
70–74	66.0	0.0	32.0	2.0	68.0	+51.5
75–79	66.0	0.0	33.0	1.0	67.0	+50.0
80–84	66.0	0.0	33.7	0.3	66.3	+49.0
85–89	66.0	0.0	33.9	0.1	66.1	+48.6
90–94	66.0	0.0	34.0	0.0	66.0	+48.5
95–100	66.0	0.0	34.0	0.0	66.0	+48.5

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

Figure 12 ($1.25/day 2005 PPP line): Households below the poverty line and all households at a given score or at or below a given score cut-off, scorecard applied to 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	0.1	100.0	0.2	Only poor targeted
5–9	0.5	100.0	0.8	Only poor targeted
10–14	2.5	98.3	3.7	57.4:1
15–19	7.6	98.0	11.2	48.1:1
20–24	15.9	97.0	23.3	32.4:1
25–29	27.4	94.8	39.3	18.1:1
30–34	40.2	92.7	56.4	12.8:1
35–39	52.6	88.9	70.8	8.0:1
40–44	64.4	84.5	82.4	5.4:1
45–49	73.6	80.6	90.0	4.2:1
50–54	82.3	76.9	96.0	3.3:1
55–59	88.3	73.5	98.3	2.8:1
60–64	92.8	70.6	99.3	2.4:1
65–69	96.0	68.6	99.9	2.2:1
70–74	98.0	67.3	100.0	2.1:1
75–79	99.0	66.7	100.0	2.0:1
80–84	99.7	66.2	100.0	2.0:1
85–89	99.9	66.1	100.0	1.9:1
90–94	100.0	66.0	100.0	1.9:1
95–100	100.0	66.0	100.0	1.9:1

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	100.0
20–24	99.8
25–29	99.4
30–34	99.3
35–39	98.7
40–44	95.5
45–49	94.2
50–54	90.0
55–59	77.4
60–64	68.6
65–69	50.0
70–74	39.4
75–79	29.3
80–84	26.1
85–89	19.1
90–94	19.1
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) from the validation sample, with confidence intervals

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	**Diff.**	**90-percent**	**95-percent**	**99-percent**
0–4	+0.0	0.0	0.0	0.0
5–9	+0.0	0.0	0.0	0.0
10–14	+0.0	0.0	0.0	0.0
15–19	+0.0	0.0	0.0	0.0
20–24	–0.0	0.2	0.2	0.2
25–29	+0.2	0.3	0.4	0.6
30–34	+0.2	0.4	0.4	0.5
35–39	–0.0	0.4	0.5	0.6
40–44	–0.5	0.8	0.9	1.1
45–49	+0.4	1.0	1.2	1.6
50–54	+1.9	1.5	1.7	2.3
55–59	–2.3	2.4	2.7	3.6
60–64	+1.1	3.0	3.6	4.9
65–69	–5.9	4.7	5.0	6.0
70–74	+4.0	4.6	5.6	7.1
75–79	+10.4	5.0	6.0	7.5
80–84	+8.6	5.6	6.7	9.0
85–89	+18.2	1.4	1.6	2.1
90–94	+19.1	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 9 ($2.50/day 2005 PPP line): Differences and precision of differences for bootstrapped estimates of households' poverty likelihoods, by sample size, scorecard applied to validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	–0.7	40.7	63.7	77.4
4	–0.1	19.4	25.5	36.2
8	–0.2	14.2	17.5	23.1
16	+0.1	10.1	12.4	17.3
32	–0.1	7.3	8.7	11.6
64	+0.2	5.1	5.9	7.7
128	+0.2	3.5	4.1	5.2
256	+0.3	2.4	3.1	4.0
512	+0.3	1.8	2.1	2.8
1,024	+0.2	1.4	1.6	1.9
2,048	+0.2	0.9	1.1	1.4
4,096	+0.2	0.7	0.8	1.1
8,192	+0.2	0.5	0.6	0.7
16,384	+0.2	0.3	0.4	0.5

Figure 11 ($2.50/day 2005 PPP line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to 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	0.1	90.5	0.0	9.3	9.5	–99.7
5–9	0.5	90.2	0.0	9.3	9.8	–98.9
10–14	2.5	88.2	0.0	9.3	11.8	–94.6
15–19	7.6	83.1	0.0	9.3	16.9	–83.3
20–24	15.9	74.8	0.0	9.3	25.2	–65.0
25–29	27.3	63.4	0.1	9.2	36.5	–39.7
30–34	39.9	50.7	0.2	9.1	49.1	–11.6
35–39	52.2	38.5	0.4	8.9	61.1	+15.5
40–44	63.6	27.1	0.8	8.5	72.1	+41.2
45–49	72.2	18.4	1.4	7.9	80.1	+60.9
50–54	79.9	10.8	2.5	6.9	86.7	+78.9
55–59	84.7	6.0	3.6	5.7	90.4	+90.8
60–64	87.8	2.9	5.0	4.3	92.1	+94.4
65–69	89.6	1.0	6.4	2.9	92.5	+92.9
70–74	90.3	0.3	7.7	1.7	92.0	+91.6
75–79	90.5	0.1	8.4	0.9	91.4	+90.7
80–84	90.7	0.0	9.0	0.3	91.0	+90.1
85–89	90.7	0.0	9.3	0.1	90.7	+89.8
90–94	90.7	0.0	9.3	0.0	90.7	+89.7
95–100	90.7	0.0	9.3	0.0	90.7	+89.7

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

Figure 12 ($2.50/day 2005 PPP line): Households below the poverty line and all households at a given score or at or below a given score cut-off, scorecard applied to 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	0.1	100.0	0.1	Only poor targeted
5–9	0.5	100.0	0.6	Only poor targeted
10–14	2.5	100.0	2.7	Only poor targeted
15–19	7.6	100.0	8.3	Only poor targeted
20–24	15.9	99.9	17.5	949.3:1
25–29	27.4	99.6	30.1	238.3:1
30–34	40.2	99.4	44.1	175.3:1
35–39	52.6	99.3	57.6	133.6:1
40–44	64.4	98.7	70.1	75.7:1
45–49	73.6	98.1	79.7	50.5:1
50–54	82.3	97.0	88.1	32.5:1
55–59	88.3	95.9	93.4	23.3:1
60–64	92.8	94.6	96.8	17.4:1
65–69	96.0	93.3	98.8	14.0:1
70–74	98.0	92.2	99.6	11.8:1
75–79	99.0	91.5	99.8	10.7:1
80–84	99.7	91.0	100.0	10.1:1
85–89	99.9	90.7	100.0	9.8:1
90–94	100.0	90.7	100.0	9.7:1
95–100	100.0	90.7	100.0	9.7:1