

Simple Poverty Scorecard[®] Poverty-Assessment Tool Malawi

Mark Schreiner

23 August 2015

This document is at SimplePovertyScorecard.com.

Abstract

The Simple Poverty Scorecard-brand poverty-assessment tool uses ten low-cost indicators from Malawi's 2010/11 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.

Version note

This paper uses 2010/11 data, replacing Schreiner (2011), which uses 2004/5 data. The new 2010/11 scorecard here should be used from now on. Existing users of Schreiner (2011) can still measure change over time using supported poverty lines with a baseline from the old 2004/5 scorecard and a follow-up from the new 2010/11 scorecard.

Acknowledgements

This paper was funded by the Private Sector Window of the Global Agriculture and Food Security Program, and by the International Finance Corporation. Data are from Malawi's National Statistical Office. Thanks go to Daniella Hawkins, Joseph Kaipa, Malumbo Mhango, and Karl Pauw. "Simple Poverty Scorecard" is a Registered Trademark of Microfinance Risk Management, L.L.C. for its brand of poverty-assessment tools.

Simple Poverty Scorecard[®] Poverty-Assessment Tool

Interview ID: _____	<u>Name</u>	<u>Identifier</u>
Interview date: _____	Participant: _____	_____
Country: <u>MWI</u>	Field agent: _____	_____
Scorecard: <u>002</u>	Service point: _____	_____
Sampling wgt.: _____	Number of household members: _____	

Indicator	Response	Points	Score
1. How many members does the household have?	A. Seven or more	0	
	B. Six	4	
	C. Five	10	
	D. Four	15	
	E. One, two, or three	31	
2. Is the (oldest) female head/spouse able to read and write in Chichewa or English?	A. No	0	
	B. Yes, only Chichewa	4	
	C. Yes, English (regardless of Chichewa)	8	
	D. No female head/spouse	13	
3. The floor of the main dwelling is predominantly made of what material?	A. Smoothed mud, or sand	0	
	B. Smooth cement, wood, tile, or other	8	
4. The outer walls of the main dwelling of the household are predominantly made of what material?	A. Mud (<i>yomata</i>), or grass	0	
	B. Mud brick (unfired)	5	
	C. Compacted earth (<i>yamdindo</i>), burnt bricks, concrete, wood, iron sheets, or other	8	
5. The roof of the main dwelling is predominantly made of what material?	A. Grass, plastic sheeting, or other	0	
	B. Iron sheets, clay tiles, or concrete	3	
6. What kind of toilet facility does the household use?	A. None, traditional latrine without roof shared with other households, or other	0	
	B. Traditional latrine without roof only for household members	4	
	C. Traditional latrine with roof shared with other households	4	
	D. Traditional latrine with roof only for household members, VIP latrine, or flush toilet	6	
7. What is the household's main source of lighting fuel?	A. Collected firewood, purchased firewood, grass, or gas	0	
	B. Paraffin, or other	8	
	C. Battery/dry cell (torch), candles, or electricity	13	
8. Do any members of the household sleep under a bed net to protect against mosquitos at some time during the year?	A. No	0	
	B. Yes	5	
9. Does the household own any tables?	A. No	0	
	B. Yes	9	
10. Does the household own any beds?	A. No	0	
	B. Yes	4	

Back-page Worksheet: Household Membership

In the scorecard header, record the unique identifier of the interview (if known), the date of the interview, and the sampling weight of the client (if known). Then record the name and identification number of the client, of yourself as the field agent, and of the service point the client uses.

Then read to the respondent: *I would like to make a complete list of the names of all the members of the household. A household is a group of people who live together, pool their money, and eat at least one meal together each day.* Give the respondent the following instructions:

First: *Please give me the first names of all the members of your immediate family who normally live and eat their meals together here.* Record the responses. List the head of the household first. For your own future use, note the (oldest) female head/spouse (if she exists). If there is more than one female spouse of the head in the household, then ask for the ages of each in order to determine who is the oldest.

Second: *Please give me the first names of any other persons related to you or other household members who normally live and eat their meals together here.* Record the responses.

Third: *Are there any other persons not here now who normally live and eat their meals here? For example, household members studying elsewhere or travelling.* Record the responses.

Fourth: *Please give me the first names of any other persons not related to you or other household members but who normally live and eat their meals together here, such as servants, lodgers, or others who are not relatives.* Record the responses.

Count the total number of household members. In the scorecard header, record this next to “Number of household members:”, and circle the response to the first scorecard indicator.

Keep in mind the full definitions of *household* and *household member* in “Guidelines for the Interpretation of Indicators”.

First name
1.
2.
3.
4.
5.
6.
7.
8.
9.
10.
Total number of household members:

**Look-up table to convert scores to poverty likelihoods:
PBM-definition poverty lines and
the line that marks the poorest half of people
below 100% of the PBM-definition national line**

Score	Poverty likelihood (%)				
	PBM-def. national lines				Poorest half of people
	Food	100%	150%	200%	<100% Govt.-def. natl. line
0-4	100.0	100.0	100.0	100.0	100.0
5-9	68.1	86.5	99.7	99.7	68.1
10-14	59.7	85.9	97.1	98.5	60.9
15-19	58.6	85.6	94.8	98.3	59.9
20-24	46.5	77.6	91.3	94.6	50.0
25-29	35.8	64.8	84.2	90.5	38.6
30-34	25.7	55.1	80.0	90.5	26.8
35-39	20.0	47.1	77.0	89.5	21.1
40-44	14.7	39.6	68.1	83.3	17.1
45-49	10.5	32.5	60.1	78.4	13.6
50-54	5.6	20.7	43.8	64.4	6.5
55-59	3.6	16.7	38.1	58.2	5.3
60-64	2.1	12.8	34.5	53.4	2.5
65-69	0.9	7.2	27.3	45.3	1.0
70-74	0.6	4.2	15.1	34.4	0.7
75-79	0.6	3.5	11.7	23.3	0.7
80-84	0.4	1.5	7.1	18.9	0.7
85-89	0.0	0.8	4.2	13.7	0.7
90-94	0.0	0.8	2.0	10.7	0.7
95-100	0.0	0.0	0.0	0.0	0.0

**Look-up table to convert scores to poverty likelihoods:
PBM-definition Intl. 2005 and 2011 PPP poverty lines**

Score	Poverty likelihood (%)						
	Intl. 2005 PPP lines					Intl. 2011 PPP lines	
	\$1.25	\$2.00	\$2.50	\$5.00	\$8.44	\$1.90	\$3.10
0–4	100.0	100.0	100.0	100.0	100.0	100.0	100.0
5–9	97.1	100.0	100.0	100.0	100.0	99.7	100.0
10–14	95.9	99.9	100.0	100.0	100.0	98.4	100.0
15–19	94.6	99.5	100.0	100.0	100.0	97.1	100.0
20–24	90.5	97.4	99.6	99.9	100.0	94.0	99.6
25–29	83.4	95.0	98.1	99.3	99.9	88.5	98.1
30–34	77.6	95.0	97.1	99.3	99.8	88.1	97.1
35–39	73.8	93.7	96.7	99.2	99.8	84.1	96.9
40–44	65.5	88.9	94.4	99.1	99.8	78.0	94.4
45–49	58.0	83.8	92.5	99.1	99.8	72.0	92.6
50–54	41.6	72.3	86.7	99.1	99.8	56.5	86.9
55–59	35.2	68.6	82.7	96.7	98.9	50.9	83.2
60–64	30.9	64.8	78.1	96.1	98.7	46.6	78.6
65–69	24.4	54.3	66.7	93.4	98.0	38.5	66.8
70–74	13.3	41.6	56.6	89.4	95.8	27.5	57.0
75–79	10.3	30.2	45.8	84.8	94.1	17.8	46.5
80–84	6.4	27.9	43.3	76.9	93.4	13.4	43.8
85–89	2.4	18.1	30.4	70.5	91.8	6.7	30.4
90–94	1.3	11.2	18.9	69.1	83.0	4.3	18.9
95–100	0.0	7.8	12.0	46.3	76.0	0.0	12.0

**Look-up table to convert scores to poverty likelihoods:
Government-definition poverty lines and
the line that marks the poorest half of people
below 100% of the government-definition national line**

Score	Poverty likelihood (%)				
	Govt.-def. national lines				Poorest half of people
	Food	100%	150%	200%	<100% Govt.-def. natl. line
0-4	100.0	100.0	100.0	100.0	100.0
5-9	81.3	97.1	99.7	100.0	81.3
10-14	71.8	95.7	98.6	100.0	73.3
15-19	70.3	94.4	98.5	100.0	72.3
20-24	55.7	87.4	96.3	98.6	60.3
25-29	49.9	80.0	94.7	98.1	51.3
30-34	36.4	72.0	94.7	97.6	37.4
35-39	27.6	70.2	90.1	96.4	29.0
40-44	21.5	57.4	82.7	93.2	23.6
45-49	16.5	47.9	76.8	90.3	18.0
50-54	9.4	30.5	57.1	80.5	9.1
55-59	5.6	24.9	48.9	74.4	6.1
60-64	4.0	20.0	44.1	67.1	3.7
65-69	2.2	12.4	34.1	53.0	2.2
70-74	1.0	6.5	23.3	38.3	0.7
75-79	0.6	5.3	17.7	31.2	0.6
80-84	0.4	2.7	11.5	24.5	0.4
85-89	0.0	1.1	3.7	13.3	0.0
90-94	0.0	1.1	2.4	8.8	0.0
95-100	0.0	0.0	0.0	0.6	0.0

**Look-up table to convert scores to poverty likelihoods:
Govt.-definition Intl. 2005 and 2011 PPP poverty lines**

Score	Poverty likelihood (%)						
	Intl. 2005 PPP lines					Intl. 2011 PPP lines	
	\$1.25	\$2.00	\$2.50	\$5.00	\$8.44	\$1.90	\$3.10
0-4	100.0	100.0	100.0	100.0	100.0	100.0	100.0
5-9	99.7	100.0	100.0	100.0	100.0	100.0	100.0
10-14	98.6	100.0	100.0	100.0	100.0	100.0	100.0
15-19	98.5	100.0	100.0	100.0	100.0	100.0	100.0
20-24	95.8	99.3	99.9	100.0	100.0	97.9	99.9
25-29	92.9	98.8	99.6	100.0	100.0	96.4	99.6
30-34	91.3	98.5	99.6	100.0	100.0	96.4	99.6
35-39	88.5	97.5	99.5	100.0	100.0	94.5	99.6
40-44	79.8	96.3	98.9	100.0	100.0	88.7	99.2
45-49	71.4	93.4	97.0	100.0	100.0	84.9	97.2
50-54	52.1	86.2	92.2	99.8	100.0	69.1	92.2
55-59	46.2	81.5	88.2	98.7	99.8	64.3	88.3
60-64	41.1	72.3	82.4	96.5	98.9	56.5	82.7
65-69	30.5	63.3	77.0	95.8	98.4	43.6	77.3
70-74	20.8	48.3	62.4	90.6	97.4	31.1	62.5
75-79	15.8	39.2	52.1	87.0	95.7	24.5	52.6
80-84	9.7	34.0	44.8	82.4	93.5	18.0	47.1
85-89	3.7	15.9	23.6	73.3	90.8	9.4	23.6
90-94	2.1	11.2	15.3	57.8	80.0	6.1	15.3
95-100	0.0	3.4	10.1	49.6	69.1	0.6	10.1

**Look-up table to convert scores to poverty likelihoods:
Old-definition Intl. 2005 PPP poverty lines**

	Poverty likelihood (%)	
	<u>Old-def. intl. 2005 PPP lines</u>	
<u>\$10.00</u>	<u>\$1.25</u>	<u>\$2.50</u>
100.0	100.0	100.0
100.0	99.7	100.0
100.0	98.6	100.0
100.0	98.5	100.0
100.0	95.8	99.9
100.0	93.9	99.6
100.0	93.4	99.6
100.0	89.1	99.6
100.0	81.0	99.2
100.0	73.9	97.5
100.0	54.6	93.0
99.8	47.5	88.5
99.1	42.7	84.1
98.8	32.2	78.2
97.8	20.9	63.7
96.4	16.5	54.1
95.8	10.8	47.4
92.6	3.7	23.9
85.4	2.1	16.2
72.5	0.0	12.9

Note on measuring changes in poverty rates over time with the old 2004/5 and new 2010/11 scorecards

This paper uses data from Malawi's 2010/11 Integrated Household Survey (IHS).

It supports three definitions of *poverty*:

- The official “government” definition for national lines in 2004/5 and 2010/11
- An “old” definition for international 2005 PPP lines used in 2004/5 that uses government-definition regional-price deflators and that has two mistakes that—for backward compatibility—are reproduced for 2010/11
- An improved “PBM” definition for 2004/5 and 2010/11 for both national lines and international 2005 PPP lines (Pauw, Beck, and Mussa, forthcoming)

The new 2010/11 scorecard here replaces the one in Schreiner (2011) that uses data from the 2004/5 IHS and supports only the government and old definitions of *poverty*. The new 2010/11 scorecard should be used from now on.

Some organizations in Malawi already use the old 2004/5 scorecard. Even after switching to the new 2010/11 scorecard, these legacy users can still estimate hybrid changes in poverty rates over time with existing baseline estimates from the old 2004/5 scorecard and follow-up estimates from the new 2010/11 scorecard.¹ This is possible because the new 2010/11 scorecard is calibrated not only to the new PBM definition of *poverty* but also to some poverty lines under the government and old definitions of *poverty* in the 2010/11 IHS data. Given the assumption that the government- and old-definition poverty lines are properly adjusted for changes in prices between the 2004/5 and 2010/11 IHS, valid hybrid estimates of change can be found for the government

¹ See the appendix for a step-by-step guide to the calculations.

and old definitions of *poverty* with a baseline measure from the old 2004/5 scorecard and a follow-up measure from the new 2010/11 scorecard.

Furthermore, a hybrid estimate of change based on the government or old definitions of *poverty* can be spliced together with a non-hybrid estimate of change based solely on the PBM definition of *poverty* if poverty rates change at the same rate under both the government (or old) definition and the PBM definition. This is the “parallel lines” assumption.

For Malawi from 2004/5 and 2010/11, the “parallel-lines” assumption does not hold well. Indeed, PBM developed their definition of *poverty* precisely because the government definition has known problems and gives a (small) estimated change in poverty that does not square with common sense nor with other triangulations. In particular, the estimated decrease in the head-count poverty rate by the national poverty line between 2004/5 to 2010/11 is 1.7 percentage points by the government definition and 8.2 percentage points by the PBM definition.

In sum, both first-time and legacy users should use the new 2010/11 scorecard and the PBM definition of *poverty* (as well as the government definition of *poverty*) from now on. Looking forward, this establishes a baseline with the best definition of *poverty* (PBM) as well as a baseline with the definition that is most likely to be supported in the next IHS (government). Looking backward, legacy users of Malawi's old 2004/5 scorecard can salvage existing estimates to find hybrid measures of change in government-definition and old-definition poverty rates over time.

Simple Poverty Scorecard[®] Poverty-Assessment Tool Malawi

1. Introduction

The Simple Poverty Scorecard poverty-assessment tool is a low-cost way for pro-poor programs in Malawi to estimate the likelihood that a household has consumption below a given poverty line, to measure groups' poverty rates at a point in time, to track changes in groups' poverty rates over time, and to segment clients for targeted services.

The new scorecard here uses data from Malawi's 2010/11 Integrated Household Survey (IHS); it replaces the old scorecard in Schreiner (2011) that uses data from the 2004/5 IHS. For now on, only the new 2010/11 scorecard should be used. The new 2010/11 scorecard can estimate a household's poverty likelihood by any or all of three definitions of *poverty*:

- The “government” definition for national lines in 2004/5 and 2010/11 and—by treating the government-definition lines as regional-price deflators—for international 2005 PPP lines
- An “old” definition for international 2005 PPP lines used in 2004/5 that uses government-definition regional-price deflators and that has two mistakes that are reproduced here for 2010/11 for backward compatability
- An improved “PBM” definition for 2004/5 and 2010/11 for national lines and—via its regional-price deflators—for international 2005 PPP lines (Pauw, Beck, and Mussa, forthcoming)

This means that existing users of the old 2004/5 scorecard do not have to start over from scratch; they can estimate changes in government- or old-definition poverty

rates over time with a baseline from the old 2004/5 scorecard and a follow-up from the new 2010/11 scorecard. From now on, existing users should record poverty-scoring results for both the government and PBM definitions, as it is not now known which of these—if any—will be supported for the next round of the IHS.

The direct approach to poverty measurement via consumption surveys is difficult and costly. As a case in point, Malawi’s 2010/11 IHS has 156 pages and includes several hundred items, many of which may be asked multiple times (for example, for each household member, each consumption item, each agricultural plot, or each crop). An enumerator visits a sampled household two or three times over four days, completing interviews at a rate of about one household per day (National Statistical Office, 2010a).

In comparison, the scorecard’s indirect approach is simple, quick, and low-cost. It uses ten verifiable indicators (such as “Is the (oldest) female head/spouse able to read and write in Chichewa or English?” and “What type of toilet facility does the household use?”) to get a score that is highly correlated with poverty status as measured by the exhaustive IHS survey.

The scorecard differs from “proxy-means tests” (Coady, Grosh, and Hoddinott, 2004) in that it is transparent, it is freely available,² and it is tailored to the capabilities

² The Simple Poverty Scorecard tool for Malawi is not, however, in the public domain. Copyright is held by the sponsor and by Microfinance Risk Management, L.L.C.

and purposes not of national governments but rather of local, pro-poor organizations. The feasible poverty-measurement options for local organizations are typically blunt (such as rules based on land ownership or housing quality) or subjective and relative (such as participatory wealth ranking facilitated by skilled field workers). Poverty measures from these approaches may be costly, their accuracy is unknown, and they are not comparable across places, organizations, nor time.

The scorecard can be used to measure the share of a program’s participants who are below a given poverty line, for example, the Millennium Development Goals’ line of \$1.25/day at 2005 purchase-power parity (PPP). USAID microenterprise partners in Malawi can use scoring with the PBM-definition \$1.25/day 2005 PPP line to report how many of their participants are “very poor”.³ Scoring can also be used to measure net movement across a poverty line over time. In all these applications, the scorecard provides a consumption-based, objective tool with known accuracy. While consumption surveys are costly even for governments, some local pro-poor organizations may be able to implement a low-cost poverty-assessment tool to help with monitoring poverty and (if desired) segmenting clients for targeted services.

³ USAID defines a household as *very poor* if its daily per-capita consumption is less than the highest of the PBM-definition \$1.25/day line—MWK133.90 in average prices for all of Malawi in February/March 2010—or the line (MWK63.65) that marks the poorest half of people below 100% of the PBM-definition national line (Figure 1). USAID (2014, p. 8) has approved the scorecard—when re-branded as a Progress Out of Poverty Index[®]—for use by its microenterprise partners.

The statistical approach here aims to be understood by non-specialists. After all, if managers are to adopt the scorecard on their own and apply it to inform their decisions, then they must first trust that it works. Transparency and simplicity build trust. Getting “buy-in” matters; proxy-means tests and regressions on the “determinants of poverty” have been around for three decades, but they are rarely used to inform decisions by local, pro-poor organizations. This is not because they do not work, but because they are often presented (when they are presented at all) as tables of regression coefficients incomprehensible to non-specialists (with cryptic indicator names such as “LGHHSZ_2” and with points with negative values and many decimal places). Thanks to the predictive-modeling phenomenon known as the “flat maximum”, simple, transparent scoring approaches are usually about as accurate as complex, opaque ones (Schreiner, 2012a; Caire and Schreiner, 2012).

Beyond its simplicity and transparency, the scorecard’s technical approach 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 the 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 scorecards.

The scorecard is based on data from the 2010/11 IHS from Malawi’s National Statistical Office (NSO). 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
- Applicable in all regions of Malawi

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 about ten minutes.

The scorecard can be used to estimate three basic quantities. First, it can estimate a particular household's *poverty likelihood*, that is, the probability that the household has per-capita consumption below a given poverty line.

Second, the scorecard can estimate the poverty rate of a group of households at a point in time. This estimate is the average of poverty likelihoods among the households in the group.

Third, the scorecard can estimate changes in the poverty rate between two points in time. With two independent samples from the same population, this estimate is the change in the average poverty likelihood in the baseline group versus the average likelihood in the follow-up group. With one sample in which each household is scored twice, this estimate is the average of each household's change from baseline to follow-up (Schreiner, 2015).

The scorecard can also be used to segment participants for targeted services. To help managers choose appropriate targeting cut-offs 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 with the PBM definition of *poverty* applied to data from the 2010/11 IHS. Scores from

this one scorecard are calibrated with data from the 2010/11 IHS to poverty likelihoods for 22 poverty lines:⁴

- Five PBM-definition national lines
- Five PBM-definition international 2005 PPP lines
- Five government-definition national lines (two of which are also supported by the old 2004/5 scorecard)
- Five government-definition international 2005 PPP lines
- Two old-definition international 2005 PPP lines (both of which are also supported by the old 2004/5 scorecard)

The new 2010/11 scorecard is constructed using half of the data from the 2010/11 IHS. That same half of the 2010/11 data is also used to calibrate scores to poverty likelihoods for all three definitions of *poverty*. The other half of the 2010/11 IHS data is used to validate the scorecard's accuracy for estimating households' poverty likelihoods, for estimating groups' poverty rates at a point in time, and for segmenting clients. Furthermore, the accuracy of estimates of changes in poverty rates over time is tested using the validation sample from the 2010/11 IHS (baseline) and all the data from the 2004/5 IHS (follow-up).

All three scoring-based estimators (the poverty likelihood of a household, the poverty rate of a group of households at a point in time, and the change in the poverty rate between two points in time) are *unbiased*. That is, they match the true value on average in repeated samples when constructed from (and applied to) a single, unchanging population in which the relationship between scorecard indicators and poverty is unchanging. Like all predictive models, the scorecard here is constructed from

⁴ Section 2 below discusses the three definitions of *poverty* and the 22 poverty lines.

a single sample and so misses the mark to some unknown extent when applied (in this paper) to validation samples. Furthermore, it is biased when applied (in practice) to a different population or when applied before or after 2010/11 (because the relationships between indicators and poverty change over time).⁵

Thus, while the indirect scoring approach is less costly than the direct survey approach, it is also biased when applied in practice. (The survey approach is unbiased by definition.) There is bias because the scorecard necessarily assumes that future relationships between indicators and poverty in all possible groups of households will be the same as in the construction data. Of course, this assumption—inevitable in predictive modeling—holds only partly.

On average across 1,000 bootstraps of $n = 16,384$ from the 2010/11 validation sample, the difference between scorecard estimates of groups' poverty rates versus the true rates at a point in time for the PBM-definition national poverty line is -1.0 percentage point. Across all 22 poverty lines under all three definitions of *poverty*, the average absolute difference is about 0.6 percentage points, and the maximum absolute difference is 1.4 percentage points. These differences reflect estimation errors due to sampling variation, not bias; the average difference would be zero if the whole 2010/11 IHS survey was to be repeatedly re-fielded and divided into sub-samples before repeating the entire process of constructing and validating scorecards.

⁵ Important cases include nationally representative samples at a later point in time or sub-groups that are not nationally representative (Diamond *et al.*, 2016; Tarozzi and Deaton, 2009).

With $n = 16,384$, the 90-percent confidence intervals are ± 0.6 percentage points or less across all poverty lines under all definitions. For $n = 1,024$, the 90-percent intervals are ± 2.6 percentage points or less.

To check the accuracy of estimates of changes in poverty rates over time, the new 2010/11 scorecard is applied to data from the 2010/11 validation sample (as a baseline) and to all the data from the 2004/5 IHS (as a follow-up).

Across 1,000 bootstraps with $n = 16,384$, the average absolute error across 20 poverty lines for estimates of change is about 2.2 percentage points. For comparison, the average absolute true change is about 4.1 percentage points. A given estimate's 90-percent confidence interval (with $n = 1,024$) includes the true value for 14 of 20 lines. The estimated direction of change is correct for 16 of 20 lines.⁶ Finally, the estimated direction is correct and "statistically significant" (its 90-percent confidence interval with $n = 1,024$ does not include zero) for 14 of 20 lines.⁷

The largest errors are for the government- and PBM-definition food lines (the lowest lines). These errors are consistent with the possibility that income increased for many of households in Malawi from 2004/5 to 2010/11 near the food lines but that the additional income served not to increase current consumption but rather to improve

⁶ The exceptions are the highest lines with household poverty in excess of 95 percent.

⁷ The additional two exceptions are the lowest poverty lines.

homes and to acquire durable assets (which increase future consumption).⁸ That is, while both the government- and the improved PBM-definition food lines show an unexpected increase in the poverty rate by the food line, the scorecard suggests that long-term quality-of-life nevertheless improved for the poorest.

Section 2 below documents data and poverty lines. Sections 3 and 4 describe scorecard construction and offer guidelines for use in practice. Sections 5 and 6 tell how to estimate households' poverty likelihoods and groups' poverty rates at a point in time. Section 7 discusses estimating changes in poverty rates over time. Section 8 covers targeting. Section 9 places the scorecard here in the context of related exercises for Malawi. The last section is a summary.

The appendix gives step-by-step instructions for how to compute hybrid estimates of change with government- and old-definition poverty lines that combine a baseline from the old 2004/5 scorecard and a follow-up from the new 2010/11 scorecard.

The “Guidelines for the Interpretation of Indicators” tells how to ask questions (and how to interpret responses) so as to mimic practice in Malawi’s IHS as closely as possible. These “Guidelines” (and the “Back-page Worksheet”) are integral parts of the Simple Poverty Scorecard tool.

⁸ It is also consistent with the possibility that the relationship between scorecard indicators and poverty differs for households near the food line vis-à-vis households near 100% of the PBM-definition national line (the line used to construct the scorecard).

2. Data, definitions of *poverty*, and poverty lines/rates

This section discusses the data used to construct and validate the scorecard. It also documents the three definitions of *poverty* used here and the 22 poverty lines to which scores are calibrated.

2.1 Data

Indicators and points for the new 2010/11 scorecard are selected (*constructed*) based on a random half of the data from the 12,271 households in the 2010/11 IHS, Malawi's most recent national consumption survey.

The half of the 2010/11 data that is used in scorecard construction is also used to associate (*calibrate*) scores to poverty likelihoods for all poverty lines under the three definitions of *poverty*.

To test the accuracy and precision of scorecard estimates, data from two validation samples are used:

- The half of the 2010/11 IHS not used in construction/calibration
- All 11,280 households in the 2004/5 IHS

Fieldwork for the 2010/11 IHS ran from 21 March 2010 to 20 March 2011.

Consumption is in MWK in average prices for Malawi as a whole as of February/March 2010.

For the 2004/5 IHS, fieldwork ran from March 2004 to April 2005, and consumption is in average prices for Malawi as of February/March 2004.

2.2 Poverty rates at the household, person, or participant level

A *poverty rate* is the share of units in households in which total household consumption (divided by the number of household members) is below a given poverty line. The unit of analysis is either the household itself or a person in the household. Each household member has the same poverty status (or estimated poverty likelihood) as the other household members.

To illustrate, suppose a program serves two households. The first household is poor (its per-capita consumption is less than a given poverty line), and it has three members, one of whom is a program participant. The second household is non-poor and has four members, two of whom are program participants.

Poverty rates are in terms of either households or people. If the program defines its *participants* as households, then the household level is relevant. The estimated household-level poverty rate is the weighted⁹ average of poverty statuses (or estimated poverty likelihoods) across households with participants. This is

$$\frac{1 \cdot 1 + 1 \cdot 0}{1 + 1} = \frac{1}{2} = 0.5 = 50 \text{ percent.}$$

In the “1 · 1” term in the numerator, the first “1” is the first household’s weight, and the second “1” is the first household’s poverty status (poor). In the “1 · 0” term in the numerator, the “1” is the second household’s weight, and the “0” is the second household’s poverty status (non-poor). The “1 + 1” in the

⁹ The example here assumes simple random sampling at the household level. This means that each household has the same weight, taken here to be one (1).

denominator is the sum of the weights of the two households. Household-level weights are used because the unit of analysis is the household.

Alternatively, a person-level rate is relevant if a program defines all people in households that benefit from its services as *participants*. In the example here, the person-level rate is the household-size-weighted¹⁰ average of poverty statuses for households with participants, or $\frac{3 \cdot 1 + 4 \cdot 0}{3 + 4} = \frac{3}{7} = 0.43 = 43$ percent. In the “3 · 1” term in the numerator, the “3” is the first household’s weight because it has three members, and the “1” is its poverty status (poor). In the “4 · 0” term in the numerator, the “4” is the second household’s weight because it has four members, and the zero is its poverty status (non-poor). The “3 + 4” in the denominator is the sum of the weights of the two households. A household’s weight is its number of members because the unit of analysis is the household member.

As a final example, a program might count as *participants* only those household members with whom it deals with directly. For the example here, this means that some—but not all—household members are counted. The person-level rate is now the participant-weighted average¹¹ of the poverty statuses of households with participants,

or $\frac{1 \cdot 1 + 2 \cdot 0}{1 + 2} = \frac{1}{3} = 0.33 = 33$ percent. The first “1” in the “1 · 1” in the numerator is the

¹⁰ Given simple random sampling, a household’s person-level weight is the number of people in the household.

¹¹ Given simple random sampling, a household’s participant-level weight is the number of participants in the household.

first household’s weight because it has one participant, and the second “1” is its poverty status (poor). In the “ $2 \cdot 0$ ” term in the numerator, the “2” is the second household’s weight because it has two participants, and the zero is its poverty status (non-poor). The “ $1 + 2$ ” in the denominator is the sum of the weights of the two households. Each household’s weight is its number of participants because the unit of analysis is the participant.

To sum up, estimated poverty rates are weighted averages of households’ poverty statuses (or estimated poverty likelihoods), where—assuming simple random sampling—the weights are the number of relevant units in the household. When reporting, organizations should make explicit the unit of analysis—household, household member, or participant—and explain why that unit is relevant.

Figure 1 reports poverty lines and poverty rates for households and people in the 2004/5 and 2010/11 IHS for Malawi as a whole, for each of Malawi’s four poverty-line regions, and for the construction/calibration and validation sub-samples.¹² Household-level poverty rates are reported because—as shown above—household-level poverty likelihoods can be straightforwardly converted into poverty rates for other units of analysis. This is also why the scorecard is constructed, calibrated, and validated with

¹² Figure 1, 8, and 9 have five versions. The first has PBM-definition national lines as well as the line that marks the poorest half of people below 100% of the PBM-definition national line. The second has PBM-definition 2005 PPP lines. The third has government-definition national lines as well as the line that marks the poorest half of people below 100% of the government-definition national line. The fourth has government-definition 2005 PPP lines. The fifth has the two old-definition lines.

household weights. Person-level poverty rates are also included in Figure 1 because these are the rates reported by the government of Malawi and by Pauw, Beck, and Mussa (PBM, forthcoming). Furthermore, person-level rates are usually used in policy discussions.

For the PBM definition in 2004/5 and 2010/11 (Figure 1), the all-Malawi person-level poverty rates for the food line (17.1 and 17.9 percent) and for 100% of the national line (47.0 and 38.8 percent) match those in PBM.

Likewise, person-level poverty rates in 2004/5 and 2010/11 for the government-definition food line (22.3 and 24.5 percent) and for 100% of the government-definition national line (52.4 percent and 50.7 percent) match those in NSO (2012, pp. 206 and 210).

2.3 Definitions of *poverty*

Poverty is whether a household is poor or non-poor. In Malawi, poverty status is determined by whether per-capita aggregate household consumption is below a given poverty line. Thus, a definition of *poverty* has two aspects: a measure of aggregate household consumption, and a poverty line.

2.3.1 Government

Following the cost-of-basic-needs approach (Ravallion, 1998), the government-definition national poverty line for the 2004/5 IHS is defined as the sum of a food

component and a non-food component. The food line¹³ is the cost of 2,400 Calories from the food basket consumed in the 2004/5 IHS by people in the fifth and sixth deciles of the distribution of per-capita aggregate household consumption (World Bank, 2005).¹⁴ In all-Malawi average prices in February/March 2004, this is MWK27.25 per person per day (Figure 1). NSO updates this line to prices as of February/March 2010 for use with the 2010/11 IHS by multiplying by a factor of 2.148, giving MWK58.52.¹⁵

The government-definition national line¹⁶ is then this food line, plus a non-food component defined as a weighted average¹⁷ of the non-food consumption of the ten percent of people in the 2004/5 IHS whose food consumption is centered on the food line. In prices as of February/March 2004, this government-definition national (food-plus-non-food) line is MWK43.92 per person per day (Figure 1). Like the government-definition food line, the government-definition national line is taken to prices for February/March 2004 by multiplying by 2.148 (MWK94.33).

¹³ NSO calls this the “ultra poverty line”, and PBM call it the “extreme poverty line”.

¹⁴ The government definition adjusts for differences in cost-of-living across four poverty-line regions, but it uses the same food basket in all four regions.

¹⁵ NSO believes that this is the inflation faced by the poor (PBM, p. 8). It comes not from data in the 2004/5 and 2010/11 IHS but rather from a (major) revision of the official all-Malawi CPI. The factor in PBM (2.289) differs from the 2.148 here because this paper adjust poverty lines—rather than consumption—for regional cost-of-living differences and because the person-weighted average regional deflators are not 1.0 but rather 0.9916 (in 2004/5) and 0.9305 (2010/11).

¹⁶ NSO calls this “the poverty line”, and PBM call it the “normal poverty line”.

¹⁷ Weights are greater for people whose food consumption is closer to the food line.

NSO (2012, p. 203)¹⁸ treats the government-definition of *poverty* in the 2004/5 IHS and in the 2010/11 IHS as the same, comparing their poverty-rate estimates across the two surveys without caveats. The estimated decrease in the six years between IHS rounds in the person-level poverty rate by 100% of the government-definition national line is $52.4 - 50.7 = 1.7$ percentage points.

2.3.2 PBM

Pauw, Beck, and Mussa (forthcoming) develop their definition of *poverty* because the government-definition estimate of the decrease in poverty in Malawi seems too low. Drilling down, the urban poverty rate by the government definition decreased by 8.1 percentage points, while rural poverty increased by 0.7 percentage points. This seems odd because (PBM, forthcoming):

- Per-capita growth in the six-year period was rapid (about 3.5 percent/year)
- Large-scale fertilizer subsidies and good weather doubled maize yields
- A scorecard (Mathiassen, 2006) applied to Welfare Monitoring Surveys between IHS rounds estimated a large fall in poverty rates (NSO, 2010b)
- Subjective assessments of well-being improved a lot between the two IHS rounds

¹⁸ “The [poverty] methodology replicates as much as possible that employed in the poverty analysis of the 2004/05 IHS in order to guarantee comparability over time.”

Beyond these cross-checks, PBM report a few technical issues in government-definition consumption and poverty lines:

- Poverty-line deflators differ from Malawi’s official Consumer Price Index (CPI) by a factor of about 1.5¹⁹
- This single deflator is applied in all regions and for both food and non-food despite evidence that price changes over time are not uniform in these dimensions
- Conversion factors for non-metric units of food items are off by factors of:
 - About 10 for sachets of cooking oil
 - About 5.5, 21, and 7.4 in the North region for cassava, dried fish, and fresh fish (together accounting for more than one-third of Calories in the North)

The measure of consumption in the PBM definition of *poverty* uses better conversion factors. To improve the definition of the food basket, PBM also:

- Derive food baskets and caloric requirements for each poverty-line region
- Ensure that the food baskets provide consistent utility to people in different regions and times (Arndt and Simler, 2010)
- Allow the Caloric value of a region’s food basket to vary more closely with its demographic composition
- Derive a region’s food-basket reference group simultaneously with its poverty line (which depends on the food basket and which determines the reference group) (Pradhan *et al.*, 2001)
- Derives regional and temporal price deflators not from the all-Malawi CPI but rather from the 2004/5 and 2010/11 IHS

The PBM-definition food line is MWK26.02 per person per day in 2004/5 and 59.59 in 2010/11, corresponding to person-level poverty rates of 17.1 and 17.9 percent (Figure 1). This is an increase of 0.8 percentage points (versus an increase of 2.2

¹⁹ The NSO deflates consumption—not poverty lines—by poverty-line region. This paper instead deflates poverty lines and leaves consumption in nominal units. This does not affect poverty status nor estimated poverty rates, and it makes cost-of-living adjustments more transparent to non-specialists. Because the person-weighted average of regional-price deflators is not exactly 1.0, this difference leads to PBM reporting a factor of 1.7 rather than 1.5.

percentage points for the government-definition food line). PBM (p. 33) conclude that “economic policies in Malawi appear to have neglected the ultra-poor.” As noted above, however, scoring’s estimate of change over time for poverty by the food line (regardless of definition) is consistent with increased income for households who started below the food line being translated into home improvements and asset acquisition rather than additional current consumption.

PBM derive the non-food component in their cost-of-basic needs approach (separately for each poverty-line region and survey round) as the average²⁰ non-food expenditure of the 20 percent of people whose total (food-plus-non-food, not just food) consumption is centered on the food line.

The PBM-definition national (food-plus-non-food) line for Malawi as a whole is MWK43.23 per person per day in 2004/5 and MWK92.82 in 2010/11, giving person-level poverty rates of 47.0 and 38.8 percent (Figure 1). This is a decrease of 8.2 percentage points (versus a decrease of 1.7 percentage points for the government-definition national line). The larger decrease by the PBM definition fits evidence from non-IHS sources better.

²⁰ People closer to the food line are assigned greater weight.

2.3.3 Old

Schreiner (2011) documents the old-definition \$1.25/day and \$2.50/day 2005

PPP poverty lines for the 2004/5 IHS. The lines have two errors:

- They take the person-weighted average of the government-definition regional price deflators in the 2004/5 IHS as 1.0, rather than 0.9916
- They use average prices from March 2004 to March 2005 rather than as of February/March 2004

For compatibility with legacy estimates, scores from the new 2010/11 scorecard here are calibrated to these lines, without fixing their errors. The 2010/11 old-definition lines (MWK137.74 and MWK275.47) are the 2004/5 lines (MWK63.60 and MWK127.20), updated with a factor of 2.148 to average prices in all of Malawi as of February/March 2010 with the government-definition deflator of 2.148. They are then adjusted for regional differences in cost-of-living in 2010/11 using government-definition deflators and accounting for the fact that these deflators' person-weighted average is 0.9305.

These two old-definition lines have known errors, so they are only for use by legacy users who want to estimate changes in poverty over time with a baseline with these lines from the old 2004/5 scorecard and a follow-up from the new 2010/11 scorecard. Other users should not use old-definition lines. Instead, they should use government-definition lines or PBM-definition lines.

2.4 Supported poverty lines

Because pro-poor organizations in Malawi may want to use different or various poverty lines, this paper calibrates scores from its single new 2010/11 scorecard to poverty likelihoods for 22 lines:

- PBM-definition:
 - Food
 - 100% of national
 - 150% of national
 - 200% of national
 - Line marking the poorest half of people below 100% of the PBM-definition national line
 - \$1.25/day 2005 PPP
 - \$2.00
 - \$2.50
 - \$5.00
 - \$8.44
- Government-definition:
 - Food
 - 100% of national
 - 150% of national
 - 200% of national
 - Line marking the poorest half of people below 100% of the government-definition national line
 - \$1.25/day 2005 PPP
 - \$2.00
 - \$2.50
 - \$5.00
 - \$8.44
- Old-definition:
 - \$1.25/day 2005 PPP
 - \$2.50

For a given definition of *poverty*, the lines for 150% and 200% of national are multiples of the national line.

For a given definition of *poverty*, the line that marks the poorest half of people below 100% of the national line is defined—separately in each of Malawi’s four poverty-line regions in a given IHS round—as the median aggregate household per-capita consumption of people (not households) below 100% of the national line (U.S. Congress, 2004).

Both the PBM- and government-definition \$1.25/day lines use the 2005 PPP factor of MWK56.922 per USD1 (World Bank, 2008). They also use the same average CPI for 2005 (197.204) and for February/March 2004 (172.850).²¹ The price deflator from February/March 2004 to February/March 2010 is taken as the ratio of the national poverty line between the two rounds ($92.82 \div 43.23 = 2.147$ for the PBM definition, and $94.33 \div 43.92 = 2.148$ for the government definition, Figure 1).

Under both the PBM definition and the government definition, the \$1.25/day 2005 PPP poverty line in average prices in Malawi overall on average in February/March 2004 is (Sillers, 2006):

$$1.25 \cdot \text{PPP} \cdot \left(\frac{\text{CPI}_{\text{Feb/Mar '04}}}{\text{CPI}_{2005}} \right) = 1.25 \cdot 56.922 \cdot \left(\frac{172.850}{197.204} \right) = \text{MWK}62.37.$$

For 2010/11, the PBM-definition \$1.25/day line in February/March 2010 is $62.37 \cdot 2.147 = \text{MWK}133.90$. For the government definition, \$1.25/day in February/March 2010 is almost the same ($62.37 \cdot 2.148 = \text{MWK}133.98$)²²

²¹ rbm.mw/inflation_rates_detailed.aspx, retrieved 30 July 2015.

²² This differs by two *tambala* from the MWK133.96 in Figure 1 due to rounding in the presentation in the text that does not occur in the actual calculations.

Although the all-Malawi \$1.25/day 2005 PPP lines are the same for both the PBM and government definitions in a given IHS round, the regional values of the \$1.25/day lines differ because the two definitions use different regional-price deflators. In particular, for a given definition of *poverty* in a given poverty-line region, the \$1.25/day line is the all-Malawi \$1.25/day line for that definition, multiplied by the national poverty line for that region and definition, divided by the average all-Malawi national line for that definition.

For example, the PBM-definition \$1.25/day 2005 PPP line in 2010/11 in the North Rural poverty-line region is the all-Malawi PBM-definition \$1.25/day line of MWK133.90, multiplied by 100% of the PBM-definition national line in North Rural of MWK95.90, divided by the average all-Malawi national line of MWK92.82. This gives a PBM-definition \$1.25/day line in the North Rural poverty-line region of $133.90 \times 95.90 \div 92.82 = \text{MWK}138.34$ (Figure 1).

Likewise, the government-definition \$1.25/day line in 2010/11 in North Rural is the all-Malawi government-definition \$1.25/day line of MWK133.96 (Figure 1), multiplied by 100% of the government-definition national line in North Rural of MWK104.55, divided by the average all-Malawi government-definition national line of MWK94.33. This gives a government-definition \$1.25/day line in North Rural of $133.96 \times 104.55 \div 94.33 = \text{MWK}148.47$ (Figure 1).

For a given definition of *poverty*, the \$2.00, \$2.50, \$5.00, and \$8.44/day lines are multiples of the \$1.25/day line. The \$8.44/day line is the 75th percentile of world-wide per-capita income (not consumption) as measured by Hammond *et al.* (2007).

The old-definition \$1.25 and \$2.50/day 2005 PPP lines are the 2004/5 lines from Schreiner (2011), updated to average prices in all of Malawi as of February/March 2010 with the government-definition temporal deflator of 2.148 and adjusted for regional differences in cost-of-living in 2010/11 using government-definition regional-price deflators and accounting for the fact that these deflators' person-weighted average is 0.9305.

For 2004/5 and 2010/11, the World Bank's PovcalNet²³ reports \$1.25/day 2005 PPP person-level rates of 75.0 and 72.2 percent. This compares with 73.2 and 68.8 percent (government definition) and 69.0 and 58.7 percent (PBM definition, Figure 1).

Because PovcalNet does not document its \$1.25/day 2005 PPP line in MWK nor its derivation, the PBM-definition \$1.25/day estimates here are to be preferred (Schreiner, 2014).

²³ iresearch.worldbank.org/PovcalNet/index.htm, retrieved 30 July 2015.

2.5 The USAID “very poor” poverty line

USAID microenterprise partners in Malawi who use the scorecard to report poverty rates to USAID should use the PBM-definition \$1.25/day 2005 PPP line. This is because USAID defines the “very poor” as those people in households whose per-capita consumption is below the highest of the following poverty lines:

- The line that marks the poorest half of people below 100% of the PBM-definition national line (MWK63.65 per person per day in 2010/11, with a person-level poverty rate of 19.4 percent, Figure 1)
- PBM-definition \$1.25/day 2005 PPP (MWK133.90, person-level rate of 58.7 percent)

2.6 “Parallel-lines” assumption

If the “parallel-lines” assumption holds, then it is valid to splice together two estimates of change over time in which the follow-up estimate of change is a non-hybrid (using PBM-definition poverty lines in both a baseline and a follow-up from the new 2010/11 scorecard) and in which the baseline estimate of change is a hybrid (using government- or old-definition poverty lines with a baseline from the old 2004/5 scorecard and a follow-up from the new 2010/11 scorecard).

The “parallel lines” assumption is that *changes* in poverty rates over time are the same regardless of the definition of *poverty*, even though the *levels* of the estimates at a point in time may differ by the definition of *poverty*.

For Malawi, the “parallel lines” assumption can be checked; between 2004/5 and 2010/11, the person-level poverty rate decreased by (Figure 1):

- 1.7 percentage points for 100% of the government-definition national line
- 8.2 percentage points for 100% of the PBM-definition national line

Thus, the “parallel-lines” assumption does not hold well from 2004/5 to 2010/11.

Of course, it may hold worse (or better) in the future, but the known differences in the definitions of *poverty* do not give reason to hope for improvement. Furthermore, if the “parallel lines” assumption does not hold well in the past, then it is more likely to not hold well in the future than if it did hold well in the past.

3. Scorecard construction

For Malawi, about 80 candidate indicators are initially prepared in the areas of:

- Household composition (such as the number of members)
- Education (such as the literacy of the (oldest) female head/spouse)
- Housing (such as the type of floor)
- Ownership of durable assets (such as tables or beds)
- Employment (such as whether the male head/spouse works)
- Agriculture (such as the ownership of goats)

Figure 2 lists the candidate indicators, ordered by the entropy-based “uncertainty coefficient” (Goodman and Kruskal, 1979) that measures how well a given indicator predicts poverty status on its own.²⁴

One possible application of the scorecard is to measure *changes* in poverty through time. Thus, when selecting indicators and holding other considerations constant, preference is given to more sensitive indicators. For example, the ownership of a table is probably more likely to change in response to changes in poverty than is the age of the male head/spouse.

The scorecard itself is built using 100% of the PBM-definition national poverty line and Logit regression on the 2010/11 construction sub-sample. Indicator selection uses both judgment and statistics. The first step is to use Logit to build one scorecard for each candidate indicator. Each scorecard’s power to rank households by poverty status is measured as “c” (SAS Institute Inc., 2004).

²⁴ The uncertainty coefficient is not used as a criterion when selecting scorecard indicators; it is just a way to order the candidate indicators in Figure 2.

One of these one-indicator scorecards is then selected based on several factors (Schreiner *et al.*, 2014; Zeller, 2004). These include improvement in accuracy, likelihood of acceptance by users (determined by simplicity, cost of collection, and “face validity” in terms of experience, theory, and common sense), sensitivity to changes in poverty, variety among indicators, applicability across regions, tendency to have a slow-changing relationship with poverty over time, relevance for distinguishing among households at the poorer end of the distribution of consumption, and verifiability.

A series of two-indicator scorecards are then built, each adding a second indicator to the one-indicator scorecard selected from the first round. The best two-indicator scorecard is then selected, again using judgment to balance “c” with the non-statistical criteria. These steps are repeated until the scorecard has 10 indicators that work well together.²⁵

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).

²⁵ For Malawi, the selection of the final 10 indicators was also informed by feedback from a field test by MicroLoan Foundation.

This algorithm is similar to common R^2 -based stepwise least-squares regression. It differs from naïve stepwise in that the selection of indicators considers both statistical²⁶ and non-statistical criteria. The use of non-statistical criteria can improve robustness through time and helps ensure that indicators are simple, sensible, and acceptable to users.

The single scorecard here applies to all of Malawi. Tests for Indonesia (World Bank, 2012), Bangladesh (Sharif, 2009), India and Mexico (Schreiner, 2006 and 2005a), Sri Lanka (Narayan and Yoshida, 2005), and Jamaica (Grosh and Baker, 1995) suggest that segmenting scorecards by urban/rural does not improve targeting accuracy much.²⁷ In general, segmentation may improve the accuracy of estimates of poverty rates (Diamond *et al.*, 2016; Tarozzi and Deaton, 2009), but it may also increase the risk of overfitting (Haslett, 2012).

²⁶ The statistical criterion for selecting an indicator is not the p values of its coefficients but rather the indicator's contribution to the ranking of households by poverty status.

²⁷ See Section 9 for an example of this result in Malawi.

4. Practical guidelines for scorecard use

The main challenge of scorecard design is not to maximize statistical accuracy but rather to improve the chances that the scorecard is actually used (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 train and convince its employees to use the scorecard properly (Schreiner, 2002). After all, most reasonable scorecards have similar targeting accuracy, thanks to the empirical phenomenon known as the “flat maximum” (Caire and Schreiner, 2012; Hand, 2006; Baesens *et al.*, 2003; Lovie and Lovie, 1986; Kolesar and Showers, 1985; Stillwell, Barron, and Edwards, 1983; Dawes, 1979; Wainer, 1976; Myers and Forgy, 1963). The bottleneck is less technical and more human, not statistics but organizational-change management. Accuracy is easier to achieve than adoption.

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

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

- Only 10 indicators
- Only "multiple-choice" indicators
- Only simple points (non-negative integers, and no arithmetic beyond addition)

A field worker using Malawi's new 2010/11 scorecard would:

- Record the interview identifier, interview date, county code ("MWI"), scorecard code ("002") and the sampling weight assigned by the survey design to the household of the participant
- Record the names and identifiers of the participant (who may not be the same as the respondent), field agent, and relevant organizational service point
- Complete the back-page worksheet with each household member's first name
- Record household size in the scorecard header next to "Number of household members:", and record the response to the first scorecard indicator based on the number of household members listed on the back-page worksheet
- Read the second question, drawing a circle around the relevant response and its points, and writing the point value in the far right-hand column
- For the third, fourth, and fifth indicators, observe the predominant construction material of floor, outer wall, and roof the household's main dwelling, drawing a circle around the relevant response and its points, and writing the point value in the far right-hand column. These questions are not asked of the respondent unless the enumerator cannot identify the predominant material on his/her own with certainty
- Read each of the remaining five questions one-by-one from the scorecard, drawing a circle around the relevant responses and their points, and writing each point value in the far right-hand column
- Add up the points to get a total score
- Implement targeting policy (if any)
- Deliver the paper scorecard to a central office for data entry and filing

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 managers or funders reward them for higher poverty rates), then it is wise to do on-going quality

control via data review and random audits (Matul and Kline, 2003).²⁸ IRIS Center (2007a) 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 alternative ways of measuring poverty, it is still absolutely difficult. Training and explicit definitions of terms and concepts in the scorecard are essential, and field workers should scrupulously study and follow the “Guidelines for the Interpretation of Indicators” found at the end of this paper, as the “Guidelines”—along with the “Backpage Worksheet”—are an integral part of the Simple Poverty Scorecard tool.²⁹

For the example of Nigeria, one study (Onwujekwe, Hanson, and Fox-Rushby, 2006) found distressingly low inter-rater and test-retest correlations for indicators as seemingly simple as whether the household owns an automobile. At the same time, Grosh and Baker (1995) suggest that gross underreporting of assets does not affect targeting. For the first stage of targeting in a conditional cash-transfer program in

²⁸ If a program does not want field workers and respondents to know the points associated with responses, then it can use a version of the scorecard that does not display the points and then apply the points and compute scores later at a central office. Even if points are hidden, field workers and respondents can apply common sense to guess how response options are linked with poverty. Schreiner (2012b) argues that hiding points in Colombia (Camacho and Conover, 2011) did little to deter cheating and that, in any case, cheating by the user’s central office was more damaging than cheating by field workers and respondents.

²⁹ The guidelines here are the only ones that organizations should give to field workers. All other issues of interpretation should be left to the judgment of field workers and respondents, as this seems to be what Malawi’s NSO does in the IHS.

Mexico, Martinelli and Parker (2007, pp. 24–25) find that “underreporting [of asset ownership] is widespread but not overwhelming, except for a few goods . . . [and] overreporting is common for a few goods”. Still, as is done in Mexico in the second stage of its targeting process, most false self-reports can be corrected (or avoided in the first place) by field workers who make a home visit. This is the recommended procedure for local, pro-poor organizations who use scoring for targeting in Malawi.

In terms of implementation and 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 questions to be answered, and the budget. The main goal should be to make sure that the sample is representative of a well-defined population and that the scorecard will inform an issue that matters to the organization.

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

- Employees of the organization
- Third parties

Responses, scores, and poverty likelihoods can be recorded on:

- Paper in the field, and then filed at a central office
- Paper in the field, and then keyed into a database or spreadsheet at a central office
- Portable electronic devices in the field, and then uploaded to a database

Given a population of participants relevant for a particular business question, the participants to be scored can be:

- All relevant participants (a census)
- A representative sample of relevant participants
- All relevant participants in a representative sample of relevant field offices
- A representative sample of relevant participants in a representative sample of relevant field offices

If not determined by other factors, the number of participants to be scored can be derived from sample-size formulas (presented later) to achieve a desired confidence level and a desired confidence interval. The focus, however, should not be on having a sample size large enough to achieve some arbitrary level of statistical significance but rather on having a representative sample from a well-defined population so that the analysis of the results can have a chance to meaningfully inform questions that matter to the organization.

The frequency of application can be:

- As a once-off project (precluding measuring change)
- Every two years (or at any other fixed or variable time interval, allowing measuring change)
- Each time a field worker visits a participant at home (allowing measuring change)

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

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

An example set of choices is illustrated by BRAC and ASA, two microfinance organizations in Bangladesh who each have about 7 million participants and who declared their intention to apply the Simple Poverty Scorecard tool for Bangladesh (Schreiner, 2013) with a sample of about 25,000. Their design is that all 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. They record responses on paper in the field before sending the forms to a central office to be entered into a database and converted to poverty likelihoods.

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 poor, the scores themselves have only relative units. For example, doubling the score decreases the likelihood of being below a given poverty line, but it does not cut it in half.

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 100% of the PBM-definition national line, scores of 35–39 correspond with a poverty likelihood of 47.1 percent, and scores of 40–44 correspond with a poverty likelihood of 39.6 percent (Figure 3).

The poverty likelihood associated with a score varies by poverty line. For example, scores of 35–39 are associated with a poverty likelihood of 47.1 percent for 100% of the PBM-definition national line but of 73.8 percent for the PBM-definition \$1.25/day line.³⁰

³⁰ Starting with Figure 3, many figures have 22 versions, covering the five broad types of poverty lines:

- Five PBM-definition national lines
- Five PBM-definition 2005 PPP lines
- Five government-definition national lines
- Five government-definition 2005 PPP lines
- Two old-definition 2005 PPP lines

To keep them straight, lines are grouped by type. Figures pertaining to all lines of a given type are placed with the figures for 100% of the PBM-definition national line.

5.1 Calibrating scores with poverty likelihoods

A given score is associated (“calibrated”) with a poverty likelihood by defining the poverty likelihood as the share of households in the calibration sub-sample who have the score and who have per-capita consumption below a given poverty line.

For the example of 100% of the PBM-definition national line (Figure 4), there are 9,445 (normalized) households in the 2010/11 calibration sub-sample with a score of 35–39. Of these, 4,444 (normalized) are below the poverty line. The estimated poverty likelihood associated with a score of 35–39 is then 47.1 percent, because $4,444 \div 9,445 = 47.1$ percent.

To illustrate with 100% of the PBM-definition national line and a score of 40–44, there are 10,368 (normalized) households in the 2010/11 calibration sample, of whom 4,101 (normalized) are below the line (Figure 4). The poverty likelihood for this score range is then $4,101 \div 10,368 = 39.6$ percent.

The same method is used to calibrate scores with estimated poverty likelihoods for all 22 poverty lines.³¹

Even though the scorecard is constructed partly based on judgment related to non-statistical criteria, the calibration process produces poverty likelihoods that are objective, that is, derived from quantitative poverty lines and from survey data on

³¹ To ensure that poverty likelihoods never increase as scores increase, likelihoods across series of adjacent scores are sometimes iteratively averaged before grouping scores into ranges. This preserves unbiasedness while keeping users from balking when sampling variation in score ranges with few households would otherwise lead to higher scores being linked with higher poverty likelihoods.

consumption. The calibrated poverty likelihoods would be objective even if the process of selecting indicators and points did not use any data at all. In fact, objective scorecards of proven accuracy are often constructed using only expert judgment to select indicators and points (Fuller, 2006; Caire, 2004; Schreiner *et al.*, 2014). 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 objectivity depends on using data in score calibration, not on using data (and nothing else) in scorecard construction.

Although the points in the Malawi scorecard are transformed coefficients from a Logit regression, (untransformed) 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. Going from scores to poverty likelihoods in this way requires no arithmetic at all, just a look-up table. This approach to calibration can also improve accuracy, especially with large samples.

5.2 Accuracy of estimates of households' poverty likelihoods

As long as the relationships between indicators and poverty do not change over time, and as long as the scorecard is applied to households who are representative of the same population from which the scorecard was originally constructed, then this calibration process produces unbiased estimates of poverty likelihoods. *Unbiased* means that in repeated samples from the same population, the average estimate matches the true value. Given the assumptions above, the scorecard also produces unbiased estimates of poverty rates at a point in time and unbiased estimates of changes in poverty rates between two points in time.³²

Of course, the relationships between indicators and poverty do change to some unknown extent over time and also across sub-national groups in Malawi's population. Thus, the scorecard will generally be biased when applied after March 2011 (the last month of fieldwork for the 2010/11 IHS) or when applied with sub-groups that are not nationally representative.

³² This follows because these estimates of groups' poverty rates are linear functions of the unbiased estimates of households' poverty likelihoods.

How accurate are estimates of households' poverty likelihoods, given the assumption of unchanging relationships between indicators and poverty over time and the assumption of a sample that is representative of Malawi as a whole? To find out, the scorecard is applied to 1,000 bootstrap samples of size $n = 16,384$ from the 2010/11 validation sample. Bootstrapping means to:

- Score each household in the 2010/11 validation sample
- Draw a 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 with consumption below a poverty line
- For each score, record the difference between the estimated poverty likelihood (Figure 4) and the true poverty likelihood in the bootstrap sample
- Repeat the previous three steps 1,000 times
- For each score, report the average difference between estimated and true poverty likelihoods across the 1,000 bootstrap samples
- For each score, report the two-sided intervals containing the central 900, 950, and 990 differences between estimated and true poverty likelihoods

For each score range and for $n = 16,384$, Figure 5 shows the average difference between estimated and true poverty likelihoods as well as confidence intervals for the differences.

For the example of 100% of the PBM-definition national line, the average poverty likelihood across bootstrap samples for scores of 35–39 in the 2010/11 validation sample is too low by 6.6 percentage points. For scores of 40–44, the estimate is too high by 0.6 percentage points.³³

³³ These differences are not zero, despite the estimator's unbiasedness, because the scorecard comes from a single sample from the 2010/11 IHS. The average difference by score range would be zero if the IHS was repeatedly applied to samples of the

The 90-percent confidence interval for the differences for scores of 35–39 is ± 4.4 percentage points (100% of the PBM-definition national line, Figure 5). This means that in 900 of 1,000 bootstraps, the difference between the estimate and the true value is between -11.0 and -2.2 percentage points (because $-6.6 - 4.4 = -11.0$, and $-6.6 + 4.4 = -2.2$). In 950 of 1,000 bootstraps (95 percent), the difference is -6.6 ± 4.7 percentage points, and in 990 of 1,000 bootstraps (99 percent), the difference is -6.6 ± 5.2 percentage points.

Several differences between estimated poverty likelihoods and true values in Figure 5 are large. There are differences because the validation 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 differences in the score ranges just above and below the targeting cut-off. This mitigates the effects of bias and sampling variation on targeting (Friedman, 1997). Sections 8 and 9 below look at targeting accuracy in detail.

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

population of Malawi and then split into sub-samples before repeating the entire process of scorecard construction/calibration and validation.

Another possible source of differences between estimates and true values is overfitting. The scorecard here is unbiased, but it may still be *overfit* when applied after the end of the IHS fieldwork in March 2011. That is, the scorecard may fit the data from the 2010/11 IHS so closely that it captures not only some real patterns but also some random patterns that, due to sampling variation, show up only in the 2010/11 IHS but not in the overall population of Malawi. Or the scorecard may be overfit in the sense that it is not robust when relationships between indicators and poverty change over time or when the scorecard is applied to samples that are not nationally representative.

Overfitting can be mitigated by simplifying the scorecard and by not relying only on data but rather also considering theory, experience, and judgment. Of course, the scorecard here does this. Combining scorecards can also reduce overfitting, at the cost of greater complexity.

Most errors in individual households' likelihoods do balance out in the estimates of groups' poverty rates for nationally representative samples (see the next two sections). Furthermore, at least some of the differences in change-through-time estimates may come from non-scorecard sources such as changes in the relationships between indicators and poverty, sampling variation, changes in poverty lines, inconsistencies in data quality across time, and imperfections in cost-of-living adjustments across time and across geographic regions. These factors can be addressed only by improving the availability, frequency, quantity, and quality of data from

national consumption surveys (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 an organization samples three households on 1 January 2016 and that they have scores of 20, 30, and 40, corresponding to poverty likelihoods of 77.6, 55.1, and 39.6 percent (100% of the PBM-definition national line, Figure 3). The group's estimated poverty rate is the households' average poverty likelihood of $(77.6 + 55.1 + 39.6) \div 3 = 57.4$ percent.

Be careful; the group's poverty rate is *not* the poverty likelihood associated with the average score. Here, the average score is 30, which corresponds to a poverty likelihood of 55.1 percent. This differs from the 57.4 percent found as the average of the three individual poverty likelihoods associated with each of the three scores. Unlike poverty likelihoods, scores are ordinal symbols, like letters in the alphabet or colors in the spectrum. Because scores are not cardinal numbers, they cannot meaningfully be added up or averaged across households. Only three operations are valid for scores: conversion to poverty likelihoods, analysis of distributions (Schreiner, 2012a), or comparison—if desired—with a cut-off for targeting. The safest rule to follow is: Always use poverty likelihoods, never scores.

Scores from the new 2010/11 scorecard are calibrated with data from the 2010/11 IHS for all 22 poverty lines. The process of calibrating scores to poverty likelihoods and the approach to estimating poverty rates is exactly the same for all

lines, regardless of their definition. For users, the only difference is in the specific look-up table used to convert scores to poverty likelihoods.

Existing users of the old 2004/5 scorecard who switch to the new 2010/11 scorecard can salvage existing poverty-rate estimates for measuring change over time by using the supported government- or old-definition lines to estimate poverty rates for use in hybrid estimates of change with a baseline from the old 2004/5 scorecard and a follow-up from the new 2010/11 scorecard. From now on, all users of the new 2010/11 scorecard should estimate poverty rates with both the government- and PBM-definition lines, as these are the definitions of *poverty* most likely to be supported in the next IHS. The appendix describes the process of splicing together hybrid estimates of change looking backwards and non-hybrid estimates of change going forward, as well as the assumptions required for such estimates to be valid.

6.1 Accuracy of estimated poverty rates at a point in time

For the new 2010/11 scorecard applied to 1,000 bootstraps of $n = 16,384$ from the 2010/11 validation sample and 100% of the PBM-definition national poverty line, the average difference between the estimated poverty rate at a point in time versus the true rate is -1.0 percentage points (Figure 7, summarizing Figure 6 across all poverty lines). Across the 10 PBM-definition poverty lines in the 2010/11 validation sample, the maximum absolute difference is 1.0 percentage points, and the average absolute

difference is about 0.4 percentage points. At least part of these differences is due to sampling variation in the division of the 2010/11 IHS into two sub-samples.

When estimating poverty rates at a point in time, the bias reported in Figure 7 should be subtracted from the average poverty likelihood to make the estimate unbiased. For the example of the new 2010/11 scorecard and 100% of the PBM-definition national line in the 2010/11 validation sample, bias is -1.0 percentage points, so the unbiased estimate in the three-household example above is $57.4 - (-1.0) = 58.4$ percent.

For government-definition lines, the maximum absolute difference is 1.4 percentage points, and the average absolute difference is about 0.8 percentage points (Figure 7).³⁴

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 better for all poverty lines (Figure 7). This means that in 900 of 1,000 bootstraps of this size, the estimate (after subtracting off bias) is within 0.6 percentage points of the true value.

For example, suppose that the average poverty likelihood in a sample of $n = 16,384$ with the Malawi scorecard and 100% of the PBM-definition national line is 57.4 percent. Then estimates in 90 percent of such samples would be expected to fall in the range of $57.4 - (-1.0) - 0.6 = 57.8$ percent to $57.4 - (-1.0) + 0.6 = 59.0$ percent, with

³⁴ For the two old-definition lines, the maximum absolute difference is 0.9 percentage points, and the average absolute difference is about 0.7 percentage points.

the most likely true value being the unbiased estimate in the middle of this range, that is, $57.4 - (-1.0) = 58.4$ percent. This is because the original (biased) estimate is 57.4 percent, bias is -1.0 percentage points, and the 90-percent confidence interval for 100% of the PBM-definition national line in the 2010/11 validation sample with this sample size is ± 0.6 percentage points (Figure 7).

6.2 Formula for standard errors for estimates of poverty rates

How precise are the point-in-time estimates? Because these estimates are averages, they have (in “large” samples) a Normal distribution and can be characterized by their average difference vis-à-vis true values (*bias*), together with their standard error (*precision*).

Schreiner (2008a) proposes an approach to deriving a formula for the standard errors of estimated poverty rates at a point in time from indirect measurement via poverty-assessment tools. It starts with Cochran's (1977) textbook formula of $\pm c = \pm z \cdot \sigma$ that relates confidence intervals with standard errors in the case of direct measurement of ratios, where:

$\pm c$ is a confidence interval as a proportion (*e.g.*, 0.02 for ± 2 percentage points),

z is from the Normal distribution and is $\begin{cases} 1.04 \text{ for confidence levels of 70 percent} \\ 1.28 \text{ for confidence levels of 80 percent,} \\ 1.64 \text{ for confidence levels of 90 percent} \end{cases}$

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

\hat{p} is the estimated proportion of households below the poverty line in the sample,

ϕ is the finite population correction factor $\sqrt{\frac{N - n}{N - 1}}$,

N is the population size, and

n is the sample size.

For example, Malawi's 2010/11 IHS gives a direct-measurement estimate of the household-level poverty rate for 100% of the PBM-definition national line in the 2010/11 validation sample of $\hat{p} = 32.4$ percent (Figure 1). If this estimate came from a sample of $n = 16,384$ households from a population N of 3,072,524 (the number of households in Malawi in 2010/11 according to the IHS sampling weights), then the

finite population correction ϕ is $\sqrt{\frac{3,072,524 - 16,384}{3,072,524 - 1}} = 0.9973$, which very close to $\phi =$

1. If the desired confidence level is 90-percent ($z = 1.64$), then the confidence interval

$$\pm c \text{ is } \pm z \cdot \sqrt{\frac{\hat{p} \cdot (1 - \hat{p})}{n}} \cdot \sqrt{\frac{N - n}{N - 1}} = \pm 1.64 \cdot \sqrt{\frac{0.324 \cdot (1 - 0.324)}{16,384}} \cdot \sqrt{\frac{3,072,524 - 16,384}{3,072,524 - 1}} =$$

± 0.598 percentage points. (If ϕ were taken as 1, then the interval is ± 0.600 percentage points.)

Scorecards, however, do not measure poverty directly, so this formula is not applicable. To derive a formula for the new 2010/11 Malawi scorecard, consider Figure 6, which reports empirical confidence intervals $\pm c$ for the differences for the scorecard applied to 1,000 bootstraps of various sizes from the 2010/11 validation sample. For example, with $n = 16,384$ and 100% of the PBM-definition national line in the 2010/11 validation sample, the 90-percent confidence interval is ± 0.624 percentage points.³⁵

Thus, the 90-percent confidence interval with $n = 16,384$ is ± 0.624 percentage points for the Malawi scorecard and ± 0.598 percentage points for direct measurement. The ratio of the two intervals is $0.624 \div 0.598 = 1.04$.

Now consider the same exercise, but with $n = 8,192$. The confidence interval under direct measurement and 100% of the PBM-definition national line in the 2010/11

$$\text{validation sample is } \pm 1.64 \cdot \sqrt{\frac{0.324 \cdot (1 - 0.324)}{8,192}} \cdot \sqrt{\frac{3,072,524 - 8,192}{3,072,524 - 1}} = \pm 0.847$$

percentage points. The empirical confidence interval with the Malawi scorecard (Figure

³⁵ Due to rounding, Figure 6 displays 0.6, not 0.624.

6) is ± 0.912 percentage points. Thus for $n = 8,192$, the ratio of the two intervals is $0.912 \div 0.847 = 1.08$.

This ratio of 1.08 for $n = 8,192$ is close to the ratio of 1.04 for $n = 16,384$. Across all sample sizes of 256 or more in Figure 6, these ratios are generally close to each other, and the average of these ratios in the 2010/11 validation sample turns out to be 1.03, implying that confidence intervals for indirect estimates of poverty rates via the Malawi scorecard and 100% of the PBM-definition national poverty line are—for a given sample size—about 3-percent wider than confidence intervals for direct estimates via the 2010/11 IHS. This 1.03 appears in Figure 7 as the “ α factor” because if $\alpha = 1.03$, then the formula for confidence intervals c for the Malawi scorecard is $\pm c = \pm z \cdot \alpha \cdot \sigma$.

That is, the formula for the standard error σ for point-in-time estimates of poverty rates

via scoring is
$$\alpha \cdot \sqrt{\frac{\hat{p} \cdot (1 - \hat{p})}{n}} \cdot \sqrt{\frac{N - n}{N - 1}}.$$

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. It turns out that α is less than 1.00 for 12 of 22 poverty lines in Figure 7.

The formula relating confidence intervals with standard errors for the scorecard can be rearranged to give a formula for determining sample size before measurement. If \bar{p} is the expected poverty rate before measurement, then the formula for sample size n from a population of size N that is based on the desired confidence level that corresponds to z and the desired confidence interval $\pm c$ is

$n = N \cdot \left(\frac{z^2 \cdot \alpha^2 \cdot \bar{p} \cdot (1 - \bar{p})}{z^2 \cdot \alpha^2 \cdot \bar{p} \cdot (1 - \bar{p}) + c^2 \cdot (N - 1)} \right)$. If the population N is “large” relative to the

sample size n , then the finite-population correction factor ϕ can be taken as one (1),

and the formula becomes $n = \left(\frac{\alpha \cdot z}{c} \right)^2 \cdot \bar{p} \cdot (1 - \bar{p})$.

To illustrate how to use this, suppose the population N is 3,072,524 (the number of households in Malawi in 2010/11), suppose $c = 0.04864$, $z = 1.64$ (90-percent confidence), and the relevant poverty line is 100% of the PBM-definition national line so that the most sensible expected poverty rate \bar{p} is Malawi’s overall poverty rate for that line in 2010/11 (32.4 percent at the household level, Figure 1). The α factor is 1.03 (Figure 7). Then the sample-size formula gives

$$n = 3,072,524 \cdot \left(\frac{1.64^2 \cdot 1.03^2 \cdot 0.324 \cdot (1 - 0.324)}{1.64^2 \cdot 1.03^2 \cdot 0.324 \cdot (1 - 0.324) + 0.04864^2 \cdot (3,072,524 - 1)} \right) = 265,$$

which is not far from the sample size of 256 observed for these parameters in Figure 6 for 100% of the PBM-definition national line. Taking the finite population correction

factor ϕ as one (1) gives the same result, as $n = \left(\frac{1.03 \cdot 1.64}{0.04864} \right)^2 \cdot 0.324 \cdot (1 - 0.324) =$

265.³⁶

³⁶ Although USAID has not specified confidence levels nor intervals, IRIS Center (2007a and 2007b) says that a sample size of $n = 300$ is sufficient for USAID reporting. USAID microenterprise partners in Malawi should report using the PBM-definition \$1.25/day line. Given the α factor of 0.98 for this line in 2010/11 (Figure 7), an expected before-measurement household-level poverty rate of 51.1 percent (the all-Malawi rate in

Of course, the α factors in Figure 7 are specific to Malawi, its poverty lines, its poverty rates, and its scorecard. The derivation of the formulas for standard errors using the α factors, however, is valid for any poverty-assessment tool following the approach in this paper.

In practice after the end of fieldwork for the IHS in March 2011, a program would select a poverty line (say, 100% of the PBM-definition national line), note its participants' population size (for example, $N = 10,000$ participants), select a desired confidence level (say, 90 percent, or $z = 1.64$), select a desired confidence interval (say, ± 2.0 percentage points, or $c = \pm 0.02$), make an assumption about \bar{p} (perhaps based on a previous measurement such as the household-level poverty rate for 100% of the PBM-definition national line for Malawi of 32.4 percent in the 2010/11 IHS in Figure 1), look up α (here, 1.03 in Figure 7), assume that the scorecard will still work in the future and for sub-groups that are not nationally representative,³⁷ and then compute the required sample size. In this illustration,

$$n = 10,000 \cdot \left(\frac{1.64^2 \cdot 1.03^2 \cdot 0.324 \cdot (1 - 0.324)}{1.64^2 \cdot 1.03^2 \cdot 0.324 \cdot (1 - 0.324) + 0.02^2 \cdot (10,000 - 1)} \right) = 1,352.$$

2010/11, Figure 1), and a confidence level of 90 percent ($z = 1.64$), then $n = 300$ implies a confidence interval of $\pm 1.64 \cdot 0.98 \cdot \sqrt{\frac{0.511 \cdot (1 - 0.511)}{300}} = \pm 4.6$ percentage points.

³⁷ This paper reports accuracy for the scorecard applied to its validation samples, but it cannot test accuracy for later years or for sub-groups. Performance after March 2011 will resemble that in the 2010/11 IHS with deterioration over time to the extent that the relationships between indicators and poverty status change.

7. Estimates of changes in 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.

This section discusses non-hybrid estimates of change in which both the baseline and follow-up use the new 2010/11 scorecard with same poverty line.

Because the new 2010/11 scorecard is calibrated to government- and old-definition lines as well as PBM-definition lines, existing users of the old 2004/5 scorecard—after switching to the new 2010/11 scorecard—can still find hybrid estimates of change in poverty rates over time for supported government- or old-definition lines with a baseline from the old 2004/5 scorecard and a follow-up from the new 2010/11 scorecard. The appendix (not this section) explains the step-by-step mechanics of that calculation.

To give an idea of how accurate the new 2010/11 Malawi scorecard might be when used to measure changes in poverty rates over time from now on, this section looks at how accurate the scorecard would have been, had it been applied between:

- 2010/11 validation sample (as baseline)
- All of the 2004/5 data (as follow-up)

The tests here are stringent because:

- They compare scorecard estimates with known, true values from the IHS
- Poverty rates in Malawi from 2004/5 to 2010/11 fell a little (by the government definition) or a lot (by the PBM definition). The long time frame increases the risk of inaccuracy due to changing relationships between indicators and poverty
- The tests are *out-of-sample* in that they use only IHS data that is not also used in construction or calibration of the new 2010/11 scorecard
- The tests are *out-of-time* in that the follow-up is from a different time (2004/5) than the baseline (2010/11)

Of course, these backward-looking tests—the only ones possible for estimates of changes in poverty rates—can only give a rough idea of how accurate the scorecard might be when used from now on. After all, the factors that mattered in the past will differ in type, degree, and extent from the factors that will matter in the future. This is the unfortunate-but-inevitable nature of scorecards.

Because estimates from the scorecard are unbiased when applied to an unchanging population in which there are unchanging relationships between indicators and poverty, inaccuracies in estimates of change between the two IHS rounds must be due to:

- Sampling variation
- Inconsistent data quality
- Inaccuracy in the adjustment for changes in prices over time
- Change in the relationships between indicators and poverty over time
- Change in the composition of Malawi's population

7.1 Warning: Change is not necessarily 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: the scorecard simply estimates change, and it does not, in and of itself, indicate the reason for the change. In particular, estimating the impact of 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, the scorecard can help estimate the impact of participation only if there is some way to know—or explicit assumptions about—what would have happened in the absence of participation. And that information must come from beyond the scorecard.

7.2 Estimating changes in poverty rates over time

Consider the illustration begun in the previous section. On 1 January 2016, an organization samples three households who score 20, 30, and 40 and so have poverty likelihoods of 77.6, 55.1, and 39.6 percent (100% of the PBM-definition national line, Figure 3). Adjusting for the known bias in the validation sample of -1.0 percentage points (Figure 7), the group's baseline estimated poverty rate is the households' average poverty likelihood of $[(77.6 + 55.1 + 39.6) \div 3] - (-1.0) = 58.4$ percent.

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

- Score a new, independent sample from the same population
- Score the same sample that was scored at baseline

By way of illustration, suppose that two years later on 1 January 2018, the organization samples three additional households who are in the same population as the three original households and finds that their scores are 25, 35, and 45 (poverty likelihoods of 64.8, 47.1, and 32.5 percent, 100% of the PBM-definition national line, Figure 3). Adjusting for the known bias, the average poverty likelihood at follow-up is $[(64.8 + 47.1 + 32.5) \div 3] - (-1.0) = 49.1$ percent, an improvement of $58.4 - 49.1 = 9.3$ percentage points.³⁸ Supposing that exactly two years passed between the average baseline interview and the average follow-up interview, the estimated annual decrease in poverty is $9.3 \div 2 = 4.7$ percentage points per year. About one in 11 participants in this hypothetical example cross the poverty line in 2016/8.³⁹ Among those who start below the line, about one in six ($9.3 \div 58.4 = 15.9$ percent) on net end up above the line.⁴⁰

Alternatively, suppose that the three original households who were scored at baseline are scored again on 1 January 2018. Given scores of 25, 35, and 45, their follow-up poverty likelihoods are 64.8, 47.1, and 32.5 percent. The average across households of the difference in each given household's baseline poverty likelihood and its follow-up poverty likelihood is $[(77.6 - 64.8) + (55.1 - 47.1) + (39.6 - 32.5)] \div 3 = 9.3$

³⁸ Of course, such a huge reduction in poverty in two years is highly unlikely, but this is just an example to show how the scorecard can be used to estimate change.

³⁹ This is a net figure; some start above the line and end below it, and vice versa.

⁴⁰ The scorecard does not reveal the reasons for this change.

percentage points. Assuming in this example that there are exactly two years between each household’s interviews, the estimated annual decrease in poverty is (again) $9.3 \div 2 = 4.7$ percentage points per year.

Both approaches to estimating change through time are unbiased. In general (and unlike in the simple example here), however, they will give different estimates due to differences in the timing of interviews, in the composition of the samples, and in the nature of two samples being scored once versus one sample being scored twice.

7.3 Accuracy for estimated change in two independent samples

The accuracy of scoring’s estimates of changes in poverty rates over time is checked using IHS data from 2004/5 and 2010/11. While one cannot “drive by looking in the rear-view mirror”, historical accuracy is the best-available—but inevitably imperfect—indicator of future accuracy.

Across 20 poverty lines⁴¹ under the three definitions applied to the 2010/11 validation sample (baseline) and all of the 2004/5 data (follow-up), the average absolute error is 2.2 percentage points, while the average absolute true change is 4.1 percentage points.

For 14 of 20 lines, the true value is in the estimate’s 90-percent confidence interval (given $n = 1,024$). That is, the estimated change is not statistically different

⁴¹ For the line that marks the poorest half of people below the national line, changes are not estimated, as this line is not constant in real terms over time.

from the true change with 90-percent confidence for almost three in four lines. Of course, if all of scoring's assumptions hold, then 90 percent of estimates' 90-percent confidence intervals would contain the true value.

Scoring gets the direction of change (that is, whether poverty increased or decreased) correct for 16 of 20 lines, with the two exceptions being the PBM- and government-definition lines for the two highest lines of \$5.00 and \$8.44/day.

For 14 of 20 lines, the estimated direction matches the true direction and is “statistically significant” (zero is not in the estimate's 90-percent confidence interval). The six misses are (for the government and PBM definitions) the lowest line (food) and the two highest lines (\$5.00 and \$8.44/day). In other words, for poverty lines that are not very low nor very high, scoring's estimate of the direction of change is correct and “statistically significant”.

In sum, the scorecard usually gets the *sign* of change correct. The absolute error in the estimated *size* of change is, on average, about half of the absolute true change, and almost three-fourths of the true changes are in the 90-percent confidence interval of the estimated changes.

Are these estimates of change for Malawi “good enough”? The answer depends, of course, on the context and purpose of a given analysis task. Sometimes scoring is adequate, sometimes not. While greater accuracy is always preferred and sought, a strength of the scorecard is that its accuracy is known, allowing judgments about how much trust to put in scoring estimates to be transparent and intentional. The accuracy

of estimates of change in Malawi is among the highest of the 13 countries for which such tests have been done. Of course, accuracy might be better (or worse) from now on in Malawi.

7.4 Precision for estimates of change in two samples

Beyond errors in the size and the sign of estimated magnitudes, another formal aspect of accuracy is the standard statistical concept of *precision*. Figure 8 reports precision as 90-percent confidence intervals (given $n = 16,384$) and more generally as the α factor used in formulas for standard errors.

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

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

Here, z , c , \hat{p} and N are defined as above, n is the sample size at both baseline and follow-up,⁴² 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.

⁴² This means that—for a given level of precision—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.

Given $n = 16,384$, the 90-percent confidence intervals for estimates of change over time are ± 0.9 percentage points or less (Figure 8).

Seen another way, the average α factor across the 20 poverty lines in Figure 8 is 0.94; scoring's standard errors for estimates of change are in general about 6-percent smaller than standard errors under direct measurement.

Is this precise enough? There can be no general, once-and-for-all answer as to whether the scorecard's bias and standard errors are small enough to be useful for measuring change over time. After all, accuracy requirements vary by context and purpose. The scorecard's precision is a little better than that of direct measurement. The estimated direction of change is always correct, except for the highest and lowest poverty lines. The average absolute error is about half of the average absolute true change, and two-thirds of estimates include the true value in their 90-percent confidence interval.

Is the scorecard better than feasible alternatives for measuring change over time? This question is also difficult to answer. A central strength of scoring is that its accuracy is known, while the accuracy of most alternatives is unknown or unreported.

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

$$n = 2 \cdot N \cdot \left(\frac{z^2 \cdot \alpha^2 \cdot \tilde{p} \cdot (1 - \tilde{p})}{z^2 \cdot \alpha^2 \cdot \tilde{p} \cdot (1 - \tilde{p}) + c^2 \cdot (N - 1)} \right). \text{ If } \phi \text{ can be taken as one, then the}$$

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

To illustrate the use of this formula 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 ($\pm c = \pm 0.02$), the poverty line is 100% of the PBM-definition national line, $\alpha = 0.99$ (Figure 8), $\hat{p} = 0.324$ (the household-level poverty rate in 2010/11 for 100% of the PBM-definition national line in Figure 1), and the population N is large enough relative to the expected sample size n that the finite population correction ϕ can be taken as one. Then the baseline sample size is

$$n = 2 \cdot \left(\frac{0.99 \cdot 1.64}{0.02} \right)^2 \cdot 0.324 \cdot (1 - 0.324) \cdot 1 = 2,887, \text{ and the follow-up sample size is}$$

also 2,887.

7.5 Precision for estimated change for one sample, scored twice

Analogous to previous derivations, the general formula relating the confidence interval $\pm c$ 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:⁴³

$$\pm c = \pm z \cdot \sigma = \pm z \cdot \alpha \cdot \sqrt{\frac{\hat{p}_{12} \cdot (1 - \hat{p}_{12}) + \hat{p}_{21} \cdot (1 - \hat{p}_{21}) + 2 \cdot \hat{p}_{12} \cdot \hat{p}_{21}}{n}} \cdot \sqrt{\frac{N - n}{n - 1}},$$

where z , c , α , N , and n are defined as usual, \hat{p}_{12} is the share of all sampled households that move from below the poverty line to above it, and \hat{p}_{21} is the share of all sampled households that move from above the line to below it.

Because the IHS data for Malawi does not cover the same households in more than one round (except by pure chance, and even then, there is no way to identify such households), it is not possible to estimate values of α here.

The formula for confidence intervals 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 \tilde{p}_{12} and \tilde{p}_{21} . Before measurement, a conservative assumption is that the change in the poverty rate will be zero, which implies $\tilde{p}_{12} = \tilde{p}_{21} = \tilde{p}_*$, giving:

$$n = 2 \cdot \left(\frac{\alpha \cdot z}{c} \right)^2 \cdot \tilde{p}_* \cdot \sqrt{\frac{N - n}{n - 1}}.$$

⁴³ See McNemar (1947) and Johnson (2007). John Pezzullo helped find this formula.

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

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

Given this, a sample-size formula for a group of households to whom the new 2010/11 Malawi scorecard is applied twice (once after March 2011 and then again later) is

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

In Peru (the only source of a data-based estimate, Schreiner, 2009), 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 ($\pm c = \pm 0.02$), the poverty line is 100% of the PBM-definition national line, the sample will first be scored in 2016 and then again in 2019 ($y = 3$), and the population N is so large relative to the expected sample size n that the finite population correction ϕ can be taken as one. The pre-baseline poverty rate p_{2016} is taken as 32.4 percent (Figure 1), and α is assumed to be 1.30. Then the baseline sample size is

$$n = 2 \cdot \left(\frac{1.30 \cdot 1.64}{0.02} \right)^2 \cdot \{-0.02 + 0.016 \cdot 3 + 0.47 \cdot [0.324 \cdot (1 - 0.324)]\} \cdot 1 = 2,976. \text{ The}$$

same group of 2,976 households is scored at follow-up as well.

8. Targeting

When an organization uses the scorecard for segmenting clients for targeted services, households with scores at or below a cut-off are labeled *targeted* and treated—for program purposes—as if they are below a given poverty line. Households with scores above a cut-off are labeled *non-targeted* and treated—for program purposes—as if they are above a given poverty line.

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

Households who score at or below a given cut-off should be labeled as *targeted*,⁴⁴ not as *poor*. After all, unless all targeted households have poverty likelihoods of 100 percent, some of them are non-poor (their consumption is above a given poverty line). With the scorecard, the terms *poor* and *non-poor* have specific definitions. Using these same terms for targeting status is incorrect and misleading.

⁴⁴ A label is acceptable as long as it describes the segment and does not confuse targeting status (having a score below a program-selected cut-off) with poverty status (having consumption below an externally defined poverty line). Examples of acceptable labels include *Groups A, B, and C*; *Households scoring 29 or less, 30 to 69, or 70 or more*; and *Households who qualify for reduced fees, or do not qualify for reduced fees*.

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 9 depicts these four possible targeting outcomes. Targeting accuracy varies by the cut-off score; a higher cut-off has better inclusion (but worse leakage), while a lower cut-off has better exclusion (but worse 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 10 shows the distribution of households by targeting outcome for Malawi. For an example cut-off of 39 or less, outcomes for 100% of the PBM-definition national line in the 2010/11 validation sample are:

- Inclusion: 19.2 percent are below the line and correctly targeted
- Undercoverage: 13.1 percent are below the line and mistakenly not targeted
- Leakage: 12.8 percent are above the line and mistakenly targeted
- Exclusion: 54.8 percent are above the line and correctly not targeted

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

- Inclusion: 23.3 percent are below the line and correctly targeted
- Undercoverage: 9.1 percent are below the line and mistakenly not targeted
- Leakage: 19.1 percent are above the line and mistakenly targeted
- Exclusion: 48.5 percent are above the line and correctly not targeted

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

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

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

- Assign benefits and costs to possible outcomes, based on its values and mission
- Tally total net benefits for each cut-off using Figure 10 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 the “hit rate”, where total net benefit is the number of households correctly included or correctly excluded:

Hit rate =	1	x	Households correctly included	–
	0	x	Households mistakenly undercovered	–
	0	x	Households mistakenly leaked	+
	1	x	Households correctly excluded.	

Figure 10 shows the hit rate for all cut-offs for the new 2010/11 scorecard for Malawi. For 100% of the PBM-definition national line in the 2010/11 validation sample, total net benefit is greatest (74.1) for a cut-off of 39 or less, with about three in four households in Malawi correctly classified.

The hit rate weighs successful inclusion of households below the line the same as successful exclusion of households above the line. If a program values inclusion more (say, twice as much) than exclusion, then it can reflect this by setting the benefit for inclusion to 2 and the benefit for exclusion to 1. Then the chosen cut-off will maximize $(2 \times \text{Households correctly included}) + (1 \times \text{Households correctly excluded})$.⁴⁵

⁴⁵ Figure 10 also reports BPAC, the Balanced Poverty Accuracy Criteria adopted by USAID for certifying poverty-assessment tools. It is discussed in Section 9.

As an alternative to assigning benefits and costs to targeting outcomes and then choosing a cut-off to maximize total net benefits, a program could set a cut-off to achieve a desired poverty rate among targeted households. The third column of Figure 11 (“% targeted HHs who are poor”) shows, for Malawi’s new 2010/11 scorecard applied to the 2010/11 validation sample, the expected poverty rate among households who score at or below a given cut-off. For the example of 100% of the PBM-definition national line, targeting households in the 2010/11 validation sample who score 39 or less would target 32.0 percent of all households (second column) and would be associated with a poverty rate among those targeted of 60.1 percent (third column).

Figure 11 also reports two other measures of targeting accuracy. The first is a version of coverage (“% poor HHs who are targeted”). For the example of 100% of the PBM-definition national line with the 2010/11 validation sample and a cut-off of 39 or less, 59.4 percent of all poor households are covered.

The final targeting measure in Figure 11 is the number of successfully targeted poor households for each non-poor household mistakenly targeted (right-most column). For 100% of the PBM-definition national line with the 2010/11 validation sample and a cut-off of 39 or less, covering 1.5 poor households means leaking to 1 non-poor household.

9. Context for poverty-assessment tools in Malawi

This section discusses twelve existing poverty-assessment tools for Malawi in terms of their goals, methods, definitions of *poverty*, data, indicators, bias, precision, and cost. In general, the advantages of the scorecard are its:

- Use of data from the most recent available nationally representative consumption survey
- Reporting bias and precision for estimates of poverty rates at a point in time from out-of-sample tests, including formulas for standard errors
- Reporting bias and precision for estimates of changes in poverty rates between two points in time from out-of-sample/out-of-time tests, including formulas for standard errors
- Fewer and lower-cost indicators
- Use of a consumption-based definition of *poverty* that is widely understood and that is used by government of Malawi
- Targeting accuracy that is similar to that of alternative approaches
- Feasibility for local, pro-poor programs, due to its simplicity and transparency

9.1 Morris *et al.*

Morris *et al.* (2000) use data on 707 rural households from a special-purpose survey in central Malawi in 1998 to test an approach to poverty assessment that measures “socioeconomic position” inexpensively enough to be included in health surveys and epidemiological studies.

They report that their indicators cover 22 assets and nine types of livestock. Each indicator’s value is defined as the number of units 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. Thus, rarer items get more points. (For example, if one-third of households own a bicycle, then each bicycle a household owns gets $1 \div (1 \div 3) = 3$

points.) The index is the logarithm of the sum of each indicator multiplied by its points. They then measure accuracy as the correlation coefficient between the index and their measure of socioeconomic status (an incomplete measure of consumption).

The scorecard 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 to improve their social-performance management. In contrast, Morris *et al.* have purely methodological aims (thus, they do not report indicators or points).

Second, the new 2010/11 scorecard here is based on a nationally representative database that is newer and that samples more households.

Third, the new 2010/11 scorecard defines *socioeconomic status* as whether per-capita household consumption is below a given poverty line. This is a more-common and complete definition than that of Morris *et al.*

Fourth, the new scorecard produces poverty likelihoods that have absolute units, whereas index values 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 2010/11 scorecard is tested *out-of-sample* on data that is not used in its construction. In contrast, Morris *et al.* test their tool *in-sample* with the same data that was used for construction. In-sample tests overstate accuracy. Beyond correlation coefficients, this paper reports differences between estimates and true values as well as formulas for standard errors.

Sixth, the new 2010/11 scorecard is less costly to use than Morris *et al.* (10 indicators versus about 30). It is also simpler for non-specialists to understand, as it eschews reciprocals and logarithms.

9.2 Gwatkin *et al.*

Gwatkin *et al.* (2007) construct a poverty-assessment tool for Malawi with an approach that they use in 56 countries with Demographic and Health Surveys (Rutstein and Johnson, 2004). They use Principal Components Analysis to make an asset index from simple, low-cost indicators available for the 14,213 households in Malawi’s 2000 DHS.⁴⁶ The PCA index is like the scorecard here except that, because the DHS does not collect data on consumption, the index is based on a different conception of *poverty*, its accuracy vis-à-vis consumption-based poverty is unknown, and it can only be assumed to be a proxy for long-term wealth/economic status.⁴⁷ Well-known examples of the PCA asset-index approach include Stifel and Christiaensen (2007), Zeller *et al.* (2006), Sahn and Stifel (2003 and 2000), Henry *et al.* (2003), and Filmer and Pritchett (2001).

⁴⁶ All DHS datasets for Malawi since 1993 include each household’s asset-index value (dhsprogram.com/topics/wealth-index/Wealth-Index-Construction.cfm, retrieved 1 August 2015).

⁴⁷ Nevertheless, the indicators are similar and the “flat maximum” is important, so carefully built PCA indexes and consumption-based poverty-assessment tools may pick up the same underlying construct (perhaps “permanent income”, see Bollen, Glanville, and Stecklov, 2007), and they may rank households much the same. Comparisons of rankings by PCA indexes, directly-measured consumption, and consumption-based poverty-assessment tools include Filmer and Scott (2012), Howe *et al.*, (2009), Lindelow (2006), Sahn and Stifel (2003 and 2000), Wagstaff and Watanabe (2003), and Montgomery *et al.* (2000).

The 12 indicators in Gwatkin *et al.* are similar to those in the scorecard in terms of their simplicity, low cost, and verifiability:

- 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 a domestic worker not related with the head
- Whether any household members work their own or family's agricultural land
- Ownership of consumer durables:
 - Radios
 - Televisions
 - Bicycles
 - Motorcycles or scooters
 - Cars or trucks

Gwatkin *et al.* suggest three possible uses for their index:

- Segmenting households by the quintile of their index value to see how health varies with socio-economic status
- Monitoring (via exit surveys) how well local health-service posts reach the poor
- Measuring local coverage of health services via small-scale surveys

The first goal is akin to targeting, and the last two goals deal with performance monitoring, so the asset index would be used much like the scorecard here.

Still, the Gwatkin *et al.* index is more costly and difficult-to-use than the scorecard. The index has 12 indicators (versus 10), and while the scorecard requires adding up 10 integers (some of them usually zeroes), Gwatkin *et al.*'s index requires adding up 76 numbers, each with five decimal places and half with negative signs.

A strength of asset indexes is that, because they do not require consumption 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. In comparison, the scorecard is linked directly to a consumption-based poverty line. Thus, while both approaches can rank households, only the scorecard can estimate consumption-based poverty status.

In essence, Gwatkin *et al.*—like all asset indexes—define *poverty* in terms of the indicators and the points in the index itself. Thus, the index is not a proxy standing in for something else (such as consumption); rather, it is a direct measure of a non-consumption-based definition of *poverty*. There is nothing wrong—and a lot right—about defining *poverty* in this way, but it is not as common as a consumption-based definition. It also means that ranks from different asset indexes are not comparable, because the definition of *poverty* changes when the indicators and points in a country’s asset index change.

The asset-based approach defines people as *poor* if their assets (physical, human, financial, and social) fall below a threshold. Arguments for an asset-based view of development include Carter and Barrett (2006), Schreiner and Sherraden (2006), Sahn and Stifel (2003), and Sherraden (1991). The main advantages of the asset-based view are that:

- Asset ownership is easier to measure accurately than consumption
- Access to resources in the long term—and thus capacity to produce income and to consume—depends on the control of assets
- Assets get at capability more directly, the difference between, say, “Would income allow for adequate sanitation?” versus “Does the toilet drain to a septic tank?”

While the asset view and the income/consumption view are distinct, they are also tightly linked. After all, income and consumption are flows of resources received/consumed from the use of stocks of assets. Both views are low-dimensional simplifications—due to practical limits on definitions and measurement—of a higher-dimensional and more complete conception of the production of human well-being.

9.3 Doctor

Doctor (2004) uses Malawi's 1987 and 1998 Population and Housing Censuses with PCA to check how a "living standards index" (segmented by urban/rural) is associated with child mortality. This index is close kin to that in Gwatkin *et al.* (2007), although it is based on different data.

Derived from census data, Doctor's 13 indicators—like those here and those in Gwatkin *et al.*—are simple, quick-to-collect, and verifiable:

- Characteristics of the residence:
 - Source of drinking water
 - Type of floor
 - Type of wall
 - Type of roof
 - Type of toilet arrangement
- Ownership of consumer durables:
 - Radios
 - Bicycles
 - Motorbikes
 - Motor vehicles
- Sector of occupation of the household head
- Education of the household head
- Source of energy for:
 - Cooking
 - Lighting

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.

Doctor does not report targeting accuracy for his PCA-based index because his goal is to relate living standards with child mortality for research purposes, not to provide a way to help target services to households with high risk of child mortality.

9.4 Howe, Hargreaves, and Huttly

Howe, Hargreaves, and Huttly (“HHH”, 2008) focus on methods, asking whether PCA-based indexes (like those in Doctor 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-measurement 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. Furthermore, it does not seem to improve performance over PCA (Booyesen *et al.*, 2008).

The appeal of the indexes with dichotomized indicators is that they are simpler than the (already simple) indicators in Gwatkin *et al.* and here.

For each of the six approaches, HHH rank households by quintiles and then compare the extent of agreement with quintile ranks based on consumption. All of the approaches have about the same targeting accuracy. 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 [as a substitute for PCA-based indexes].”

The constancy of accuracy across approaches is such a common result in the predictive-modeling literature that it 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 reduction in accuracy is low.

9.5 Mukherjee and Benson

Mukherjee and Benson (2003) use the 1997/8 IHS to construct a poverty tool “to assess the likely impact on poverty of a number of poverty-reduction policy interventions” (p. 339). That is, Mukherjee and Benson seek not to estimate poverty rates nor to reach poor households with targeted programs. Instead, their goal is to measure the causal effects of indicators on poverty so as to guide policies that might address the drivers that determine poverty.⁴⁸ Thus, their 28 indicators include only things that affect current poverty but that are not affected by current poverty:

- Household demographics:
 - Age of the household head
 - Sex of the household head
 - Number of members ages 9 or younger
 - Number of members ages 10 to 17
 - Number of women ages 18 to 59
 - Number of men ages 18 to 59
 - Number of members ages 60 or older
 - Number of members of all ages, squared

⁴⁸ World Bank (2007, chapter 2 of the full report) has a similar goal and approach, but it uses the 2004/5 IHS instead of the 1997/8 IHS.

- Education:
 - Maximum education among all member ages 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:
 - District
 - Average maize yield
 - Interaction of district with average maize yield
 - Availability of agricultural inputs
 - Availability of electricity
 - Availability of a public-works program
 - Average travel time to the nearest 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 consumption to build four regional poverty-assessment 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 values of 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 person-level poverty rate (by the national line and holding all else constant) by 23.1 percent in urban areas and by 12.5 to 15.0

percent in rural areas. Removing small children would also reduce the number of poor people in a purely mechanical way.

Likewise, Mukherjee and Benson find it “very encouraging” (p. 353) that increasing by one the number of women with a 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 already knows that smaller families and greater education for girls would reduce poverty. Mukherjee and Benson’s contribution is to quantify the magnitude of the reductions. Still, poverty reduction in practice is usually constrained not by technical knowledge of what drives poverty but rather by political, financial, and organizational factors. Why would the people in Malawi’s government prioritize poverty reduction in the first place? How could they fund a family-planning campaign or administer a quantum increase in secondary education for girls? In general, how could the Malawi government design effective policies and then implement them effectively? In practice in development, the challenge is rarely knowing what to do; instead, it is in actually having the incentives and organizational/political/financial capacity to actually try do it.

Thus, the scorecard differs from Mukherjee and Benson chiefly in focus. Rather than seek to identify poverty drivers, it seeks to identify poor households, both for targeting and for monitoring. Rather than identify promising policies, it aims to help managers in local programs to implement a given pro-poor policy more 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.

9.6 Benson

Benson (2002) uses “poverty mapping” (Elbers, Lanjouw, and Lanjouw, 2003) with the 1997/8 IHS to estimate person-level poverty rates for Malawi’s 3 regions, 27 districts, 368 Traditional Authorities/urban-administrative wards, 851 local-government wards, and 9,218 Enumeration Areas. The main goal is to inform poverty policy and to give Malawi’s government a guide for directing more resources and attention more accurately to poorer areas.

For each of the 24 strata in the 1997/8 IHS (the 23 IHS strata, plus a group of rural Enumeration Areas that are really urban), Benson uses stepwise least-squares regression of the logarithm of per-capita consumption with indicators found both in the 1997/8 IHS and in the September 2008 Population and Housing Census (including Enumeration-Area census means). The 24 poverty-assessment tools are then applied to census data with the national poverty line associated with the 1997/8 IHS (MWK10.47

per person per day). This set-up gives poverty estimates for smaller areas than is possible with only the 1997/8 IHS. Finally, Benson reports estimated poverty rates and their standard errors for all regions, districts, and Traditional Authorities/urban-administrative wards.

Poverty mapping in Benson has much in common with the scorecard here in that they both:

- Build poverty-assessment tools with data that is representative of a population (a single scorecard for all of Malawi here, and 24 tools for Benson's poverty map) and then apply the tools to other data on groups that are not, in general, representative of the same populations
- Estimate poverty rates for groups
- Test accuracy empirically
- Report bias and standard errors
- Have similar accuracy
- Provide unbiased estimates when their assumptions hold
- Seek to be useful in practice and so aim to be understood by non-specialists

Strengths of poverty mapping include that it:

- Has formally established theoretical properties
- Can be applied straightforwardly to distributional measures of well-being (such as the poverty gap or the Gini coefficient) that go beyond head-count poverty rates
- Accounts for uncertainty in the estimation of tool points when estimating standard errors
- Requires data on fewer households for construction and calibration
- Includes indicators at the level of the Enumeration Area, decreasing bias and increasing precision
- Uses only indicators that are in a census

Strengths of the scorecard include that it:

- Uses simple, verifiable indicators that are quick and inexpensive to collect
- Is simpler in terms of both construction and application
- Associates poverty likelihoods with scores non-parametrically
- Surfaces estimates of poverty likelihoods for individual households
- Reduces overfitting by selecting indicators with statistical and non-statistical criteria and by having only a single, all-Malawi scorecard⁴⁹
- Tests accuracy *out-of-sample* (that is, with data not used in scorecard construction)
- Reports confidence intervals and simple formulas for standard errors
- Aims to be transparent to non-specialists

The basic difference between the two approaches is that poverty mapping seeks to help governments to target pro-poor policies, while the scorecard seeks to help local, pro-poor organizations to manage their social performance.⁵⁰ On a technical level,

⁴⁹ According to Mahadevan, Yoshida, and Praslova (2013, pp. 6–7) “the latest recommendation from poverty-map experts in the World Bank Research Department is not to use multiple [poverty-assessment tools] to predict household consumption” because they can be “problematic since the number of observations for each area becomes small and, as a result, the regression coefficients become less stable.” To reduce this overfitting, Haslett (2012) likewise recommends that poverty maps be based on a single, all-country tool. Benson (2002) is prone to overfitting, with 24 tools based on data from fewer than 7,000 households. Indeed, Benson worries a lot about the accuracy of estimates in Enumeration Areas with less than 500 households, but 22 of the 24 poverty-map tools used to produce those estimates are themselves based on data on less than 500 households. The danger is that while the 24 tools may fit the 1997/8 IHS data well, their accuracy may degrade a lot when applied outside the IHS (for example, with census data). Unfortunately, there is no way to check the extent of degradation; the census does not collect consumption data (and if it did, then the poverty map’s tools would be superfluous anyway).

⁵⁰ Another apparent difference is that the developers of poverty mapping (Elbers, Lanjouw, and Lanjouw, 2003; Demombynes *et al.*, 2004) say that poverty mapping is too inaccurate to be used for targeting at the household level. In contrast, Schreiner (2008b) supports household-level targeting as a legitimate, potentially useful application of the scorecard. In Elbers *et al.* (2007), the developers of poverty mapping seem to take a step back from their previous position.

Benson estimates consumption directly, whereas the scorecard estimates poverty likelihoods.

Benson's 24 tools use an average of about 14 indicators from among the following 49:

- Demographics:
 - Age of the household head (and its square)
 - Whether the head is female
 - Marital status of the head
 - Number of household members
 - Number of members who are not part of the nuclear family
 - Number of male members ages 5 or younger
 - Number of male members ages 6 to 14
 - Number of male members ages 15 to 29
 - Number of male members ages 30 to 49
 - Number of male members ages 50 or older
 - Number of female members (any age)
 - Number of female members ages 5 or younger
 - Number of female members ages 6 to 14
 - Number of female members ages 15 to 29
 - Number of female members ages 30 to 49
 - Number of female members ages 50 or older
 - Number of fertile female members (ages 15 to 45)
 - Average years between births for female members who have given birth
 - Total children born to fertile female members
 - Mortality rate among children born to fertile female members
- Education:
 - Years of schooling completed by the head
 - Number of members who have finished primary school
 - Highest class finished by a male member
 - Highest class finished by a female member
 - Number of primary-school-aged children who are in primary school

- Employment:
 - Whether the head is an employee
 - Whether the head is an employer
 - Whether the household has a family business
 - Number of household members working in a secondary industry
 - Number of household members working in a tertiary industry
 - Number of household members working in services
 - Number of household members working in an “other” occupation
- Quality of the residence:
 - Tenancy status
 - Type of cooking fuel
 - Source of drinking water
- Whether the household owns any bicycles
- Enumeration-Area census means:
 - Population density
 - Education:
 - Maximum education level among household members
 - Net enrollment rate
 - Distance (straight-line, in kilometers) to:
 - Boma
 - Nearest health facility
 - Nearest market center
 - Nearest primary or secondary road
 - Nearest urban center
 - Quality of residence:
 - Whether the residence is rented
 - Whether the residence has an improved toilet
 - Whether the residence is made of permanent materials
 - Rooms per person
 - Agriculture:
 - Average 20-year maize yield
 - Difference in 1997/8 maize yield from long-term average

Benson's poverty map is not meant for use by local, pro-poor organizations. For example, there are 24 tools, complicating administration if an organization works in more than a single Traditional Authority/urban administrative ward. Furthermore, an organization's back-office would have to match up a household and its Enumeration Area with its average census values.

Benson reports bias and standard errors for estimated poverty rates at the levels of regions, districts, and Traditional Authorities/urban administration wards. For all of Malawi, in-sample bias for Benson's 24-tool poverty map is -1.0 percentage points, versus (out-of-sample) +0.1 percentage points with the old 2004/5 scorecard in Schreiner (2011). Other accuracy comparisons are not possible—even though the old 2004/5 scorecard uses the same data as Benson—because Benson's three regions do not match up with the four in Schreiner (2011).

9.7 Tatem *et al.*

Tatem *et al.* (2014) make a poverty map⁵¹ of Malawi with consumption data from the 2010/11 IHS with \$1.25 and \$2.00/day 2005 PPP poverty lines.⁵² Unlike Benson, Tatem *et al.* do not construct many stratum-level poverty-assessment tools by linking consumption by IHS households with indicators in the IHS that are also in a census. Instead, they make a single all-Malawi tool with a Bayesian regression of the logarithm of average per-capita aggregate consumption for geo-coded IHS households against satellite-sensed environmental indicators. The environmental indicators are known for small areas of Malawi, so they can be fed into the resulting all-Malawi tool to estimate average consumption—and poverty rates—in all areas not covered by the IHS.

⁵¹ Called a *high-resolution gridded poverty surface* with an estimate for every km².

⁵² Tatem *et al.* do not report poverty lines in MWK, nor do they report all-Malawi person-level poverty rates. Still, their lines probably do not match PovcalNet nor the government-definition 2005 PPP lines here. Regional price deflation is already built into the measure of consumption—in prices in February/March 2010—supplied by the NSO. But Tatem *et al.* report deriving their \$1.25/day line by multiplying the 2005 PPP factor for Malawi (56.922) by \$1.25 and by the ratio of the average CPI for 2011 and 2005 reported by PovcalNet (1.676), which matches the ratio of the average 2011 CPI to the average 2005 CPI of $331.304 \div 197.204 = 1.680$ (rbm.mw/inflation_rates_detailed.aspx, retrieved 30 July 2015.) This suggests that Tatem *et al.*'s lines are $1.25 \times 56.922 \times 1.676 = \text{MWK}119.25$ per person per day in average prices during calendar-year 2011 (when consumption is in prices as of February/March 2010). PovcalNet's line in MWK is unreported, and the government-definition \$1.25/day line here (in prices as of February/March 2010) is MWK133.90 (Figure 1). Furthermore, as of 1 August 2015, the average 2005 and 2011 CPIs on PovcalNet are not actually the average annual CPIs, but rather some figures (246.35 and 114.8) that give the ratio of 2.146, which is almost the 2.148 used here, by the NSO, and by PBM to take prices from February/March 2004 to February/March 2010.

Tatem *et al.* use 12 indicators, most of them non-intuitive to non-specialists (or at least presented that way):

- Accessibility to cities with more than 50,000 people via all transport methods
- Accessibility to three sizes of settlements via major or minor roads
- The Afripop measure of urban population
- The GRUMP measure of population density
- The GRUMP population count
- Average annual aridity
- Average annual potential envirotranspiration
- Night-time lights
- Elevation
- “evi”
- “lst”
- “midir”

Like Benson, Tatem *et al.* intend for their map to help governments to target resources more finely.

The poverty map’s accuracy is tested out-of-sample via ten-fold cross-validation. Tatem *et al.* state that the estimates “for Malawi are essentially unbiased (mean square error = -0.0213 [\$2.00/day line] and -0.0316 [\$1.25/day]), indicating no overall tendency to over- or under-predict the poverty-headcount ratio” (p. 34). They also say (p. 35) that “The correlation between predicted and actual values was around 0.73, indicating a good degree of linear association.” Of course, the goodness of the association depends on the purpose and context in which the estimates are used.

9.8 Mathiassen

The approach in Mathiassen (2006, with the 2004/5 IHS) is similar to that here (with the 2010/11 IHS) and in Schreiner, 2011 (with the 2004/5 IHS).⁵³ Both build simple, inexpensive poverty-assessment tools with the explicit goal of measuring poverty rates at a point in time.⁵⁴ Both estimate poverty rates as the average of individual households' poverty likelihoods.⁵⁵ And both divide the data into two sub-samples, one for construction and one for testing accuracy.⁵⁶

Before Malawi's 2005 Welfare Measurement Survey (WMS) was designed, Mathiassen used data from the first six months of fieldwork for the 2004/5 IHS to build four tools. The 28 indicators that appeared in one or more of the four tools were then put in the WMS with the express purpose of using the tool to estimate Malawi's poverty rate in 2005—and in later years—without the WMS' having to incur the cost of measuring consumption. Thus, Mathiassen is a rare example of a poverty-assessment tool actually being used for a commonly proposed purpose: to update poverty estimates between consumption surveys using “light” surveys (for example, Fofack, 2000).

⁵³ The approach is also discussed in Mathiassen (2009).

⁵⁴ This paper also seeks to measure changes in poverty rates and to help with targeting. Mathiassen could also be used for these purposes, although they are not discussed.

⁵⁵ Mathiassen 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 logarithm of per-capita consumption. Mathiassen's approach is a correct, valid, and sometimes preferred alternative to Logit (Coudouel, Hentschel, and Wodon, 2002).

⁵⁶ Both papers report bootstrapped average differences between estimated and true poverty rates, and both report standard errors or (equivalently) confidence intervals.

Mathiassen differs from the scorecard mainly in the number of indicators (28 across the four tools versus 10 here), in the verifiability of indicators, and in the number of tools (one for each of the four poverty-line regions versus one for all of Malawi). Fourteen of Mathiassen's 28 indicators deal with consumption or past spending and thus are not verifiable. Two others are ratios and thus would be difficult to calculate on paper in the field:

- Demographics:
 - Number of household members
 - Number of members younger than 15
 - Dependency ratio (number of members younger than 15 or older than 60 divided by the total number of members)
 - Age of head
- Highest educational qualification for a household member
- Characteristics of the residence:
 - Type of floor
 - Type of roof
 - Number of members per room
- Ownership of consumer durables:
 - Whether the head sleeps under sheets
 - Number of changes of clothes for the head
 - Number of radios
 - Beds
 - Irons
 - Refrigerators
 - Mobile telephones
- Whether the household used/consumed an item in the 30 days (transport) or past 7 days (food items):
 - Transport
 - Eggs
 - Meat
 - Rice
 - Bread
 - Fresh milk
 - Cooking oil
 - Sugar
 - Toothpaste

- Expenses in MWK in the past seven days:
 - Cooking oil
 - Sugar
- Whether the household purchased clothing items in the past three months:
 - Men's clothing
 - Shoes

To measure bias, Mathiassen divides the 2004/5 IHS data in half, re-constructs the four models, and applies them to the other half of the data. Bias in Malawi's four poverty-line regions is +0.9 percentage points (Urban), +1.4 (North Rural), +1.1 (Centre Rural), and +0.7 (South Rural). With the same 2004/5 data and the same testing approach, bias for the single all-Malawi scorecard in Schreiner (2011) is zero, +3.1, +7.7, and -6.2 percentage points. Together, Mathiassen's four tools have much less bias.

When applied to yearly WMS data from 2005 to 2009, Mathiassen's tools estimate a large decrease in the person-level poverty rate for the national line: from 50 percent in 2005, to 40 percent in 2007, and to 29 percent in 2009 (NSO, 2010b, p. 85). PBM note that this is consistent with the possibility that the government definition of *poverty* understates improvements in living standards between the 2004/5 and 2010/11 IHS.

In sum, the approach here and in Mathiassen are similar in construction and testing. The scorecard here, however, uses fewer indicators and only verifiable indicators, so it is less costly to use and more difficult to game. Mathiassen has less bias, and there is a single all-Malawi scorecard, versus four tools for Mathiassen.

9.9 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 poverty-assessment tool to measure change over time, thus matching all three uses here. 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 2010/11 IHS)
- Segment tools by urban/rural (versus no segmenting)
- Use 17 indicators for rural and 7 for urban (versus 10 for all-Malawi)
- Estimate per-capita consumption directly via a stepwise least-squares regression (versus poverty likelihoods from a Logit regression based on statistical accuracy as well as non-statistical criteria)
- Report accuracy for person-level estimates (versus household-level)

As noted earlier, the scorecard is not segmented because tests elsewhere find that segmentation does not improve targeting accuracy much. Segmentation may improve (or worsen) the accuracy of estimated poverty rates.

This paper estimates poverty likelihoods, rather than consumption, because the poverty-likelihood approach transparently quantifies the error inherent in estimates for individual households. Furthermore, the poverty-likelihood approach here produces unbiased estimates of poverty rates. In contrast, poverty rates derived from assigning

poverty likelihoods of either zero or 100 percent based on an estimate of consumption and a cut-off (such as a poverty line) are biased.⁵⁷ Benson *et al.* note that converting their estimates of consumption to poverty likelihoods—as recommended by Mathiassen (2009) and by Coudouel, Hentschel, and Wodon (2002)—would eliminate bias.

Benson *et al.* use indicators that resemble those here in terms of simplicity and verifiability, save for two:

- Household size (and its square)
- Education of the head
- Number of salaried household members
- Type of fuel for:
 - Cooking
 - 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
- Having purchased sugar in the past two to four weeks

It is difficult to verify the recent purchase of sugar (and the relevant time period is not well-defined), and households sometimes cannot easily report acres cultivated.

⁵⁷ There is bias because the function converting the estimate of consumption to a poverty likelihood is not continuous. For a given poverty line, there is a cut-off that gives unbiased estimates, but it differs from the poverty line. For example, the cut-off for the USAID Poverty Assessment Tool (see below) is set to make estimates unbiased.

Are Benson *et al.*'s urban and rural poverty-assessment tools more accurate than the single scorecard for all of Malawi? An apples-to-apples comparison is not possible, as the two use data from different IHS rounds. Also, Benson *et al.* build and test with person weights, versus household weights. Furthermore, the poverty rate must be held constant, as does the share of people targeted. These differences put the scorecard at a disadvantage, as it must adjust in the test to match Benson *et al.* More important, however, is that the 1997/8 IHS has worse data quality than later rounds, and it collected fewer good poverty indicators.⁵⁸ This puts Benson *et al.* at a net disadvantage.

The table on the next page compares Benson *et al.* (with 1997/8 IHS data) with Schreiner (2011, with 2004/5 IHS data) in terms of the bias of estimated poverty rates as well as in terms of targeting inclusion, exclusion, and the hit rate. The comparison uses person-level weights, adjusts the scorecard's poverty line and targeting cut-off to match Benson *et al.*'s poverty rate and to target the same share of people as in Benson *et al.* The test covers urban and rural areas for two poverty lines (food and national).

As the "flat maximum" would predict, targeting accuracy is about the same. The scorecard has better inclusion in three of four cases, better exclusion in three of four cases, and a better hit rate in all cases. But differences are small, especially considering that Benson *et al.* start at a net disadvantage. These results are consistent with

⁵⁸ Benson *et al.* identify poverty indicators suggested by Malawi households and link them with community wealth ranks. On Benson *et al.*'s advice, the NSO added many of these to later IHS rounds (and some appear now in scorecards).

Schreiner (2014), which compares accuracy in a large number of countries for the scorecard versus an approach similar to Benson *et al.*

For estimated poverty rates, bias is similar in two of four cases, and the scorecard has less bias in the other two (zero versus +5.1 percentage points in urban areas with the national line, and -1.0 percentage points versus -11.6 in rural areas with the food line). This is remarkable, as Benson *et al.* use urban and rural tools rather than a single all-Malawi scorecard.

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

**Comparison of accuracy, Benson *et al.* (1997/8 IHS data) versus Schreiner
(2011, 2004/5 IHS data)**

Area	Line	Poverty	Share	Bias		Inclusion		Exclusion		Hit rate	
		rate	targeted	Benson	Card	Benson	Card	Benson	Card	Benson	Card
Urban	Food	21.9	20.9	-0.7	+0.3	13.0	12.9	70.5	71.0	83.5	83.9
Urban	National	46.5	51.6	+5.1	0.0	37.9	38.6	39.8	44.4	77.7	83.0
Rural	Food	23.5	12.3	-11.6	-1.0	7.1	7.7	71.4	71.6	78.5	79.3
Rural	National	53.4	54.8	+1.4	1.8	39.1	46.7	30.9	24.5	70.0	71.2

All figures are percentages of people.

Benson *et al.* is constructed and tested (out-of-sample) with data from the 1997/8 IHS.

Schreiner (2011) is constructed and tested (out-of-sample) with data from the 2004/5 IHS.

The poverty line in the validation sample for Schreiner (2011) is set so that the poverty rate matches Benson *et al.*

The targeting cut-off for Schreiner (2011) is set so that the share of people targeted matches Benson *et al.*

9.10 IRIS Center

USAID commissioned IRIS Center (2012) to build a “Poverty Assessment Tool” (PAT) using data from the 2004/5 IHS so that USAID’s microenterprise partners in Malawi could report the share of their participants who are “very poor”. In general, the PAT for Malawi is like the scorecard, except that the PAT:

- Estimates consumption directly (rather than poverty likelihoods) and then converts estimated consumption into a poverty likelihood of either 0 or 100 percent (rather than between 0 and 100)
- Has more indicators (19 rather than 10)

The PAT supports two 2005 PPP poverty lines:

- \$1.25/day
- \$2.50/day

IRIS tests four regression-based approaches in both one-stage and two-stage versions (IRIS, 2005), settling on a one-step quantile regression that estimates the 57th percentile of the logarithm of per-capita household consumption. It uses 19 indicators (IRIS, 2012):

- Demographics:
 - Number of household members (and its square)
 - Age of the head (and its square)
 - Marital status of the head
- Education:
 - Share of household members (excluding head) who have never attended school or who have no education
 - Share of adult members who can read

- Characteristics of the residence:
 - Presence of electricity
 - Number of rooms
 - Type of floor
 - Type of toilet arrangement
- Ownership of consumer durables:
 - Coffee tables
 - Beds
 - Tape players, CD players, or hi-fis
 - Irons
 - Bicycles
- Location:
 - Poverty-line region
 - Urban/rural
- Agriculture:
 - Cultivates a *dimba* garden
 - Owns goats
- Household purchased powered laundry detergent in the past month

Except for past purchases of laundry detergent, these indicators are simple, inexpensive, and verifiable.

Schreiner (2014) reports an apples-to-apples comparison of accuracy for IRIS (2012) versus the old 2004/5 scorecard from Schreiner (2011).⁵⁹ In out-of-sample tests, the PAT and the scorecard have about the same absolute bias (0.3 versus 0.0 percentage points).⁶⁰ The PAT is less precise (α of 0.94 versus 0.85). For targeting, the PAT correctly classifies 1.3 more people per 100 than does the scorecard. Thus, in terms of accuracy in Malawi, the PAT and the old 2004/5 scorecard are about tied.

⁵⁹ Schreiner (2014) corrects the mistakes in the scorecard’s 2004/5 old-definition \$1.25/day 2005 PPP poverty line, and replicates an error in this poverty line by the PAT (it fails to adjust for regional-price differences). This ensures that the household-level poverty rates in the test are the same for both the scorecard and the PAT.

⁶⁰ When bias is known, it can be removed, so both the PAT and scorecard are unbiased.

IRIS also reports accuracy in terms of the Balanced Poverty Accuracy Criterion. IRIS Center (2005) introduces BPAC, and USAID adopted it as its criterion for approving poverty-assessment tools for use by its microenterprise partners. BPAC considers accuracy in terms of targeting inclusion and in terms of the absolute difference between undercoverage and leakage (which, under the PAT’s approach, is equal to the absolute value of the bias of the estimated poverty rate). The formula is:

$$\text{BPAC} = 100 \cdot \left(\frac{\text{Inclusion} - |\text{Undercoverage} - \text{Leakage}|}{\text{Inclusion} + \text{Undercoverage}} \right).$$

Because bias (in the PAT approach) is the difference between undercoverage and leakage, and because the normalization term $\frac{100}{\text{Inclusion} + \text{Undercoverage}}$ is possibly relevant only when comparing poverty-assessment tools across populations with different poverty rates (but irrelevant when selecting among alternative tools for a given country in a given year for a given poverty line), the simpler formula $\text{BPAC} = \text{Inclusion} - |\text{Bias}|$ ranks poverty-measurement tools the same as the more complex formula.

Expressing BPAC as $\text{Inclusion} - |\text{Bias}|$ helps to show why BPAC is not useful for comparing the PAT with the scorecard (Schreiner, 2014). Given the assumptions discussed earlier,⁶¹ the scorecard produces unbiased estimates of poverty rates, regardless of whether undercoverage differs from leakage. While BPAC can be used to compare alternative poverty-assessment tools that use the PAT’s consumption-

⁶¹ The unbiasedness of the PAT also requires these assumptions.

estimation approach, it does not make sense to apply BPAC to the scorecard's likelihood-estimation approach. This is because—unlike the PAT—the scorecard does not use a single cut-off to classify households as either 100-percent poor or 0-percent poor. Instead, households have an estimated poverty likelihood somewhere between 0 to 100 percent. If a scorecard user sets a targeting cut-off, then that cut-off matters only for targeting, without affecting the estimation of poverty rates at all.

Although IRIS reports the PAT's targeting accuracy and although the BPAC formula considers targeting accuracy in terms of inclusion, IRIS says that the PAT should not be used for targeting.⁶²

IRIS also doubts that the PAT can be useful for measuring change over time, 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 unless the tools are exceptionally accurate, then the changes identified are likely to be contained within the margin of error.”⁶³

That is, IRIS asserts⁶⁴ that the confidence interval for estimates of change—for some unstated confidence level, some unstated sample size, and some unstated true change—will usually include zero. In Malawi for the new 2010/11 scorecard applied out-of-sample to the 2010/11 validation sample (baseline) and out-of-sample/out-of-time to the entire 2004/5 IHS (follow-up), estimates of change have the same sign as the true

⁶² povertytools.org/faq/faq.html#11, retrieved 19 February 2009.

⁶³ povertytools.org/faq/faq2.html, retrieved 7 December 2012.

⁶⁴ IRIS has never reported the PAT's accuracy for estimates of change over time.

change and are statistically different from zero with $n = 1,024$ and 90-percent confidence for 14 of 20 poverty lines. Likewise, the 90-percent confidence interval ($n = 1,024$) of the estimated change includes the true change for 14 of 20 lines.

In the same way, targeting and estimating changes over time are possible uses that are supported for the scorecard, despite IRIS' doubts. In particular, this paper reports targeting accuracy so users can decide for themselves whether scoring targets adequately for their purposes.

9.11 Cnobloch and Subbarao

Cnobloch and Subbarao (2015) follow IRIS, Benson *et al.*, and Mathiassen in that they make a poverty-assessment tool with stepwise least-squares regression on the logarithm of per-capita consumption, selecting indicators based on R^2 and on the p values of their coefficients.⁶⁵ The tool—derived from data from the 2004/5 IHS—then assigns households 0- or 100-percent poverty likelihoods based on whether estimated consumption is below a poverty line. Cnobloch and Subbarao aim to provide “clear criteria with which to identify beneficiaries consistently across [Malawi]” to fulfill “a need to reach out to the poor with safety-net programs” and to help “ensure that program coverage is adequate, program implementation is efficient, and targeting is effective” (p. 129).

To accomplish this targeting purpose, Cnobloch and Subbarao “develop an objective measure for identifying potential beneficiaries for safety-net programs . . . using a proxy-means test formula [a poverty-assessment tool] for targeting the poorest (bottom 10 percent [of people]) and the extreme poor (bottom 25 percent) of the consumption distribution” (p. 131). They test the following models:

- All-Malawi with a single cut-off
- All-Malawi with separate urban/rural cut-offs
- Segmented urban/rural models with separate cut-offs

⁶⁵ Cnobloch and Subbarao know of the old 2004/5 scorecard, noting that “the approaches are similar” and echoing Schreiner (2011) in parts of their text.

Cnobloch and Subbarao focus on the all-Malawi model with a single cut-off, noting that the segmented models perform only “marginally better” [and sometimes worse]⁶⁶ and saying that the number of observations in the 2004/5 IHS is inadequate to support two models.

Cnobloch and Subbarao’s all-Malawi tool uses 18 indicators:

- Demographics:
 - Number of household members
 - Number of members ages 4 or younger
 - Number of members ages 5 to 12
 - Number of members ages 13 to 18
 - Number of members ages 65 or older
 - Structure of household headship
 - Marital status of the head
- Characteristics of the residence:
 - Presence of electricity
 - Type of floor
 - Type of outer wall
 - Type of roof
 - Source of drinking water
 - Type of toilet arrangement
 - Number of rooms per member
- Agriculture:
 - Whether the household cultivated tobacco in the last cropping season
 - Acres of rain-fed land cultivated in the last cropping season
 - Acres of *dimba* land cultivated in the last dry season
- District of residence

These “do not include hard-to-verify indicators (such as self-employment in agriculture or in the non-agriculture sector) or easy-to-conceal indicators (such as ownership of a radio or mobile phone)” (p. 133). The included indicators are all simple

⁶⁶ This is consistent with the “flat maximum” as well as with targeting experience with poverty-assessment tools segmented by urban/rural as cited earlier in this paper.

and verifiable, except the two on acreage cultivated; in practice, some households find these difficult to answer because they do not know the land area or because they know it only in non-acre units. Also, the inclusion of the location (district) of residence may spark political objections.

How does the targeting accuracy of Cnobloch and Subbarao's tool compare with that of the old 2004/5 scorecard in Schreiner (2011)? The test is unusually clean, as both use the same data, although it is still not perfect. For the poverty line that identifies the poorest 10 percent of people, Cnobloch and Subbarao's tool targets 16.5 percent of all people, with inclusion of 5.2 percent, exclusion of 78.7 percent, and a hit rate of 83.9 percent. Using the same poverty line and targeting the same share of people, the old 2004/5 scorecard has inclusion of 5.1 percent, exclusion of 79.0 percent, and a hit rate of 84.1 percent. As usual, the tools are tied in terms of targeting accuracy.

This tie—while not unexpected—is still remarkable, both because Cnobloch and Subbarao are at a disadvantage (because they exclude consumer durables while the scorecard includes them) and because Cnobloch and Subbarao are at an advantage because they use:

- 18 indicators (versus 10)
- Location of residence (which the scorecard excludes)
- In-sample tests (versus out-of-sample)
- Person-level weights in both construction and testing (the scorecard uses household-level weights in construction and then person-level weights for this test)

9.12 McBride and Nichols

McBride and Nichols (MN, 2015) use data from Malawi's 2004/5 IHS to test the targeting accuracy of random quantile-regression forests. They select indicators from among those in the IRIS PAT (2012), and they compare targeting accuracy with the PAT. Given that the PAT and the scorecard are similar in both approach and accuracy (Schreiner, 2014), it makes sense to compare/contrast MN with Malawi's old 2004/5 scorecard in Schreiner (2011) as well.

MN use a random quantile-regression forest to estimate a quantile of the distribution of a household's consumption, given the values of that household's indicators. The household is then assigned an estimated poverty likelihood of either 0 percent or 100 percent, according to whether the forest's estimate of the 57th quantile⁶⁷ of its conditional distribution of consumption is below a given poverty line (in MN, the \$2.50/day 2005 PPP line that has a household-level poverty rate of 65.0 percent).⁶⁸

The random quantile-regression forest is a collection of random regression trees. MN start the construction of the forest by following the PAT and the scorecard by dividing data from the 2004/5 IHS into a construction and validation sample.⁶⁹ They then draw 500 bootstrap samples from the construction sample, dividing each bootstrap sample into a sub-construction sample (two-thirds of the data) and a sub-validation

⁶⁷ MN choose the 57th quantile because it maximizes BPAC.

⁶⁸ Schreiner (2011) reports a 66.1 percent household-level poverty rate for this line. Schreiner (2014) documents two errors in the \$1.25/day line in Schreiner (2011).

⁶⁹ MN note that with random quantile-regression forests, this division is not needed to estimate bias and precision, leading to an understatement of their approach's accuracy.

sample (one-third of the data). They then derive 500 random regression trees. A regression tree splits the sub-construction data recursively into two segments, choosing an indicator and a split between the ordered values of the chosen indicator so as to maximize a measure of the difference between the two segments in terms of their observed consumption (Breiman *et al.*, 1984). MN’s regression trees are “random” because only one-third of the indicators used by the PAT—selected at random—are available for any given split.

The random quantile-regression forest combines data for all households in all 500 random regression trees to estimate the quantile of the distribution of a given household’s consumption.⁷⁰

Random quantile-regression forests are not unbiased, but—as IRIS does with the PAT—the quantile of estimation can be chosen to make estimates of poverty rates unbiased. Also, it is known that combining estimates—as forests do as a sort of poverty-assessment tool of poverty-assessment tools—often produces more accurate estimates than does a single tool (Palm and Zellner, 1992). Furthermore, the out-of-sample accuracy of forests is straightforwardly found via cross-validation, and the approach obviates the need for a single initial division of the data, a division that increases bias and reduces precision for the PAT and the scorecard.

Does MN’s random quantile-regression forest target better than scorecard? It might or might not. In particular, MN handicap themselves by limiting indicators only

⁷⁰ See MN and Meinshausen (2006) for details.

to those in the PAT. The comparison here also uses MN's more conservative estimates. Furthermore, even though a random quantile-regression forest is more accurate than a single random regression tree, it might not do better than a single poverty-assessment tool derived from another approach.

Holding constant the household-level poverty rate (65.04 percent) and the share of people targeted (67.3 percent) in MN, the random quantile-regression forest with Malawi's 2004/5 IHS (MN, Table A2, p. 24) has inclusion of 56.3 percent (versus 55.9 for the scorecard) and exclusion of 24.0 (versus 22.7), giving a hit rate of 80.3 (versus 78.6). Thus, MN's forest correctly classifies about two more people per 100 than does the old 2004/5 scorecard. Most of the edge comes from reduced leakage.⁷¹

Could the current Logit-and-look-up-table approach of the scorecard be replaced with random quantile-regression forests? It is not out of the question, as long as the number of indicators in the underlying trees is limited (to keep interviewing costs low). Targeting would likely be sharper, and the poverty-likelihood approach (rather than the 0- or 100-percent likelihood approach) probably could be maintained. On the other hand, random quantile-regression forests are more difficult to explain to non-specialists. With hundreds of trees in a forest, the scorecard could no longer be presented on a single page, its points would no longer be transparent, and pro-poor programs would need canned software to compute estimates. While it is feasible to grow forests and

⁷¹ Bias for the estimated person-level poverty rate is +2.3 percentage points for MN versus +0.2 for the scorecard.

build software, it is not clear to what extent users would embrace a less-simple, less-transparent tool, nor whether greater accuracy would compensate for possibly lower user acceptance.

10. Conclusion

Pro-poor programs in Malawi can use the scorecard to segment clients for targeted services as well as to estimate:

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

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

The scorecard is constructed with half of the data from Malawi's 2010/11 IHS. Its scores are then calibrated with that same data to poverty likelihoods for 10 PBM-definition poverty lines, 10 government-definition lines, and two old-definition lines. The support for government- and old-definition lines allows existing users of Malawi's old 2004/5 scorecard (Schreiner, 2011) to switch to the new 2010/11 scorecard here and to find hybrid estimates of changes in poverty rates over time with a baseline with the old 2004/5 scorecard and a follow-up with the new 2010/11 scorecard. In general, the new 2010/11 scorecard is more accurate and more relevant, so it—with government-definition and PBM-definition poverty lines (the two definitions of *poverty* most likely to be supported in the next IHS)—should be used from now on.

The accuracy of the new 2010/11 scorecard is tested on data from the 2004/5 and 2010/11 IHS that is not used in construction or calibration. Bias and precision are reported for estimates of households' poverty likelihoods, populations' poverty rates at a

point in time, and changes in populations' poverty rates over time. Of course, the scorecard's estimates of change are not necessarily the same as estimates of program impact. Targeting accuracy is also reported.

On average when the scorecard is applied to the 22 poverty lines under the three definitions of *poverty* with the 2010/11 validation sample, the maximum absolute bias for estimates versus true poverty rates for groups of households at a point in time is 1.4 percentage points. The average absolute bias is about 0.6 percentage points. Unbiased estimates may be had by subtracting the known bias for a given poverty line from the original estimates.

For $n = 16,384$ and 90-percent confidence, the precision of point-in-time estimates of poverty rates is ± 0.6 percentage points or better. With $n = 1,024$, the 90-percent confidence interval is ± 2.6 percentage points or better.

This paper also tests the accuracy of scorecard estimates of changes in poverty rates over time, using data from both the 2004/5 and 2010/11 IHS. The estimated direction of change is correct except for the highest poverty lines. For the 20 poverty lines for which estimating a change makes sense, the average absolute error is about 2.2 percentage points, which is about half of the average absolute true change of 4.1 percentage points. For 14 of 20 lines and given $n = 1,024$, the true change is within the 90-percent confidence interval of the estimated change. Furthermore, 14 of 20 estimates of the direction of change are both correct and "statistically significant" in that zero is outside of the estimate's 90-percent confidence interval.

The biggest errors are for the food line, and they are consistent with the possibility that income for the poorest in Malawi increased between 2004/5 and 2010/11 but that the increase was used to improve residences and to acquire other assets rather than to increase current consumption.

If an organization wants to use the scorecard for segmenting clients for targeted services, then the results here provide useful information for selecting a cut-off that fits its values and mission.

Although the statistical technique is innovative, and although technical accuracy is important, the design of the scorecard focuses on transparency and ease-of-use. After all, accuracy is irrelevant if an organization's managers feel so daunted by a scorecard's complexity or its cost that they do not even try to use it.

For this reason, the scorecard uses ten indicators that are straightforward, low-cost, and verifiable. Points are all zeros or positive integers, and scores range from 0 (most likely below a poverty line) to 100 (least likely below a poverty line). Scores are converted to poverty likelihoods via simple look-up tables, and targeting cut-offs are likewise straightforward to apply. The design attempts to facilitate voluntary adoption by helping managers to understand and trust scoring and by allowing non-specialists to add up scores quickly in the field.

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

References

- Adams, Niall M.; and David J. Hand. (2000) “Improving the Practice of Classifier Performance Assessment”, *Neural Computation*, Vol. 12, pp. 305–311.
- Arndt, Channing; and Kenneth R. Simler. (2010) “Estimating Utility-Consistent Poverty Lines, with Applications to Egypt and Mozambique”, *Economic Development and Cultural Change*, Vol. 58, pp. 449–474.
- 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) *Malawi: Atlas of Social Statistics*, ebrary.ifpri.org/cdm/ref/collection/p15738coll12/id/83961, retrieved 1 August 2015.
- ; Payongayong, Ellen; Ahmed, Akhter U.; and Manohar Sharma. (2006) “Simple Household Poverty Assessment Models for Malawi”.
- 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.
- Booyesen, Frikkie; van der Berg, Servaas; Burger, Ronelle; von Maltitz, Michael; and Gideon du Rand. (2008) “Using an Asset Index to Assess Trends in Poverty in Seven Sub-Saharan African Countries”, *World Development*, Vol. 36, No. 6, pp. 1113–1130.
- Breiman, Leo; Friedman, Jerome H.; Olshen, Richard A.; and Charles J. Stone. (1984) *Classification and Regression Trees*.
- Caire, Dean. (2004) “Building Credit Scorecards for Small-Business Lending in Developing Markets”, microfinance.com/English/Papers/Scoring_SMEs_Hybrid.pdf, retrieved 4 August 2015.
- ; and Mark Schreiner. (2012) “Cross-Tab Weighting for Credit Scorecards in Developing Markets”, business-school.ed.ac.uk/crc/conferences/conference-archive?a=46055, retrieved 4 August 2015.

- Camacho, Adriana; and Emily Conover. (2011) “Manipulation of Social-Program Eligibility”, *American Economic Journal: Economic Policy*, Vol. 3, No. 2, pp. 41–65.
- Carter, Michael R.; and Christopher B. Barrett. (2006) “The Economics of Poverty Traps and Persistent Poverty: An Asset-Based Approach”, *Journal of Development Studies*, Vol. 42, No. 2, pp. 178–199.
- 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 *PRSP Sourcebook*.
- Cnobloch, Rodica; and Kalanidhi Subbarao. (2015) “Targeting Methods to Identify the Poorest in Malawi”, pp. 129–157 in Carlo del Ninno and Bradford Mills (eds), *Safety Nets in Africa: Effective Mechanisms to Reach the Poor and Most Vulnerable*.
- 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 Özler. (2004) “Producing an Improved Geographic Profile of Poverty: Methodology and Evidence from Three Developing Countries”, pp. 154–176 in Anthony Shorrocks and Rolph van der Hoeven (eds.) *Growth, Inequality, and Poverty*.
- Diamond, Alexis; Gill, Michael; Rebolledo Dellepiane, Miguel Angel; Skoufias, Emmanuel; Vinha, Katja; and Yiqing Xu. (2016) “Estimating Poverty Rates in Target Populations: An Assessment of the Simple Poverty Scorecard and Alternative Approaches”, World Bank Policy Research Working Paper No. 7793, hdl.handle.net/10986/25038, retrieved 11 January 2017.
- Doctor, Henry V. (2004) “The Effect of Living Standards on Childhood Mortality in Malawi”, *African Population Studies*, Vol. 19, Supplement A, pp. 249–263.
- Elbers, Chris; Fujii, Tomoki; Lanjouw, Peter; Özler, Berk; and Wesley Yin. (2007) “Poverty Alleviation through Geographic Targeting: How Much Does Disaggregation Help?”, *Journal of Development Economics*, Vol. 83, pp. 198–213.

- ; Lanjouw, Jean O.; and Peter Lanjouw. (2003) “Micro-Level Estimation of Poverty and Inequality”, *Econometrica*, Vol. 71, No. 1, pp. 355–364.
- 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.
- ; and Kinnon Scott. (2012) “Assessing Asset Indices”, *Demography*, Vol. 49, pp. 359–392.
- Fofack, Hippolyte. (2000) “Combining Light Monitoring Surveys with Integrated Surveys to Improve Targeting for Poverty Reduction: The Case of Ghana”, *World Bank Economic Review*, Vol. 14, No. 1, pp. 195–219.
- 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 4 August 2015.
- Goodman, Leo A.; and Kruskal, William H. (1979) *Measures of Association for Cross Classification*.
- Grosh, Margaret; and Judy L. Baker. (1995) “Proxy-Means Tests for Targeting Social Programs: Simulations and Speculation”, World Bank LSMS Working Paper No. 118, go.worldbank.org/W90WN57PD0, retrieved 13 May 2016.
- Gwatkin, Davidson R.; Rutstein, Shea; Johnson, Kiersten; Suliman, Eldaw; Wagstaff, Adam; and Agbessi Amouzou. (2007) “Socio-Economic Differences in Health, Nutrition, and Population: Malawi”, go.worldbank.org/T6LCN5A340, retrieved 1 August 2015.
- Hammond, Allen L.; Kramer, William J.; Katz, Robert S.; Tran, Julia T.; and Courtland Walker. (2007) *The Next 4 Billion: Market Size and Business Strategy at the Base of the Pyramid*, wri.org/publication/next-4-billion, retrieved 4 August 2015.
- Hand, David J. (2006) “Classifier Technology and the Illusion of Progress”, *Statistical Science*, Vol. 22, No. 1, pp. 1–15.

- Haslett, Stephen. (2012) “Practical Guidelines for the Design and Analysis of Sample Surveys for Small-Area Estimation”, *Journal of the Indian Society of Agricultural Statistics*, Vol. 66, No. 1, pp. 203–212.
- Henry, Carla; Sharma, Manohar; Lapenu, Cecile; and Manfred Zeller. (2003) “Microfinance Poverty Assessment Tool”, CGAP Technical Tool No. 5, cgap.org/publications/microfinance-poverty-assessment-tool, retrieved 5 August 2015.
- Hoadley, Bruce; and Robert M. Oliver. (1998) “Business Measures of Scorecard Benefit”, *IMA Journal of Mathematics Applied in Business and Industry*, Vol. 9, pp. 55–64.
- Howe, Laura D.; Hargreaves, James R.; Gabrysch, Sabine; and Sharon R.A. Huttly. (2009) “Is the Wealth Index a Proxy for Consumption Expenditure? A Systematic Review”, *Journal of Epidemiology and Community Health*, Vol. 63, pp. 871–880.
- IRIS Center. (2012) “Poverty-Assessment Tool Accuracy Submission: USAID/IRIS Tool for Malawi”, 24 May, povertytools.org/countries/Malawi/USAID_PAT_Malawi.pdf, retrieved 4 August 2015.
- (2007a) “Manual for the Implementation of USAID Poverty Assessment Tools”, povertytools.org/training_documents/Manuals/USAID_PAT_Manual_Eng.pdf, retrieved 4 August 2015.
- (2007b) “Introduction to Sampling for the Implementation of PATs”, povertytools.org/training_documents/Sampling/Introduction_Sampling.pdf, retrieved 4 August 2015.
- (2005) “Notes on Assessment and Improvement of Tool Accuracy”, povertytools.org/other_documents/AssessingImproving_Accuracy.pdf, retrieved 4 August 2015.
- Johnson, Glenn. (2007) “Lesson 3: Two-Way Tables—Dependent Samples”, onlinecourses.science.psu.edu/stat504/node/96, retrieved 4 August 2015.
- Kolesar, Peter; and Janet L. Showers. (1985) “A Robust Credit-Screening Model Using Categorical Data”, *Management Science*, Vol. 31, No. 2, pp. 124–133.

- Lindelow, Magnus. (2006) “Sometimes More Equal Than Others: How Health Inequalities Depend on the Choice of Welfare Indicator”, *Health Economics*, Vol. 15, pp. 263–279.
- Lovie, Alexander D.; and Patricia Lovie. (1986) “The Flat-Maximum Effect and Linear Scoring Models for Prediction”, *Journal of Forecasting*, Vol. 5, pp. 159–168.
- Mahadevan, Meera; Yoshida, Nobou; and Larisa Praslova. (2013) “Poverty Mapping in the Kyrgyz Republic: Methodology and Key Findings”, World Bank Report No. 76690, documents.worldbank.org/curated/en/2013/04/17584758/kyrgyz-republic-poverty-mapping-methodology-key-findings, retrieved 4 August 2015.
- Martinelli, César; and Susan W. Parker. (2007) “Deception and Misreporting in a Social Program”, *Journal of the European Economic Association*, Vol. 4, No. 6, pp. 886–908.
- 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*, Milongwe, nsomalawi.mw/publications/welfare-monitoring-surveys-wms/welfare-monitoring-surveys-wms-2009.html, retrieved 4 August 2015.
- (2009) “A Model-Based Approach for Predicting Annual Poverty Rates without Expenditure Data”, *Journal of Economic Inequality*, June 2009, Vol. 7, No. 2, pp. 117–135.
- 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, mfc.org.pl/sites/mfc.org.pl/files/spotlight4.PDF, retrieved 4 August 2015.
- McBride, Linden; and Austin Nichols. (2015) “Improved Poverty Targeting through Machine Learning: An Application to the USAID Poverty Assessment Tools”, econthatmatters.com/wp-content/uploads/2015/01/improvedtargeting_21jan2015.pdf, retrieved 4 August 2015.
- McNemar, Quinn. (1947) “Note on the Sampling Error of the Difference between Correlated Proportions or Percentages”, *Psychometrika*, Vol. 17, pp. 153–157.

- Meinshausen, Nicoli. (2006) “Quantile Regression Forests”, *Journal of Machine Learning Research*, Vol. 7, pp. 983–999.
- 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 4 August 2015.
- National Statistical Office. (2012) “Household Socio-Economic Characteristics Report: Integrated Household Survey 2010–2011”, go.worldbank.org/NOXNI9YDS0, retrieved 4 August 2015.
- (2010a) “Enumerator Manual for the Household Questionnaire”, go.worldbank.org/6A7GUDQ1Q0, retrieved 4 August 2015.
- (2010b) “Report on the 2009 Welfare Monitoring Survey”, nsomalawi.mw/publications/welfare-monitoring-surveys-wms/welfare-monitoring-surveys-wms-2009.html, retrieved 4 August 2015.
- Onwujekwe, Obinna; Hanson, Kara; and Julia Fox-Rushby. (2006) “Some Indicators of Socio-Economic Status May Not Be Reliable and Use of Indexes with These Data Could Worsen Equity”, *Health Economics*, Vol. 15, pp. 639–644.
- Palm, Franz C.; and Arnold Zellner. (1992) “To Combine or not to Combine? Issues of Combining Forecasts”, *Journal of Forecasting*, Vol. 11, No. 8, pp. 687–701.

- Pauw, Karl; Beck, Ulrik; and Richard Mussa. (forthcoming) “Did Rapid Smallholder-Led Agricultural Growth Fail to Reduce Rural Poverty? Making Sense of Malawi’s Poverty Puzzle”, in Channing Arndt, Andrew McKay, and Finn Tarp (eds.) *A Study of Growth and Poverty in Sub-Saharan Africa*.
- Pradhan, Menno; Suryahadi, Asep; Sumarto, Sudarno; and Lant Pritchett. (2001) “Eating Like which ‘Joneses’? An Iterative Solution to the Choice of a Poverty Line ‘Reference Group’”, *Review of Income and Wealth*, Series 47, No. 4, pp. 473–487.
- Ravallion, Martin. (1998) “Poverty Lines in Theory and Practice”, World Bank LSMS Working Paper No. 133, go.worldbank.org/8P3IBJPQS1, retrieved 4 August 2015.
- Rutstein, Shea Oscar; and Kiersten Johnson. (2004) “The DHS Wealth Index”, DHS Comparative Reports No. 6, measuredhs.com/pubs/pdf/CR6/CR6.pdf, retrieved 4 August 2015.
- 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 4 August 2015.
- Schreiner, Mark. (2015) “The Process of Poverty-Scoring Analysis”, microfinance.com/English/Papers/Scoring_Poverty_Outline_Analysis.pdf, retrieved 4 August 2015.
- (2014) “How Do the Simple Poverty Scorecard and the PAT Differ?”, microfinance.com/English/Papers/Scorecard_versus_PAT.pdf, retrieved 4 August 2014.
- (2013) “Simple Poverty Scorecard Poverty-Assessment Tool: Bangladesh”, SimplePovertyScorecard.com/BGD_2005_ENG.pdf, retrieved 4 August 2015.

- (2012a) “An Expert-Based Poverty Scorecard for Rural China”, microfinance.com/English/Papers/Scoring_Poverty_China_EN.pdf, retrieved 4 August 2015.
- (2012b) “Simple Poverty Scorecard Poverty-Assessment Tool: Colombia”, SimplePovertyScorecard.com/COL_2009_ENG.pdf, retrieved 4 August 2015.
- (2011) “Simple Poverty Scorecard Poverty-Assessment Tool: Malawi”, SimplePovertyScorecard.com/MWI_2004_ENG.pdf, retrieved 4 August 2015.
- (2009) “Simple Poverty Scorecard Poverty-Assessment Tool: Peru”, SimplePovertyScorecard.com/PER_2007_ENG.pdf, retrieved 4 August 2015.
- (2008a) “Simple Poverty Scorecard Poverty-Assessment Tool: Peru”, SimplePovertyScorecard.com/PER_2003_ENG.pdf, retrieved 4 August 2015.
- (2008b) “Simple Poverty Scorecard Poverty-Assessment Tool: Ecuador”, SimplePovertyScorecard.com/ECU_2005_ENG.pdf, retrieved 4 August 2015.
- (2006) “Is One Simple Poverty Scorecard Poverty-Assessment Tool Enough for India?”, microfinance.com/English/Papers/Scoring_Poverty_India_Segments.pdf, retrieved 4 August 2015.
- (2005a) “La Herramienta del Índice de Calificación de la Pobreza™: México”, SimplePovertyScorecard.com/MEX_2002_SPA.pdf, retrieved 4 August 2015.
- (2005b) “IRIS Questions on the Simple Poverty Scorecard Poverty-Assessment Tool”, microfinance.com/English/Papers/Scoring_Poverty_Response_to_IRIS.pdf, retrieved 4 August 2015.
- (2002) *Scoring: The Next Breakthrough in Microfinance?* CGAP Occasional Paper No. 7, microfinance.com/English/Papers/Scoring_Breakthrough_CGAP.pdf, retrieved 4 August 2015.
- ; Matul, Michal; Pawlak, Ewa; and Sean Kline. (2014) “Poverty Scoring: Lessons from a Microlender in Bosnia-Herzegovina”, *Poverty and Public Policy*, Vol. 6, No. 4, pp. 407–428.
- ; and Michael Sherraden. (2006) *Can the Poor Save? Saving and Asset Accumulation in Individual Development Accounts*.

- Sharif, Iffath Anwar. (2009) "Building a Targeting System for Bangladesh Based on Proxy-Means Testing", World Bank Social Protection Discussion Paper No. 0914, [siteresources.worldbank.org/SOCIALPROTECTION/ Resources/SP-Discussion-papers/Safety-Nets-DP/0914.pdf](http://siteresources.worldbank.org/SOCIALPROTECTION/Resources/SP-Discussion-papers/Safety-Nets-DP/0914.pdf), retrieved 4 August 2015.
- Sherraden, Michael. (1991) *Assets and the Poor: A New American Welfare Policy*.
- Sillers, Don. (2006) "National and International Poverty Lines: An Overview", pdf.usaid.gov/pdf_docs/Pnadh069.pdf, retrieved 4 August 2015.
- 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. (2009) "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.
- Tatem, Andrew; Gething, Peter; Pezzulo, Carla; Weiss, Dan; and Samir Bhatt. (2014) "Development of High-Resolution Gridded Poverty Surfaces".
- Toohig, Jeff. (2008) "PPI Pilot Training Guide", microfinancegateway.org/sites/default/files/mfg-en-paper-progress-out-of-poverty-index-ppi-pilot-training-mar-2008.pdf, retrieved 4 August 2015.
- USAID. (2014) *Microenterprise Results Reporting: Annual Report to Congress, Fiscal Year 2013*, eads.usaid.gov/mrr/publications/ar2013.pdf, retrieved 4 August 2015.
- United States Congress. (2004) "Microenterprise Results and Accountability Act of 2004 (HR 3818 RDS)", November 20, smith4nj.com/laws/108-484.pdf, retrieved 4 August 2015.
- Wagstaff, Adam; and Naoko Watanabe. (2003) "What Difference Does the Choice of SES Make in Health-Inequality Measurement?", *Health Economics*, Vol. 12, No. 10, pp. 885–890.

- Wainer, Howard. (1976) “Estimating Coefficients in Linear Models: It Don’t Make No Nevermind”, *Psychological Bulletin*, Vol. 83, pp. 223–227.
- World Bank. (2012) *Targeting Poor and Vulnerable Households in Indonesia*, documents.worldbank.org/curated/en/2012/01/15879773/targeting-poor-vulnerable-households-indonesia, retrieved 4 August 2015.
- (2008) “International Comparison Project: Tables of Results”, siteresources.worldbank.org/ICPINT/Resources/icp-final-tables.pdf, retrieved 4 August 2015.
- (2007) *Malawi Poverty and Vulnerability Assessment Investing in Our Future, Volume I: Synthesis of Main Findings and Recommendations, Volume II: Detailed Background Papers, and Appendices*, Report No. 36546-MW, documents.worldbank.org/curated/en/docsearch/report/36546, retrieved 4 August 2015.
- (2005) “Note on Construction of Expenditure Aggregate and Poverty Lines for IHS2”, siteresources.worldbank.org/INTLSMS/Resources/3358986-1181743055198/3877319-1181928149600/ihs2agg.pdf, retrieved 4 August 2015.
- Zeller, Manfred. (2004) “Review of Poverty Assessment Tools”, pdf.usaid.gov/pdf_docs/PNADH120.pdf, retrieved 4 August 2015.
- ; 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.

Calculating Hybrid and Spliced Estimates of Change in Poverty Rates over Time

This appendix is a step-by-step process with which legacy users of the old 2004/5 scorecard in Malawi can calculate hybrid and spliced estimates of changes in poverty rates through time. The process makes use of past applications of the old 2004/5 scorecard by existing users, and it also allows all users from now on to make estimates of change based on current and future applications of the new 2010/11 scorecard.

In general, the process involves applying a scorecard at three points in time:

- *Past*: Only old 2004/5 scorecard, only with government- or old-definition poverty lines
- *Now*: Only new 2010/11 scorecard, with PBM-definition lines and with government-definition lines (and perhaps with old-definition) lines)
- *Future*: Only new 2010/11 scorecard, only with PBM- and government-definition lines

The steps are:

Past:

1. Select a government- or old-definition poverty line from among those supported both in this paper for the new 2010/11 scorecard and in Schreiner (2011) for the old 2004/5 scorecard. For the government definition, these are the food or 100% of national poverty lines; for the old definition, these are the \$1.25 or \$2.50/day 2005 PPP lines.⁷²

⁷² In Schreiner (2011), these lines are not labeled as “government-definition” or “old-definition”; they are just called “food”, “100% of national”, “\$1.25/day” and “\$2.50/day”. The poverty lines in Schreiner (2011) are not supported here and thus cannot be used in hybrid measures of change over time that use a baseline from the old 2004/5 scorecard and follow-up from the new 2010/11 scorecard.

2. Estimate a baseline poverty rate for the chosen government- or old-definition line:
 - a. Retrieve (from a paper file, spreadsheet, or database) the poverty likelihoods for the chosen government- or old-definition line for each household in the representative sample of a given population to whom the old 2004/5 scorecard has already been applied in the past. This likelihood comes from the look-up table for the chosen government- or old-definition line in Schreiner (2011), not from the look-up tables in this paper
 - b. Average the households' poverty likelihoods to estimate their baseline poverty rate for the chosen government- or old-definition line, subtracting off known bias for the chosen government- or old-definition line from Figure 8 in Schreiner (2011, p. 89).

Present:

3. Estimate a follow-up poverty rate for the chosen government- or old-definition line:
 - a. Apply the new 2010/11 scorecard to a representative sample of the same population to which the old 2004/5 scorecard was originally applied in (2a)⁷³
 - b. Add up the score for each household from the new 2010/11 scorecard
 - c. Convert each household's score to a poverty likelihood using the look-up tables for the chosen government- or old-definition line in Figure 3 this paper (not the look-up tables in Schreiner, 2011). In this paper, the government- and old-definition lines are explicitly labeled as "government-definition" or "old-definition"
 - d. Average the households' poverty likelihoods to estimate their follow-up poverty rate for the chosen government- or old-definition line, subtracting off known bias as found in Figure 7 on pp. 175–179 of this paper.

4. Find hybrid estimates of change for the chosen government- or old-definition line:
 - a. The estimated hybrid change is the estimated follow-up poverty rate (3d) minus the estimated baseline poverty rate (2b). If estimated poverty decreased (got better) through time, then the result will be a negative number. If estimated poverty increased (got worse) through time, then the result will be a positive number.
 - b. The estimated hybrid change relative to the share of participants who were under the chosen government- or old-definition line at baseline is the estimated hybrid change (4a) divided by the estimated baseline poverty rate (2b)
 - c. The estimated net number of participants who crossed from below the chosen government- or old-definition poverty line to above it since baseline is the negative of the change (4a) expressed as a proportion,⁷⁴ multiplied by the number of participants in the population at baseline

⁷³ What matters is that the sample be representative of the same population as that to which the old 2004/5 scorecard was originally applied in (2a). In particular, the new 2010/11 scorecard does not have to be applied to the same households as the old 2004/5 scorecard (although it could be).

⁷⁴ For example, 0.123 is the proportion that is equivalent to 12.3 percentage points.

To be ready to estimate on-going changes in poverty rates over time using government- and PBM-definition poverty lines, all users (legacy and new) from now on should:

5. Select a government-definition poverty line and a PBM-definition poverty line from among those supported in this paper. For the government definition, these are food; 100%, 150%, or 200% of the national line; or \$1.25, \$2.00, \$2.50, \$5.00, or \$8.44/day 2005 PPP).⁷⁵ For the PBM definition, these are food; 100%, 150%, or 200% of the national line; or \$1.25, \$2.00, \$2.50, \$5.00, or \$8.44/day 2005 PPP). Users should choose at least one government-definition line and at least one PBM-definition line because it is not known which one—if any—will be supported in the next IHS. The two lines need not be the same. For example, a user could choose the government-definition \$1.25/day line along with 100% of the PBM-definition national line.

⁷⁵ The line marking the poorest half of people below 100% of the national line is omitted because its real value changes with time. Thus, it is not meaningful when estimating changes in poverty.

6. Estimate a baseline poverty rate for both the chosen government-definition line and for the chosen PBM-definition line:
 - a. In addition to the sample of households to which the new 2010/11 scorecard was applied in (3a), apply the new 2010/11 scorecard to samples of households that are representative of any additional populations of interest
 - b. Add up (or retrieve from 3b) the score for each household to which the new 2010/11 scorecard has been applied in (3a)
 - c. Convert each household's score to a poverty likelihood using the look-up tables for the chosen government-definition line and for the chosen PBM-definition line. These score-to-poverty-likelihood look-up tables are found in this paper (not in Schreiner, 2011, as none of the look-up tables there pertain to government-definition lines nor PBM-definition lines calibrated to the new 2010/11 scorecard)
 - d. For the sample of households to which the new 2010/11 scorecard was applied in 3a (and separately for any samples of households that are representative of any additional populations of interest in 6a), average the households' poverty likelihoods (separately for the likelihoods pertaining to the chosen government-definition line and those pertaining to the chosen PBM-definition line) to estimate their baseline poverty rate for the chosen lines, subtracting off known bias as found in this paper's Figure 7, pp. 175–179

From this point on, all estimates of change are based solely on the government-definition and PBM-definition lines for the new 2010/11 scorecard in this paper.

Future:

7. Select the government-definition and PBM-definition poverty lines for which a baseline poverty rate has been estimated in 6d

8. Estimate a follow-up poverty rate for the two chosen government-definition and PBM-definition lines:
 - a. Apply the new 2010/11 scorecard to a representative sample of the same population to which the new 2010/11 scorecard was originally applied (3a, as well as any additional populations represented in 6a)

 - b. Add up the score for each household to which the new 2010/11 scorecard has just been applied (8a)

 - c. Convert each household's score to a government-definition poverty likelihood and to a PBM-definition poverty likelihood using the look-up tables for the chosen government-definition and PBM-definition lines in this paper's Figure 3 (not the look-up tables in Schreiner, 2011, none of which pertain to government-definition or PBM-definition lines calibrated to the new 2010/11 scorecard)

 - d. For the sample(s) representing a given population (8a), average the households' poverty likelihoods (separately for the chosen government-definition and PBM-definition lines) to get an estimate of their follow-up poverty rate for the chosen government-definition and PBM-definition lines, subtracting off known bias as found in this paper's Figure 7, pp. 175–179

9. Find the (non-hybrid) estimates of change for the chosen government-definition and PBM-definition lines:
 - a. The estimated change is the estimated follow-up poverty rate (8d) minus the estimated baseline poverty rate (6d), separately for both the chosen government-definition and PBM-definition lines. If estimated poverty decreased (got better) through time, then the result will be a negative number. If estimated poverty increased (got worse) through time, then the result will be a positive number.
 - b. The estimated change relative to the share of participants who were under the chosen government-definition and PBM-definition lines at baseline is (separately for each chosen line) the change (9a) divided by the estimated baseline poverty rate (6d)
 - c. The estimated net number of participants who crossed from below the chosen government-definition and PBM-definition lines to above them since baseline is (separately for each chosen line) the negative of the estimated change (9a) expressed as a proportion, multiplied by the number of participants at baseline

10. Assuming that the “parallel lines” assumption holds,⁷⁶ find the “grand” estimates of change that splice together hybrid and non-hybrid estimates:
- a. The “grand” spliced estimate of change is the hybrid estimate of change (4a) for a chosen government- or old-definition line plus the non-hybrid estimate of change for a chosen government- or PBM-definition line (9a)
 - b. The “grand” spliced estimate of change relative to the share of participants who were below the chosen government- or old-definition line in the past baseline is the “grand” estimate of change (10a) divided by the share of participants who were below the chosen government- or old-definition line in the past baseline (2b). (There is no “grand” spliced estimate of relative change for a chosen PBM-definition line because there is no estimate of the poverty rate by the chosen PBM-definition line in the past baseline)
 - c. The “grand” spliced estimate of the net number of participants who crossed from below the chosen government- or old-definition line to above it (or from below the chosen PBM-definition line to above it) since the past baseline is the negative of the “grand” estimate of change 10a expressed as a proportion, multiplied by the number of participants in the past baseline (2b)

⁷⁶ As discussed in the text, there is no evidence that the “parallel lines” assumption holds for Malawi between 2004/5 and 2010/11 between the PBM definition and the government definition, nor between the PBM definition and the old definition. Of course, the “parallel lines” assumption can be expected to hold for the government and old definitions.

The following hypothetical example illustrates the steps with specific numbers:

Past:

1. *Select a government- or old-definition poverty line from among those supported both in this paper for the new 2010/11 scorecard and in Schreiner (2011) for the old 2004/5 scorecard:*

Select 100% of the government-definition national line.

2. *Estimate a baseline poverty rate for the chosen government- or old-definition line:*
 - a. *Retrieve (from a paper file, spreadsheet, or database) the scores and the poverty likelihoods for the chosen government- or old-definition line for each household in the representative sample of a given population to whom the old 2004/5 scorecard has already been applied in the past. This likelihood comes from the look-up table for the given government- or old-definition line in Schreiner (2011), not the look-up tables in this paper*

In this hypothetical example, the scores and likelihoods for the three⁷⁷ households in the sample are:

Score	Poverty likelihood (100% of the government-definition national line)
15	88.9
20	82.5
25	70.0

The poverty likelihoods for 100% of the government-definition national line for the old 2004/5 scorecard in Schreiner (2011) come from Figure 4, p. 85 of Schreiner (2011).⁷⁸

⁷⁷ Of course, three households is an unrealistically small sample, but it is used in this hypothetical illustration to keep the arithmetic manageable.

⁷⁸ This is “Figure 4 (National line): Estimated poverty likelihoods associated with scores”, microfinance.com/English/Papers/Scoring_Poverty_Malawi_2004_EN.pdf, retrieved 5 August 2015.

- b. *Average the households' poverty likelihoods to get an estimate of their baseline poverty rate for the chosen government or old-definition line, subtracting off known bias.*

$$[(88.9 + 82.5 + 70.0) \div 3] - (+0.1) = 80.4 \text{ percent.}$$

The known bias of +0.1 percentage points for 100% of the government-definition national line comes from Figure 8, p. 89 of Schreiner (2011)

Present:

3. *Estimate a follow-up poverty rate for a chosen government- or old-definition line:*

a. *Apply the new 2010/11 scorecard to a representative sample of the same population to which the old 2004/5 scorecard was originally applied in (2a)*

Draw a new sample of three households.

b. *Add up the score for each household from the new 2010/11 scorecard*

In this hypothetical example, the scores are 21, 26, and 31.

c. *Convert each household's score to a poverty likelihood using the look-up tables for the chosen government- or old-definition line from Figure 3 in this paper (not the look-up tables in Schreiner, 2011)*

Look up poverty likelihoods for 100% of the government-definition national line on p. 249 in this paper.

Score	Poverty likelihood (100% of the government-definition national line)
21	87.4
26	80.0
31	72.0

d. *Average the households' poverty likelihoods to get an estimate of their follow-up poverty rate for the chosen government- or old-definition line, subtracting off known bias*

$$[(87.4 + 80.0 + 72.0) \div 3] - (+1.2) = 78.6 \text{ percent.}$$

Bias for 100% of the government-definition national line with the new 2010/11 scorecard is +1.2 percentage points (Figure 7 on pp. 175–179 in this paper).

4. Find hybrid estimates of change for the chosen government- or old-definition line:

- a. The estimated change is the estimated follow-up poverty rate (3d) minus the estimated baseline poverty rate (2b). If estimated poverty decreased (got better) through time, then the result will be a negative number. If estimated poverty increased (got worse), then the result is a positive number

$$78.6 \text{ percent} - 80.4 \text{ percent} = -1.8 \text{ percentage points.}$$

- b. The estimated change relative to the share of participants who were under the chosen government- or old-definition line at baseline is the estimated change (4a) divided by the estimated baseline poverty rate (2b)

$$-1.8 \text{ percentage points} \div 80.4 \text{ percentage points} = -2.2 \text{ percent.}$$

- c. The estimated net number of participants who crossed from below the chosen government- or old-definition poverty line to above it since baseline is the negative of the change (4a) expressed as a proportion, multiplied by the number of participants at baseline

Assuming for the sake of this hypothetical illustration that there were 10,000 participants in the baseline population, $-(-0.018) \times 10,000$ participants = 180 participants.

To be ready to estimate on-going changes in poverty rates over time using government- and PBM-definition lines from now on, all users (legacy and new) should:

- 5. Select a government-definition poverty line and a PBM-definition poverty line from among those supported in this paper*

Select 100% of the PBM-definition national line and 100% of the government-definition line. (From here on, this illustration will not show the work for a government-definition line.)

6. *Estimate a baseline poverty rate for both the chosen government-definition line and for the chosen PBM-definition line:*

a. *In addition to the sample of households that are representative of the same population as that to which the new 2010/11 scorecard was applied in (3a), apply the new 2010/11 scorecard to samples of households that are representative of any additional populations of interest*

In this example, no samples are drawn from additional populations. Thus the three households in (3a) are the only three households here.

b. *Add up (or retrieve from 3b) the score for each household to which the new 2010/11 scorecard has been applied*

The scores for the three households in 3b are 21, 26, and 31.

c. *Convert each household's score to a poverty likelihood using the look-up tables for the chosen government-definition line and for the chosen PBM-definition line in this paper (not the look-up tables in Schreiner, 2011, as none of the look-up tables there pertain to government-definition lines or PBM-definition lines calibrated to the new 2010/11 scorecard)*

Look up the poverty likelihoods for 100% of the PBM-definition national line in Figure 3 on p. 171 in this paper.

Score	Poverty likelihood (100% of PBM-definition national line)
21	77.6
26	64.8
31	55.1

d. *For the sample of households to which the new 2010/11 scorecard was applied in 3a (and separately for any samples of households that are representative of any additional populations of interest in 6a), average the households' poverty likelihoods (separately for the likelihoods pertaining to the chosen government-definition line and those pertaining to the chosen PBM-definition line) to estimate their baseline poverty rate for the chosen lines, subtracting off known bias as found in this paper's Figure 7*

$$[(77.6 + 64.8 + 55.1) \div 3] - (-1.0) = 66.8 \text{ percent.}$$

The known bias for 100% of the PBM-definition national line is -1.0 percentage points (Figure 7 on pp. 175–179).

Future:

From this point on, all estimates of change are based solely on the government-definition and PBM-definition lines for the new 2010/11 scorecard in this paper:

7. *Select a government-definition and PBM-definition poverty line for which a baseline poverty rate has been estimated in 6d*

For compatibility with the above, select 100% of the PBM-definition national line and 100% of the government-definition national line. The example here will illustrate the work only with 100% of the PBM-definition national line.

8. *Estimate a follow-up poverty rate for the two chosen government-definition and PBM-definition lines:*

- a. *Apply the new 2010/11 scorecard to a representative sample of the same population to which the new 2010/11 scorecard was originally applied (3a, as well as any additional populations represented in 6a)*

Draw a new sample of three households from the same population as (3a). In this illustration, no additional samples are drawn.

- b. *Add up the score for each household to which the new 2010/11 scorecard has just been applied (8a)*

In this hypothetical example, the scores are 22, 27, and 37.

- c. *Convert each household's score to a government-definition poverty likelihood and a PBM-definition poverty likelihood using the look-up tables for the chosen government-definition and PBM-definition lines in this paper (not the look-up tables in Schreiner, 2011, none of which pertain to government-definition or PBM-definition lines calibrated to the new 2010/11 scorecard)*

Look up poverty likelihoods for 100% of the PBM-definition national line in Figure 3 on p. 171 in this paper.

Score	Poverty likelihood (100% of PBM-definition national line)
22	77.6
27	64.8
37	47.1

- d. *For the sample representing a given population (8a), average the households' poverty likelihoods (separately for the chosen government-definition and the PBM-definition lines) to get an estimate of the follow-up poverty rate for the chosen government-definition and PBM-definition lines, subtracting off known bias*

$$[(77.6 + 64.8 + 47.1) \div 3] - (-1.0) = 64.2 \text{ percent.}$$

The known bias for 100% of the PBM-definition national line is -1.0 percentage points (Figure 7 on pp. 175–179).

9. Find non-hybrid estimates of change for the chosen government- or PBM-definition line:

a. The estimated change is the estimated follow-up poverty rate (8d) minus the estimated baseline poverty rate (6d), separately for both the chosen government-definition and PBM-definition lines. If estimated poverty decreased (got better) through time, then the result will be a negative number. If estimated poverty increased (got worse), then the result will be a positive number

$$64.2 \text{ percent} - 66.8 \text{ percent} = -2.6 \text{ percentage points.}$$

b. The estimated change relative to the share of participants who were under the chosen government-definition or PBM-definition lines at baseline is (separately for each chosen line) the estimated change (9a) divided by the estimated baseline poverty rate (6d)

$$-2.6 \text{ percentage points} \div 66.8 \text{ percentage points} = -3.9 \text{ percent.}$$

c. The estimated net number of participants who crossed from below the chosen government-definition or PBM-definition line to above it since baseline (separately for each chosen line) is the negative of the change (9a) expressed as a proportion, multiplied by the number of participants at baseline

Assuming for the sake of this hypothetical illustration that there were 10,000 participants in the baseline population, $-(-0.039) \times 10,000$ participants = 390 participants.

10. Assuming that the “parallel lines” assumption holds, find the “grand” spliced estimates of change that combine the hybrid and non-hybrid estimates:

- a. The “grand” spliced estimate of change is the hybrid estimate of change for a chosen government- or old-definition line (4a) plus the non-hybrid estimate of change for the chosen government-definition or PBM-definition line (9a)

$$-1.8 \text{ percentage points} + (-2.6 \text{ percentage points}) = -4.4 \text{ percentage points.}$$

- b. The “grand” spliced estimate of change relative to the share of participants who were below the chosen government- or old-definition line in the past baseline is the “grand” estimate of change 10a divided by the share of participants who were below the chosen government or old-definition line in the past baseline (2b). (There is no “grand” spliced estimate of relative change for a chosen PBM-definition line because there is no estimate of the poverty rate by the chosen PBM-definition line in the past baseline)

$$-4.4 \div 80.4 = -5.5 \text{ percent.}$$

- c. The “grand” spliced estimate of the net number of participants who crossed from below the chosen government or old-definition line to above it (or from below the chosen PBM-definition line to above it) since the past baseline is the negative of the “grand” spliced estimate of change 10a expressed as a proportion, multiplied by the number of participants in the past baseline

Assuming for the sake of this hypothetical illustration that there were 10,000 participants in the baseline population, $-(-0.044) \times 10,000 = 440$.

The following summarizes the process in the hypothetical illustration above. It focuses on estimates of changes in poverty rates.

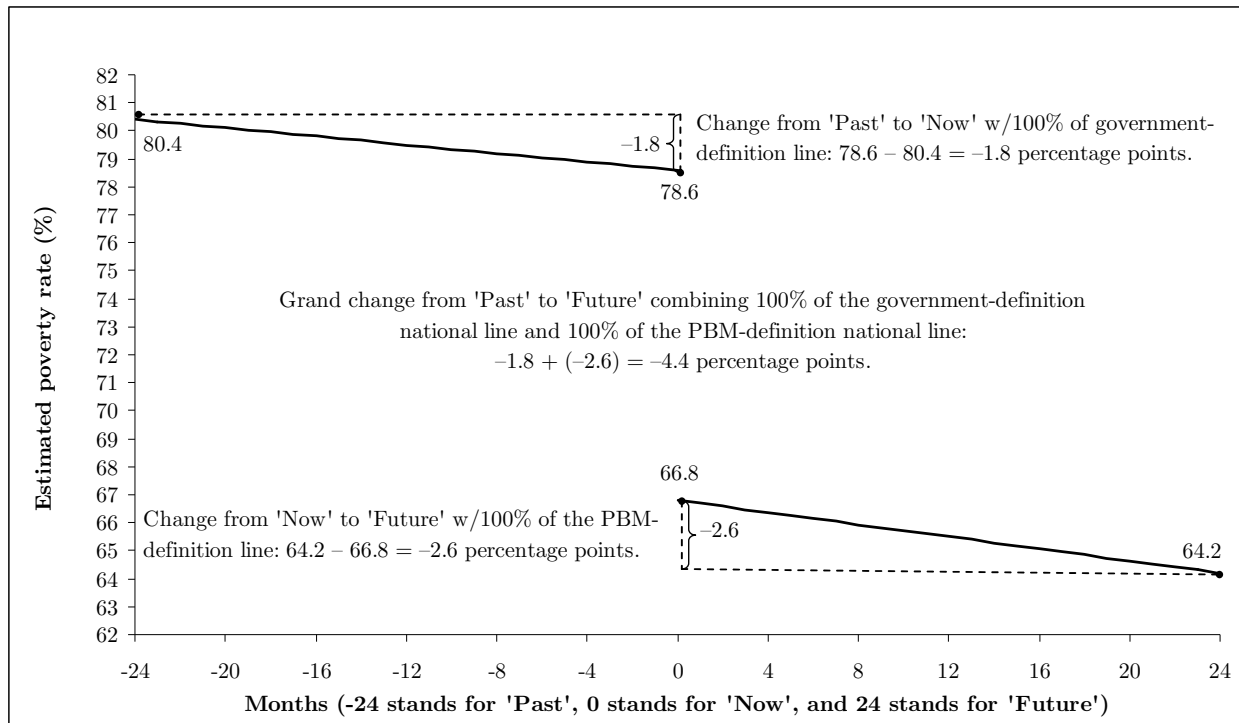
Selected poverty lines: 100% of national line (government-definition and PBM-definition)

Scores and poverty likelihoods of sampled households for 100% of the national line

Past		"Now"			Future	
Score	Pov. like. (govt.-def., old card) (%)	Score	Pov. like. (govt.-def., new card) (%)	Pov. like. (PBM-def.) (%)	Score	Pov. like. (newPBM.) (%)
15	88.9	21	87.4	77.6	22	77.6
20	82.5	26	80.0	64.8	27	64.8
25	70.0	31	72.0	55.1	37	47.1
Bias	+0.1	—	+1.2	-1.0	—	-1.0
Est. pov. rate (%)	80.4	—	78.6	66.8	—	64.2

Estimated change between:

Past and now (hybrid): $78.6 - 80.4 = -1.8$ percentage points
 Now and future (non-hybrid): $64.2 - 66.8 = -2.6$ percentage points
 Past and future ("grand" spliced): $-1.8 + (-2.6) = -4.4$ percentage points



Guidelines for the Interpretation of Indicators

The following comes from:

National Statistical Office. (2010) “Enumerator Manual for the Household Questionnaire”, [the *Manual*], go.worldbank.org/6A7GUDQ1Q0, retrieved 20 April 2012,

and

National Statistical Office. (2010) “Household Questionnaire, Third Integrated Household Survey, 2010/11”, [the *Questionnaire*], go.worldbank.org/6A7GUDQ1Q0, retrieved 20 April 2012.

When an issue arises that is not addressed here, its resolution should be left to the unaided judgment of the enumerator, as that apparently was the practice of Malawi’s NSO in the 2010/11 IHS. That is, an organization using the scorecard should not promulgate any definitions or rules (other than those in these “Guidelines”) to be used by all its field agents. Anything not explicitly addressed in these “Guidelines” is to be left to the unaided judgment of the individual enumerator.

General Guidelines

Fill out the scorecard header and the “Back-page Worksheet” first, following the directions on the “Back-page Worksheet”. In particular, do not ask the first scorecard indicator directly. Instead, use the information recorded on the “Back-page Worksheet” to determine the proper response for the first indicator.

According to p. 20 of the *Manual*, “Read the questions exactly as they are written in the questionnaire, following the established order.”

Do not ask the third, fourth, and fifth questions of the respondent. These questions are concerned with the predominant construction material of the floor, outer walls, and roof of the main dwelling. Instead of asking these questions of the respondent, you should record the answer based on your observation of the dwelling as the enumerator. If the predominant construction material is not obvious to you from your observation, then you should go ahead and ask the question of the respondent.

Do not read the response options for any questions to the respondent. Unless instructed otherwise here, read the question, and then stop; wait for a response. If the respondent asks for clarification or otherwise hesitates or seems confused, then read the question again or provide additional assistance based on these “Guidelines” or as you, the enumerator, deem appropriate.

In general, you should accept the responses given by the respondent. Nevertheless, if the respondent says something—or if you see or sense something—that suggests that the response may not be accurate, that the respondent is uncertain, or that the respondent desires assistance in figuring out how to respond, then you should read the question again and provide whatever assistance you deem appropriate based on these “Guidelines”.

While most indicators in the scorecard are verifiable, you do not—in general—need to verify responses. You should verify a response only if something suggests to you that the response may not be accurate and thus that verification might improve data quality. For example, you might choose to verify if the respondent hesitates, seems nervous, or otherwise gives signals that he/she may be lying or be confused. Likewise, verification is probably appropriate if a child in the household or a neighbor says something that does not square with the respondent’s answer. Verification is also a good idea if you can see something yourself—such as a consumer durable that the respondent avers not to have, or a child eating in the room who has not been counted as a member of the household—that suggests that the response may not be accurate.

In general, the application of the scorecard should mimic as closely as possible the application of the 2010/11 IHS. For example, poverty-scoring interviews should take place in respondents’ homestead because the 2010/11 IHS took place in respondents’ homesteads. According to p. 5 of the *Manual*, “The most reliable and complete means of gathering information is by visiting respondents at their homes.”

Questionnaire Translation:

The 2010/11 IHS left translation of the survey instrument to each individual enumerator (or to local translators). They did this translation to languages other than English—when needed—on the fly.

While the application of the scorecard should, in general, mimic the application of the 2010/11 IHS, it nevertheless makes sense to have a standard, well-done, checked translation to languages and dialects that are common in Malawi (such as Chichewa, among others). Without a standard translation, the variation in translations and interpretations across enumerators could greatly harm data quality. Any translation should reflect the meaning in the original English IHS survey instrument as closely as possible. Ideally, all organizations using the scorecard in a given dialect or language in Malawi would coordinate and use a single translation.

According to p. 13 of the *Manual*, “The questionnaire is in English. Most of the households to whom you will administer this questionnaire will not be able to respond to the questions if they are asked in English. Consequently, you must translate the questions into a language in which the survey household members are fluent. There are three points to bear in mind.

“First, there are several key terms . . . including *household* and *head of household* . . . [that] should always be translated into local languages using the exact same words. The questions have been carefully worded to ensure that the desired concept is being asked. Study the questions so that you can ask them in a consistent and natural manner. If this is not done, then responses to the same question across households may not be comparable.”

The following table translates some key terms to Chichewa, Chitumbuka, and Chiyao (pp. 79–80 of the *Manual*). The NSO provided this table to enumerators during their training for the 2010/11 IHS. The *Manual*, however, was written before the training, and it indicates that the translations would be discussed and finalized during enumerator training, and this project has no record of what terms were finally agreed upon there.

English	Chichewa	Chiyao	Chitumbuka
Head of household	Mkulu wolamulira mnyumba kapena pa banja	Mtwee waliwasa; jwakulamulila pewasa	Uyo wali namazaza
Household	Panyumba; banja	Nyumba/liwasa	Banja
Poverty	Umphawi	Kulaga; usauchi; yakunonopa	Ukavu
Respondent	Oyankha	Wakwanga iusyo	Wakuzgora

“Secondly, the questionnaire should be administered in a language that the survey household members understand fluently. If you find that you have been assigned to do interviews in an area in which most survey households are only fluent in a language in which you are not fluent . . . then an enumerator fluent in the language of that area should be assigned to do the interviews.

“Third and finally, do not assume that your skills in Chichewa will allow you to conduct interviews throughout Malawi. Although Chichewa is the national language of Malawi, many rural residents are not fluent in the language. This is particularly the case in northern Malawi where Chichewa is not commonly spoken and in the lakeshore areas, where Yao is the predominant language spoken in the villages. If you know that, because of language difficulties, you will be unable to efficiently and accurately administer the questionnaire in the area to which you have been assigned, you should immediately make this fact known to your supervisor.”

Confidentiality:

According to p. 4 of the *Manual*, “The information collected is confidential and must not be divulged to any unauthorized person.”

According to p. 5 of the *Questionnaire*, you should tell the respondent before the interview that “all of your answers will be held in confidence. The answers which you and the members of your household give me will only be used by [my organization].”

Who to interview:

According to p. 49 of the *Manual*, the poverty-scoring survey should be directed “primarily to the head of household, who may be assisted by other informed adults within the household. In the absence of the head of household, the most-informed adult member of the household should be selected as the respondent.”

According to p. 5 of the *Questionnaire*, the preferred respondent is the head of the household, or—if not the head—then the spouse of the head.

According to p. 28 of the *Manual*, the survey (or specific questions within the survey) should be administered “to the best-informed individual in the household on the topic”.

According to p. 9 of the *Manual*, “The head of household—assisted by other household members if necessary—should be asked questions that concern the household as a whole.”

According to pp. 10–11 of the *Manual*, “The *head of household* is the person commonly regarded by the household members as their head. The head is usually the main income-earner and decision-maker for the household, but you should accept the decision of the household members as to who is their head.

“There is one—and only one—head in the household. If more than one individual in a potential household claims headship, or if individuals within a potential household give conflicting statements as to who is the head, then you may be dealing with two or more households, rather than one.”

According to p. 29 of the *Manual*, “The household head need not be resident in the household. The household head is the person whom other household members designate as their head, but it is possible that this person will not be residing in the household at the time of the survey.”

According to p. 33 of the *Manual*, “It is possible that the household head may not be residing in the dwelling at the time of the interview. He or she may be living and

working—temporarily or permanently—in another part of Malawi or in another country.”

Administering the interview:

According to pp. 16–17 of the *Manual*, “Before you go to a selected household, you should ensure that you are ready to begin the interview, that is, that you are presentable and that you know how you are going to begin the interview. . . .

“The setting of the [interview] should be relatively private. Some of the questions being asked are of a personal and private nature. You should respect the desire of the respondents for privacy.

“No [unauthorized people] should come with you when you interview. If an [authorized person] does accompany you to an interview, you should always be sure to introduce him/her to the respondent, making clear to the respondent the purpose of the [presence of the authorized person]. . . .

“Any other persons unrelated to the survey or to the household should not accompany you while introducing yourself to the household or be present during the interview. If any such individuals are present when you begin your interview, you must politely request them to leave in order to respect the privacy of the survey household. If they cannot leave at that time, then you should schedule the interview for a later time or move to a more appropriate place, when or where greater privacy can be assured. In the event that the respondent requests to be accompanied by a non-household member, you have to honor that request.”

According to p. 18 of the *Manual*, “When you first arrive in a rural area, present yourself to the local village head to explain why you are going to be working in the area. You will be provided with an official letter of explanation from [your organization] to show them. In urban locations, identifying a local leader is more problematic. Make inquiries as to who might be considered local leaders when you first come to an urban area. These may be local businesses, religious leaders, community police, or political leaders.

“Do not be secretive about your work. Please explain what you are doing to all community members who ask about your activities. You should be respectful, courteous, and patient with all community members. The quality of your work is, to a large degree, dependent on the level of cooperation that you receive from the members of the communities in which the survey households reside.

“If the general community attitude towards your work is negative, then you will likely experience problems as you conduct interviews with the survey households in that community. You are expected to do all that you can in order to cultivate a courteous relationship with the community as a whole.

“While your work should not be secretive, you must, however, respect the confidentiality and privacy of the survey household respondents when administering the

questionnaire. Non-household members should not be present while you are conducting your interviews (unless specifically requested by the household).

“If you want to have a good reception from the community, then they should be clear on what exactly you are doing.” In particular, communicate to them that you are doing a survey of participants with [your organization] in order to understand better how they live.

Interactions with respondents:

According to pp. 18–19 of the *Manual*, “You should always be courteous and tactful in your dealings with respondents.

“Above all, your attitude towards respondents must be one of respect. You must always be patient towards the members of survey households. Be business-like in your conduct, never bullying, nor demanding, nor rude. Always act in a way that warrants respect and cooperation from the respondent.

“During your interviews, you should work efficiently and relatively quickly, but you should not rush the respondent nor make unnecessary mistakes.

“After each interview, you should always quickly go through your questionnaire [to check for errors or omissions] and thank each interviewee for his/her help and time. This is vital if the survey is to be carried out successfully. . . .

“It is vital that your behaviour be beyond reproach. You will find work more pleasant if you remain polite and friendly to everyone at all times.

“Be willing to answer any questions the respondents ask you about the survey and its contents. . . .

“You must seek to develop a smooth, flowing interviewing style so that you can obtain all of the information required as efficiently as possible. . . .

“Do not unnecessarily test the respondent’s patience by delaying the interview in any way, particularly through excessive probing on questions that the respondent feels that he/she has already answered to the best of his/her ability and recollection. Your interview technique will improve as you gain experience. These guidelines should help you. Find a balance between:

- Maintaining a smooth, flowing, continuous dialogue that allows you to obtain all of the information required in the shortest possible time, that is, without testing the patience of the respondents by delaying the interview in any way
- Allowing the respondents to ask any questions that they have about the survey so that they are convinced of its value and so that they are cooperative. . . . Do not encourage any questions from the respondents on issues unrelated to the survey, such as politics, religion, sports, etc.

“In conducting an interview, if it is clear that the respondent has understood the question that you have asked, then you must accept whatever response the respondent

provides. Probing questions can be used to make sure the respondent understands the key element of the question being asked.

“Never second-guess the respondent or make the assumption that you have a better understanding of the condition of the household than the respondent does. Your job is not to verify that the information provided is correct. . . . It is always possible that the respondent will lie to you or provide inaccurate information, but you, as the enumerator, should not make any judgements on the information provided. . . .

“There are exceptions, of course. At all stages of the interview, you should be alert to errors. These can be accidental or deliberate. You can never force people to give answers that they do not want to give, but you can approach the true facts through diplomatic and intelligent interviewing. For example, if the respondent says that the household has no livestock but there are chickens pecking at your feet or goats tied up nearby, you should inquire about these animals. However, do not probe excessively after seeking initial clarification.

“In any case, you should never go outside of the household to get information. This is beyond the scope of your work. Instead, you should always instill trust among the household members.

“Ultimately, you should not assess whether responses are ‘right’ or ‘wrong’. The questionnaire is being administered to the survey household members as we rightly expect that they will be able to provide the best information about their own living conditions.”

Guidelines for specific scorecard indicators

1. How many members does the household have?
 - A. Seven or more
 - B. Six
 - C. Five
 - D. Four
 - E. One, two, or three

According to pp. 10–11 of the *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 is not a current member of the household. . . .

“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 who are living in the same compound or dwelling are necessarily members of the same household. Two brothers who live in the same dwelling and who each have their own wives and children may or may not form a single housekeeping arrangement.

“There is a distinction between *family* and *household*. A *family* reflects social relationships, blood descent, and marriage. The concept of *household* is used here to identify an economic unit. While families and households are often the same, this is not always 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 may be distributed over two or more dwellings. If these dwelling units are in the same compound or nearby and if they have a common housekeeping arrangement with a common household budget, then the residents of these separate dwelling units should be treated as one household. . . .

“Keep in mind when listing household members:

- It is possible that the household head may not be residing in the dwelling at the time of the interview. He/she may be living and working, temporarily or permanently, in another part of Malawi or in another country
- Boarding-school students who are residing at a boarding school but who are still dependent on the household should be listed as *household members*
- Do not count as *household members* military personnel, prisoners, or other individuals who are residing elsewhere (in such institutions) and who are not primarily dependent on the household for their welfare
- Some household members may not be relatives of the household head. For example, a servant who lives in the household and who does not keep a household elsewhere is counted as a *household member*
- Servants, other hired workers, and lodgers (individuals who pay to reside in the dwelling of the household) should not be listed as *household members* if they have their own household elsewhere which they head or upon which they are dependent
- Children who are living with other relatives (for example, an aunt or uncle) should not be listed as *household members*”

According to p. 12 of the *Manual*, “It is possible for a single household to reside in several dwellings or for several households to reside in a single dwelling.”

According to p. 6 of the *Questionnaire*: “Make a complete list of *all individuals who normally live and eat their meals together in this household*, starting with the head of the household.”

“In order to make a comprehensive list of individuals connected to the household, use the following questions:

“First, say: ‘Please give me the names of all the members of your *immediate family* who normally live and eat their meals together here.’ Write down the names. List the head of the household first.

“Then say: ‘Please give me the names of any *other persons related to you or other household members* who normally live and eat their meals together here.’

“Then say: ‘Are there any other persons not here now who normally live and eat their meals here? For example, household members studying elsewhere or travelling.’

“Then say: ‘Please give me the names of any *other persons not related to you or other household members* but who normally live and eat their meals together here, such as servants, lodgers, or others who are not relatives.’

“Do not list servants who have a household elsewhere, nor guests who are visiting temporarily and who have a household elsewhere.”

2. Is the (oldest) female head/spouse able to read and write in Chichewa or English?
- A. No
 - B. Yes, only Chichewa
 - C. Yes, English (regardless of Chichewa)
 - D. No female head/spouse

According to pp. 10–11 of the *Manual*, “The *head of household* is the person commonly regarded by the household members as their head. The head is usually the main income-earner and decision-maker for the household, but you should accept the decision of the household members as to who is their head.

“There is one—and only one—head in the household. If more than one individual in a potential household claims headship, or if individuals within a potential household give conflicting statements as to who is the head, then it is likely that you are dealing with two or more households, rather than one.”

For the purposes of the scorecard, the (oldest) *female head/spouse* is defined as:

- The household head, if the head is female
- The (oldest) spouse/conjugal partner of the household head, if the head is male
- Non-existent, if the head is male and if he does not have a spouse/conjugal partner who is also a member of the household

If the head is male, and if he has more than one wife who is a member of the interviewed household, then for the purposes of this question you should consider only his oldest wife. This is not necessarily the wife he married first, but rather the wife who is the oldest in terms of her age since birth.

According to p. 29 of the *Manual*, “The household head need not be resident in the household. The household head is the person that other household members designate as their head, but it is possible that this person will not be residing in the household at the time of the survey.”

According to p. 33 of the *Manual*, “It is possible that the household head may not be residing in the dwelling at the time of the interview. He or she may be living and working—temporarily or permanently—in another part of Malawi or in another country.”

The *Manual* has no additional information for this indicator. In particular, it does not indicate that you should ask for any proof or verification of ability of the (oldest) female head/spouse to read and write. This suggests that you are to take the respondent’s word, unless, as discussed earlier, there is something that suggests to you that the response may not be accurate.

The question refers to being able to read *and* write, not just read *or* write. Mark response option “A. Yes only Chichewa” or “B. Yes, English (regardless of Chichewa)” only if the person can both read and write in the relevant language.

Combine the responses related to reading, writing, Chichewa, and English as follows:

Female head/spouse?	Chichewa		English		Response
	Read	Write	Read	Write	
No	N/A	N/A	N/A	N/A	D
Yes	No	No	No	No	A
Yes	Yes	No	No	No	A
Yes	No	Yes	No	No	A
Yes	Yes	Yes	No	No	B
Yes	No	No	Yes	No	A
Yes	Yes	No	Yes	No	A
Yes	No	Yes	Yes	No	A
Yes	Yes	Yes	Yes	No	B
Yes	No	No	No	Yes	A
Yes	Yes	No	No	Yes	A
Yes	No	Yes	No	Yes	A
Yes	Yes	Yes	No	Yes	B
Yes	No	No	Yes	Yes	C
Yes	Yes	No	Yes	Yes	C
Yes	No	Yes	Yes	Yes	C
Yes	Yes	Yes	Yes	Yes	C

3. The floor of the main dwelling is predominantly made of what material?
 - A. Smoothed mud, or sand
 - B. Smooth cement, wood, tile, or other

This question pertains to the finish of the floor.

According to pp. 22 and 49 of the *Manual*, you should not ask this question of the respondent. Rather, you as the enumerator should record the answer based on your observation of the dwelling.

If the answer is not clear to you from your observation, then ask the respondent.

According to p. 50 of the *Manual*, “If the enumerator cannot see the floor inside the dwelling unit, then this question could be asked to the respondent.”

According to p. 50 of the *Manual*, “the [household’s] *dwelling* is defined as any structure (permanent, semi-permanent, or traditional) where people live and sleep. It may be a hut, house, a store with a sleeping room or rooms at the back or sides, a shelter of reeds/straw such as those used by fishermen, or any other structure where people sleep.

“Dwellings made up of several separate structures are most commonly found in rural areas, as where separate sleeping huts are constructed for various members of a household.”

According to p. 50 of the *Manual*, “If two or more different types of materials are used for the floor, then you must report the material that is used in the majority.”

4. The outer walls of the main dwelling of the household are predominantly made of what material?
 - A. Mud (*yomata*), or grass
 - B. Mud brick (unfired)
 - C. Compacted earth (*yamdindo*), burnt bricks, concrete, wood, iron sheets, or other

According to pp. 22 and 49 of the *Manual*, you should not ask this question of the respondent. Rather, you as the enumerator should record the answer based on your observation of the dwelling.

If the answer is not clear to you from your observation, then ask the respondent.

According to p. 50 of the *Manual*, “If the enumerator cannot see the outer walls the dwelling unit, then this question could be asked to the respondent.”

According to p. 50 of the *Manual*, “the [household’s] *dwelling* is defined as any structure (permanent, semi-permanent, or traditional) where people live and sleep. It may be a hut, house, a store with a sleeping room or rooms at the back or sides, a shelter of reeds/straw such as those used by fishermen, or any other structure where people sleep.

“Dwellings made up of several separate structures are most commonly found in rural areas, as where separate sleeping huts are constructed for various members of a household.”

According to p. 50 of the *Manual*, “If two or more different types of materials are used for the outer walls, then you must report the material that is used in the majority.”

5. The roof of the main dwelling is predominantly made of what material?
 - A. Grass, plastic sheeting, or other
 - B. Iron sheets, clay tiles, or concrete

According to pp. 22 and 49 of the *Manual*, you should not ask this question of the respondent. Rather, you as the enumerator should record the answer based on your observation of the dwelling.

If the answer is not clear to you from your observation, then ask the respondent.

According to p. 50 of the *Manual*, “If the enumerator cannot see the roof of the dwelling unit, then this question could be asked to the respondent.”

According to p. 50 of the *Manual*, “the [household’s] *dwelling* is defined as any structure (permanent, semi-permanent, or traditional) where people live and sleep. It may be a hut, house, a store with a sleeping room or rooms at the back or sides, a shelter of reeds/straw such as those used by fishermen, or any other structure where people sleep.

“Dwellings made up of several separate structures are most commonly found in rural areas, as where separate sleeping huts are constructed for various members of a household.”

According to p. 50 of the *Manual*, “If two or more different types of materials are used for the roof, then you must report the material that is used in the majority.”

According to p. 50 of the *Manual*, “Concrete can be counted as a *roof* in the case in which there is a flat roof [when] the building has an unfinished floor above it.”

6. What kind of toilet facility does the household use?
- A. None, traditional latrine without roof shared with other households, or other
 - B. Traditional latrine without roof only for household members
 - C. Traditional latrine with roof shared with other households
 - D. Traditional latrine with roof only for household members, VIP latrine, or flush toilet

According to p. 53 of the *Manual*:

- “*Flush toilet* is characterized by the draining of human excreta by a rush of running water
- A *VIP latrine* is a Ventilated Improved Pit latrine. It is defined as an on-site means of human-excreta disposal in a hygienic, low-cost, and more-acceptable manner. The primary features of VIP latrines consist of an enclosed structure (roof and walls) with a large diameter (110mm) PVC vertical ventilation pipe running outside the structure from the pit of the latrine to vent above the roof. VIP latrines often will have concrete slabs containing the latrine hole
- A *traditional latrine with a roof* is an ordinary pit latrine built without health or hygienic-related specifications. It does not have a vent pipe, but it does have a roof
- A *traditional latrine without a roof* is an ordinary pit latrine built without health or hygienic-related specifications. It has neither a vent pipe nor a roof
- *No facility [none]* refers to the use of bushes, grass/field, or other open spaces as toilet facilities”

Combine the response related to the type of toilet arrangement with the response related with whether the toilet arrangement is shared as follows:

Toilet arrangement	Shared with other households?	Response
None	Yes	A
None	No	A
Traditional latrine without roof	Yes	A
Traditional latrine without roof	No	B
Other	Yes	A
Other	No	A
Traditional latrine with roof	Yes	C
Traditional latrine with roof	No	D
VIP latrine	Yes	D
VIP latrine	No	D
Flush toilet	Yes	D
Flush toilet	No	D

7. What is the household's main source of lighting fuel?
 - A. Collected firewood, purchased firewood, grass, or gas
 - B. Paraffin, or other
 - C. Battery/dry cell (torch), candles, or electricity

According to p. 51 of the *Manual*, "A rechargeable lamp should be counted under *other*."

8. Do any members of the household sleep under a bed net to protect against mosquitos at some time during the year?
 - A. No
 - B. Yes

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

9. Does the household own any tables?
- A. No
 - B. Yes

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

10. Does the household own any beds?

- A. No
- B. Yes

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

Figure 1: PBM-definition national poverty lines (and the line marking the poorest half of people below 100% of the PBM-definition national line) and poverty rates for all of Malawi and for each of Malawi's four poverty-line regions and for construction/validation samples, by households and people, for 2004/5 and 2010/11

Sample	Year	Line or Rate	HHs or people	HHs surveyed	% with consumption below a poverty line				
					National			Poorest half	
					Food	100%	150%	200%	below 100% natl.
All Malawi	2004/5	Line			26.02	43.23	64.85	86.46	29.88
		Rate	HHs	11,280	12.8	39.2	63.3	77.2	18.0
		Rate	People		17.1	47.0	71.1	83.2	23.5
	2010/11	Line			59.59	92.82	139.23	185.65	63.65
		Rate	HHs	12,271	14.2	32.4	53.2	67.9	15.4
		Rate	People		17.9	38.8	60.7	74.8	19.4
Urban	2004/5	Line			30.28	56.99	85.49	113.98	39.24
		Rate	HHs	1,440	6.3	30.8	51.8	64.3	14.0
		Rate	People		9.0	37.6	60.0	71.0	18.8
	2010/11	Line			65.75	128.08	192.12	256.16	89.74
		Rate	HHs	2,233	3.4	21.6	39.9	52.9	10.7
		Rate	People		4.7	27.4	47.3	61.1	13.7
North rural	2004/5	Line			29.73	46.08	69.11	92.15	29.50
		Rate	HHs	1,440	23.9	50.7	71.7	84.8	23.6
		Rate	People		30.1	59.4	79.4	89.6	29.7
	2010/11	Line			64.88	95.90	143.85	191.80	63.29
		Rate	HHs	1,758	20.8	41.4	64.8	79.0	19.3
		Rate	People		25.7	48.2	72.7	85.8	24.1
Central rural	2004/5	Line			26.12	43.76	65.64	87.52	31.19
		Rate	HHs	3,840	9.0	33.7	59.7	76.1	15.8
		Rate	People		11.8	40.0	66.5	81.4	20.0
	2010/11	Line			59.98	88.77	133.16	177.55	61.88
		Rate	HHs	3,485	13.6	29.7	50.6	66.3	14.4
		Rate	People		16.2	34.1	56.2	71.6	17.1
South rural	2004/5	Line			23.80	38.16	57.24	76.33	26.12
		Rate	HHs	4,560	15.3	43.6	67.7	80.2	19.7
		Rate	People		21.1	53.1	76.4	86.8	26.5
	2010/11	Line			55.16	81.54	122.30	163.07	54.90
		Rate	HHs	4,795	17.1	36.4	57.6	72.2	17.0
		Rate	People		22.7	45.1	67.0	80.2	22.5
Construction and calibration (Selecting indicators and weights, and associating scores with likelihoods)									
	2010/11	Rate	HHs	7,195	14.2	32.3	53.4	67.9	15.6
Validation (Measuring accuracy)									
	2004/5	Rate	HHs	11,280	12.8	39.2	63.3	77.2	18.0
	2010/11	Rate	HHs	4,085	14.2	32.4	53.0	67.9	15.1

Source: Malawi's 2004/5 and 2010/11 IHS

Poverty lines for 2004/5 are in MWK in average prices for all of Malawi in Feb./March 2004.

Poverty lines for 2010/11 are in MWK in average prices for all of Malawi in Feb./March 2010.

Consumption is daily per-capita and is measured according to the BPM definition.

Figure 1: PBM-definition international 2005 and 2011 PPP poverty lines and poverty rates for all of Malawi and for each of Malawi's four poverty-line regions and for construction/validation samples, by households and people, for 2004/5 and 2010/11

Sample	Year	Line or Rate	HHs or people	HHs surveyed	% with consumption below a poverty line						
					International 2005 PPP					International 2011 PPP	
					\$1.25	\$2.00	\$2.50	\$5.00	\$8.44	\$1.90	\$3.10
All Malawi	2004/5	Line			62.37	99.78	124.73	249.46	421.09	77.34	126.18
		Rate	HHs	11,280	61.1	82.5	88.8	97.4	99.1	71.9	89.0
		Rate	People		69.0	87.5	92.4	98.3	99.4	78.8	92.5
	2010/11	Line			133.90	214.25	267.81	535.62	904.12	166.05	270.92
		Rate	HHs	12,271	51.1	74.2	83.0	95.6	98.4	62.5	83.2
		Rate	People		58.7	80.7	88.1	97.1	98.9	69.8	88.2
Urban	2004/5	Line			82.21	131.54	164.43	328.85	555.10	101.95	166.34
		Rate	HHs	1,440	50.5	70.5	78.3	90.8	96.2	59.8	78.5
		Rate	People		58.6	76.5	83.1	92.6	97.4	67.4	83.3
	2010/11	Line			184.77	295.62	369.53	739.06	1,247.54	229.12	373.83
		Rate	HHs	2,233	37.5	59.6	70.0	88.4	95.1	48.2	70.3
		Rate	People		45.0	67.5	76.7	91.3	96.3	56.5	76.9
North rural	2004/5	Line			66.47	106.35	132.94	265.87	448.79	82.42	134.48
		Rate	HHs	1,440	70.1	89.1	93.2	99.0	99.6	80.6	93.3
		Rate	People		77.9	92.4	95.5	99.5	99.8	86.6	95.6
	2010/11	Line			138.34	221.35	276.68	553.37	934.09	171.55	279.90
		Rate	HHs	1,758	62.7	84.4	91.3	98.6	99.7	74.0	91.5
		Rate	People		70.7	90.2	94.8	98.9	99.9	81.7	94.9
Central rural	2004/5	Line			63.13	101.00	126.25	252.50	426.23	78.28	127.72
		Rate	HHs	3,840	57.3	82.0	89.5	98.3	99.6	69.7	89.7
		Rate	People		64.1	86.7	92.7	99.0	99.7	75.8	92.8
	2010/11	Line			128.06	204.90	256.12	512.24	864.66	158.80	259.10
		Rate	HHs	3,485	48.5	73.7	83.3	96.6	98.7	60.5	83.6
		Rate	People		54.1	78.9	87.5	97.8	99.1	66.1	87.7
South rural	2004/5	Line			55.05	88.09	110.11	220.21	371.72	68.27	111.39
		Rate	HHs	4,560	65.3	84.8	90.2	98.2	99.4	75.2	90.4
		Rate	People		74.2	90.2	93.9	98.9	99.7	82.7	94.0
	2010/11	Line			117.62	188.19	235.24	470.49	794.18	145.86	237.98
		Rate	HHs	4,795	55.5	77.5	85.6	96.7	99.1	66.7	85.8
		Rate	People		65.1	85.0	91.2	98.3	99.5	75.3	91.4
Construction and calibration (Selecting indicators and weights, and associating scores with likelihoods)											
	2010/11	Rate	HHs	7,195	51.2	74.4	83.1	95.5	98.4	62.6	83.3
Validation (Measuring accuracy)											
	2004/5	Rate	HHs	11,280	61.1	82.5	88.8	97.4	99.1	71.9	89.0
	2010/11	Rate	HHs	4,085	51.0	74.0	83.0	95.7	98.4	62.4	83.2

Source: Malawi's 2004/5 and 2010/11 IHS

Poverty lines for 2004/5 are in MWK in average prices for all of Malawi in Feb./March 2004.

Poverty lines for 2010/11 are in MWK in average prices for all of Malawi in Feb./March 2010.

Consumption is daily per-capita and is measured according to the BPM definition.

Figure 1: Government-definition national poverty lines (and the line marking the poorest half of people below 100% of the government-definition national line) and poverty rates for all of Malawi and for each of Malawi's four poverty-line regions and for construction/validation samples, by households and people, for 2004/5 and 2010/11

Sample	Year	Line or Rate	HHs or people	HHs surveyed	% with consumption below a poverty line				
					National				Poorest half below 100% natl.
					Food	100%	150%	200%	
All Malawi	2004/5	Line			27.25	43.92	65.87	87.83	29.75
		Rate	HHs	11,280	16.6	43.6	67.9	80.6	19.6
		Rate	People		22.3	52.4	75.9	86.5	26.2
	2010/11	Line			58.52	94.33	141.49	188.66	60.50
		Rate	HHs	12,271	19.8	43.7	64.2	77.1	20.5
		Rate	People		24.5	50.7	71.3	82.7	25.3
Urban	2004/5	Line			31.05	50.04	75.07	100.09	37.30
		Rate	HHs	1,440	5.0	19.9	40.2	54.4	9.0
		Rate	People		7.5	25.4	48.0	62.6	12.7
	2010/11	Line			62.57	100.85	151.27	201.69	72.98
		Rate	HHs	2,233	3.1	13.5	29.1	42.9	6.4
		Rate	People		4.3	17.3	36.3	50.8	8.6
North rural	2004/5	Line			29.81	48.04	72.07	96.09	31.34
		Rate	HHs	1,440	18.5	46.3	70.8	85.0	20.4
		Rate	People		25.9	56.3	79.3	90.1	28.1
	2010/11	Line			64.86	104.55	156.82	209.09	66.36
		Rate	HHs	1,758	23.3	52.3	74.5	85.9	24.0
		Rate	People		29.0	59.9	81.6	91.4	30.0
Central rural	2004/5	Line			26.03	41.96	62.94	83.93	30.25
		Rate	HHs	3,840	12.1	38.7	66.5	80.9	18.1
		Rate	People		16.1	46.7	74.2	86.7	23.4
	2010/11	Line			54.00	87.05	130.57	174.09	57.34
		Rate	HHs	3,485	18.6	43.4	65.9	80.5	21.0
		Rate	People		21.5	48.7	71.6	84.8	24.3
South rural	2004/5	Line			26.68	43.00	64.50	86.00	26.77
		Rate	HHs	4,560	23.3	53.8	76.2	86.7	23.7
		Rate	People		31.5	64.4	84.5	92.1	32.2
	2010/11	Line			59.34	95.64	143.46	191.28	56.73
		Rate	HHs	4,795	26.5	53.5	73.9	85.3	24.5
		Rate	People		34.2	63.3	82.1	91.1	31.7
Construction and calibration (Selecting indicators and weights, and associating scores with likelihoods)									
	2010/11	Rate	HHs	6,193	19.6	43.9	64.2	77.3	20.5
Validation (Measuring accuracy)									
	2004/5	Rate	HHs	11,280	16.6	43.6	67.9	80.6	19.6
	2010/11	Rate	HHs	6,078	19.9	43.4	64.3	77.0	20.4

Source: Malawi's 2004/5 and 2010/11 IHS

Poverty lines for 2004/5 are in MWK in average prices for all of Malawi in Feb./March 2004.

Poverty lines for 2010/11 are in MWK in average prices for all of Malawi in Feb./March 2010.

Consumption is daily per-capita and measured according to the government definition.

Figure 1: Government-definition international 2005 and 2011 PPP poverty lines and poverty rates for all of Malawi and for each of Malawi's four poverty-line regions and for construction/validation samples, by households and people, for 2004/5 and 2010/11

Sample	Year	Line or Rate	HHs or people	HHs surveyed	% with consumption below a poverty line						
					International 2005 PPP					International 2011 PPP	
					\$1.25	\$2.00	\$2.50	\$5.00	\$8.44	\$1.90	\$3.10
All Malawi	2004/5	Line			62.37	99.78	124.73	249.46	421.09	77.34	126.18
		Rate	HHs	11,280	64.9	84.3	90.4	97.8	99.1	75.3	90.6
		Rate	People		73.2	89.4	93.8	98.6	99.4	82.3	93.9
	2010/11	Line			133.96	214.34	267.92	535.84	904.50	166.12	271.04
		Rate	HHs	12,271	61.6	81.4	87.0	96.3	98.6	71.5	87.3
		Rate	People		68.8	86.2	90.7	97.6	90.0	77.9	90.8
Urban	2004/5	Line			71.07	113.71	142.13	284.27	479.84	88.13	143.79
		Rate	HHs	1,440	36.6	60.0	70.0	88.1	94.4	47.7	70.3
		Rate	People		44.2	67.8	76.4	90.8	95.8	55.6	76.7
	2010/11	Line			143.22	229.15	286.43	572.86	966.99	177.60	289.76
		Rate	HHs	2,233	27.0	48.5	58.2	83.3	92.6	35.8	58.7
		Rate	People		33.9	56.7	66.0	87.9	94.7	43.4	66.4
North rural	2004/5	Line			68.22	109.16	136.45	272.89	460.64	84.00	137.05
		Rate	HHs	1,440	67.9	88.6	93.4	99.4	99.8	78.6	93.8
		Rate	People		77.0	92.6	96.2	99.8	100.0	85.8	96.4
	2010/11	Line			148.47	237.55	296.93	593.87	1,002.45	184.11	300.39
		Rate	HHs	1,758	72.4	89.9	94.1	99.4	100.0	81.7	94.2
		Rate	People		80.0	93.9	96.4	99.6	100.0	88.3	96.5
Central rural	2004/5	Line			59.59	95.34	119.18	238.36	402.36	73.90	120.57
		Rate	HHs	3,840	63.0	85.2	92.1	99.0	99.8	74.8	92.2
		Rate	People		70.9	90.2	95.1	99.5	99.9	81.7	95.2
	2010/11	Line			123.62	197.79	247.23	494.46	834.66	153.29	250.11
		Rate	HHs	3,485	62.8	85.2	91.3	98.4	99.5	74.0	91.4
		Rate	People		68.4	88.7	93.8	99.1	99.8	79.2	93.9
South rural	2004/5	Line			61.06	97.70	122.13	244.26	412.30	75.72	123.55
		Rate	HHs	4,560	73.8	89.3	94.1	99.2	99.8	82.8	94.3
		Rate	People		82.5	93.9	96.8	99.6	99.9	89.4	96.9
	2010/11	Line			135.82	217.32	271.64	543.29	917.07	168.43	274.80
		Rate	HHs	4,795	71.3	88.8	92.9	98.8	99.7	80.6	93.1
		Rate	People		80.0	93.3	96.0	99.4	99.8	87.5	96.0
Construction and calibration (Selecting indicators and weights, and associating scores with likelihoods)											
	2010/11	Rate	HHs	6,193	61.4	81.5	86.9	96.5	98.6	71.5	87.1
Validation (Measuring accuracy)											
	2004/5	Rate	HHs	11,280	64.9	84.3	90.4	97.8	99.1	75.3	90.6
	2010/11	Rate	HHs	5,087	61.9	81.4	87.2	96.1	98.5	71.4	87.5

Source: Malawi's 2004/5 and 2010/11 IHS

Poverty lines for 2004/5 are in MWK in average prices for all of Malawi in Feb./March 2004.

Poverty lines for 2010/11 are in MWK in average prices for all of Malawi in Feb./March 2010.

Consumption is daily per-capita and measured according to the government definition.

Figure 1: Old-definition international 2005 PPP poverty lines and poverty rates for all of Malawi and for each of Malawi's four poverty-line regions and for construction/validation samples, by households and people, for 2004/5 and 2010/11

Sample	Year	Line or Rate	HHs or people	HHs surveyed	% with consumption < a poverty line	
					International 2005 PPP	
					\$1.25	\$2.50
All Malawi	2004/5	Line			63.60	127.20
		Rate	HHs		66.1	90.8
		Rate	People	11,280	74.2	94.0
	2010/11	Line			137.74	275.47
		Rate	HHs		62.9	87.7
		Rate	People	12,271	70.1	91.2
Urban	2004/5	Line			72.47	144.95
		Rate	HHs		38.3	70.7
		Rate	People	1,440	45.9	76.9
	2010/11	Line			147.25	294.51
		Rate	HHs		28.5	59.6
		Rate	People	2,233	35.8	67.4
North rural	2004/5	Line			69.57	139.15
		Rate	HHs		68.9	93.8
		Rate	People	1,440	77.8	96.4
	2010/11	Line			152.65	305.31
		Rate	HHs		73.1	94.8
		Rate	People	1,758	80.5	97.1
Central rural	2004/5	Line			60.77	121.54
		Rate	HHs		64.3	92.4
		Rate	People	3,840	72.1	95.3
	2010/11	Line			127.10	254.20
		Rate	HHs		64.5	91.7
		Rate	People	3,485	70.2	94.2
South rural	2004/5	Line			62.27	124.55
		Rate	HHs		74.8	94.4
		Rate	People	4,560	83.3	96.9
	2010/11	Line			139.65	279.30
		Rate	HHs		72.4	93.4
		Rate	People	4,795	80.9	96.3
Construction and calibration (Selecting indicators and weights, and associating scores with likelihoods)						
	2010/11	Rate	HHs	6,193	62.7	87.5
Validation (Measuring accuracy)						
	2004/5	Rate	HHs	11,280	66.1	90.8
	2010/11	Rate	HHs	6,078	63.1	87.9

Source: Malawi's 2004/5 and 2010/11 IHS

Poverty lines for 2004/5 are in MWK in average prices for all of Malawi in Feb./March 2004.

Poverty lines for 2010/11 are in MWK in average prices for all of Malawi in Feb./March 2010.

Figure 2: Poverty indicators by uncertainty coefficient

<u>Uncertainty coefficient</u>	<u>Indicator (Responses ordered starting with those linked with higher poverty likelihoods)</u>
978	How many household members are 16-years-old or younger? (Five or more; Four; Three; Two; One; None)
970	How many household members are 15-years-old or younger? (Five or more; Four; Three; Two; One; None)
968	How many household members are 18-years-old or younger? (Six or more; Five; Four; Three; Two; One; None)
964	How many household members are 17-years-old or younger? (Five or more; Four; Three; Two; One; None)
936	How many household members are 14-years-old or younger? (Five or more; Four; Three; Two; One; None)
882	How many household members are 12-years-old or younger? (Four or more; Three; Two; One; None)
861	How many household members are 12-years-old or younger? (Four or more; Three; Two; One; None)
854	How many members does the household have? (Seven or more; Six; Five; Four; One, two, or three)
778	How many household members are 11-years-old or younger? (Four or more; Three; Two; One; None)
633	What does the household head sleep on? (Mat (grass) on floor, cloth/sack on floor, floor (nothing else), Bed and mat (grass), or other; Mattress on floor; Bed alone, or bed and mattress)
631	Are all household members ages 6 to 12 currently attending 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 ages 6 to 12)
610	Are all household members ages 6 to 13 currently attending 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 ages 6 to 13)
604	Is the (oldest) female head/spouse able to read and write in Chichewa or English? (No; Yes, only Chichewa; Yes, English (regardless of Chichewa); No female head/spouse)
600	Are all household members ages 6 to 14 currently attending 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 ages 6 to 14)
594	Are all household members ages 6 to 15 currently attending 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 ages 6 to 15)

Figure 2 (cont.): Poverty indicators by uncertainty coefficient

<u>Uncertainty coefficient</u>	<u>Indicator (Responses ordered starting with those linked with higher poverty likelihoods)</u>
594	Are all household members ages 6 to 16 currently attending 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 ages 6 to 16)
585	Are all household members ages 6 to 11 currently attending 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 ages 6 to 11)
563	The floor of the main dwelling is predominantly made of what material? (Smoothed mud, or sand; Smooth cement, wood, tile, or other)
556	Are all household members ages 6 to 17 currently attending 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 ages 6 to 17)
550	How many working cell phones in total does the household have? (None; One; Two or more)
501	Does the household own any radios ('wireless'), tape players, or CD players/hifi? (No; Yes)
497	What class is the male head/spouse in or what is the highest class level he ever attended? (None, or nursery/pre-school; Primary standard 1 or 2; Primary standard 3; Primary standard 4; Primary standard 5; Primary standard 6; Primary standard 7; Primary standard 8; No male head/spouse; Secondary form 1, 2, or 3; Secondary form 4, or higher)
497	What class is the (oldest) female head/spouse in or what is the highest class level she ever attended? (None, or nursery/pre-school; Primary standard 1 or 2; Primary standard 3; Primary standard 4; Primary standard 5; Primary standard 6; Primary standard 7; Primary standard 8; No female head/spouse; Secondary forms 1, 2, or 3; Secondary form 4, or higher)
495	Are all household members ages 6 to 18 currently attending 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 ages 6 to 18)
493	How many household members are 6-years-old or younger? (Three or more; Two; One; None)

Figure 2 (cont.): Poverty indicators by uncertainty coefficient

<u>Uncertainty coefficient</u>	<u>Indicator (Responses ordered starting with those linked with higher poverty likelihoods)</u>
474	What is the household's main source of lighting fuel? (Collected firewood, purchased firewood, grass, or gas; Paraffin, or other; Battery/dry cell (torch), candles, or electricity)
444	The roof of the main dwelling is predominantly made of what material? (Grass, plastic sheeting, or other; Iron sheets, clay tiles, or concrete)
441	Does the household own any beds? (No; Yes)
436	What general type of construction materials are used for the dwelling? (Traditional; Semi-permanent (mix of traditional (mud, grass) and modern materials (iron sheets, cement)); Permanent)
413	How many changes of clothes does the household head own? (number of trousers for men: skirts/dresses for women) (None, one, or two; Three; Four; Five or more)
407	What is the household's main source of cooking fuel? (Collected firewood, crop residue, or animal waste; Purchased firewood; Charcoal, electricity, paraffin, gas, saw dust, or other)
399	Does the household own any irons (for pressing clothes)? (No; Yes)
399	Does the household own any tables? (No; Yes)
384	Do you have electricity working in your dwelling? (No; Yes)
383	Does the household own any radios ('wireless')? (No; Yes)
383	Does the household own any chairs, tables, upholstered chairs, sofa sets, or coffee tables? (No; Yes)
374	Does the household own any tapes or CD players/hifi? (No; Yes)
344	Does the household own any upholstered chairs or sofa sets? (No; Yes)
311	What does the household head sleep under in the hot season (October)? (Nothing, or other; <i>Chitenje</i> cloth, fertilizer or grain sack, or clothes; Blanket only; Sheets only; Blanket and sheets)
277	The outer walls of the main dwelling of the household are predominantly made of what material? (Mud (<i>yomata</i>), or grass; Compacted earth (<i>yamdindo</i>); Mud brick (unfired), burnt bricks, concrete, wood, iron sheets, or other)
271	Does the household own any chairs? (No; Yes)

Figure 2 (cont.): Poverty indicators by uncertainty coefficient

<u>Uncertainty coefficient</u>	<u>Indicator (Responses ordered starting with those linked with higher poverty likelihoods)</u>
271	Does the household own any coffee tables (for sitting room)? (No; Yes)
265	In the last seven days, did the (oldest) female head/spouse do any casual, part-time work or ganyu labour, or did he do any work on household agricultural activities (including livestock- and fishing-related activities) or fishing, whether for sale or for household food? (Yes; No; No female head/spouse)
261	What does the household head sleep under in the cold season (July)? (<i>Chitenje</i> cloth, fertilizer or grain sack, clothes, nothing, or other; Sheets only; Blanket only; Blanket and sheets)
258	In the last seven days, how many household members did any casual, part-time work or <i>ganyu</i> labour or did any work on household agricultural activities (including livestock- and fishing-related activities) or fishing, whether for sale or for household food? (Four or more; Three; Two; One; None)
258	Is the male head/spouse able to read and write in Chichewa or English? (No, neither language; No male head/spouse; Yes (either or both languages))
234	What kind of rubbish disposal facilities does the household use? (None, or other; Public rubbish heap; Burning; Rubbish pit; Collected from rubbish bin)
219	Does the household own any refrigerators, kerosene/paraffin stoves, electric or gas stoves, hot plates, or sewing machines? (No; Yes)
207	What is the present marital status of the (oldest) female head/spouse? (Divorced; Separated; Polygamous married or non-formal union; Monogamous married or non-formal union; Widow; Never-married; No female head/spouse)
201	In the last seven days, did the (oldest) female head/spouse do any work on household agricultural activities (including livestock- and fishing-related activities) or fishing, whether for sale or for household food? (No; Yes; No female head/spouse)
170	What is the structure of household headship? (Both male and female heads/spouses; Female head/spouse only; Male head/spouse only)
163	In the last seven days, did any household members do any work for a wage, salary, commission, or any payment of any kind, excluding <i>ganyu</i> ? (No; Yes)
157	How many hoes does the household own? (None; One; Two; Three or more)

Figure 2 (cont.): Poverty indicators by uncertainty coefficient

<u>Uncertainty coefficient</u>	<u>Indicator (Responses ordered starting with those linked with higher poverty likelihoods)</u>
149	What is the present marital status of the male head/spouse? (Polygamous married or non-formal union; Monogamous married or non-formal union; No male head/spouse; Widower, separated, divorced, or never-married)
149	What kind of toilet facility does the household use? (None, traditional latrine without roof shared with other households, or other; Traditional latrine without roof only for household members; Traditional latrine with roof shared with other households; Traditional latrine with roof only for household members, VIP latrine, or flush toilet)
144	Does the household own any cupboards, drawers, or bureaus? (No; Yes)
139	In the last seven days, how many household members ran or helped with a non-agricultural and non-fishing household business, big or small, for him/herself? (None; One; Two or more)
132	During the last 12 months, did the male head/spouse or the (oldest) female head/spouse run or help with a non-agricultural and non-fishing household business, big or small, for him/herself? (No; Yes)
120	How many separate rooms do the members of the household occupy? (Do not count bathrooms, toilets, storerooms, or garage) (One or none; Two; Three; Four or more)
108	Do any members of the household sleep under a bed net to protect against mosquitos at some time during the year? (No; Yes)
84	In the last seven days, how many household members did any work on household agricultural activities (including livestock- and fishing-related activities) or fishing, whether for sale or for household food? (Two or more; One; None)
83	In the last seven days, did any household members do any casual, part-time work or <i>ganyu</i> labour? (Yes; No)
82	How many bicycles, motorcycles/scooters, cars, mini-buses, or lorries does the household own? (None; One; Two or more)
60	In the last seven days, did the male head/spouse do any casual, part-time work or <i>ganyu</i> labour, or did he do any work on household agricultural activities (including livestock- and fishing-related activities) or fishing, whether for sale or for household food? (Yes; No male head/spouse; No)

Figure 2 (cont.): Poverty indicators by uncertainty coefficient

<u>Uncertainty coefficient</u>	<u>Indicator (Responses ordered starting with those linked with higher poverty likelihoods)</u>
36	How many mortars/pestles (mtondo) does the household own? (None; One; Two)
34	In the last seven days, did the male head/spouse do any work on household agricultural activities (including livestock- and fishing-related activities), whether for sale or for household food? (No; Yes; No male head/spouse)
14	Does the household own any lanterns (paraffin)? (No; Yes)
12	Does the household now own (present at the farm or away) any goats? (No; Yes)
12	Does the household now own (present at your farm or away) any goats or pigs? (No; Yes)
12	Does the household own any <i>panga</i> ? (No; Yes)
11	Does the household now own (present at your farm or away) any cows, bulls/oxen, steers/heifer, calves, goats, or pigs? (No; Yes)
6	Does the household now own (present at your farm or away) any cows, bulls/oxen, steers/heifer, or calves? (No; Yes)
6	Does the household own any axes? (No; Yes)
3	Does the household own any sickles? (No; Yes)
2	Does the household now own (present at your farm or away) any pigs? (No; Yes)
2	What was your main source of drinking water over the past month? (Communal standpipe, open public well, protected public well, pond/lake, dam, river/stream, spring, rainwater, or other; Piped into dwelling, piped into yard/plot, borehole, open well in yard/plot, protected well in yard/plot, tanker truck/bowser, or bottled water)

Source: 2010/11 IHS questionnaire and 100% of the PBM-definition national poverty line

**Figures for
100% of the PBM-Definition National Poverty Line

(and Figures Pertaining
to Multiple Poverty Lines
across All Definitions of *Poverty*)**

**Figure 3 (100% of the PBM-definition 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	86.9
10-14	85.9
15-19	85.6
20-24	77.6
25-29	64.8
30-34	55.1
35-39	47.1
40-44	39.6
45-49	32.5
50-54	20.7
55-59	16.7
60-64	12.8
65-69	7.2
70-74	4.2
75-79	3.5
80-84	1.5
85-89	0.8
90-94	0.8
95-100	0.0

**Figure 4 (100% of the PBM-definition national line):
Derivation of estimated poverty likelihoods
associated with scores**

Score	Households in range and < poverty line		All households in range		Poverty likelihood (%)
0-4	128	÷	128	=	100.0
5-9	178	÷	205	=	86.9
10-14	831	÷	967	=	85.9
15-19	1,600	÷	1,869	=	85.6
20-24	3,137	÷	4,045	=	77.6
25-29	4,479	÷	6,913	=	64.8
30-34	4,664	÷	8,459	=	55.1
35-39	4,444	÷	9,445	=	47.1
40-44	4,101	÷	10,368	=	39.6
45-49	3,173	÷	9,767	=	32.5
50-54	1,938	÷	9,346	=	20.7
55-59	1,672	÷	10,027	=	16.7
60-64	1,121	÷	8,761	=	12.8
65-69	457	÷	6,362	=	7.2
70-74	207	÷	4,963	=	4.2
75-79	109	÷	3,116	=	3.5
80-84	30	÷	1,955	=	1.5
85-89	11	÷	1,315	=	0.8
90-94	7	÷	932	=	0.8
95-100	0	÷	1,055	=	0.0

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

Figure 5 (100% of the PBM-definition national line): For each score range, average differences between estimated and true poverty likelihoods for households, with confidence intervals, from 1,000 bootstraps of $n = 16,384$, 2010/11 scorecard applied to the 2010/11 validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	0.0	0.0	0.0	0.0
5-9	+26.2	15.3	18.2	24.2
10-14	-4.2	3.9	4.7	6.3
15-19	+4.6	4.0	4.6	6.1
20-24	+10.9	3.4	4.0	5.1
25-29	+2.8	2.7	3.1	4.1
30-34	-1.5	2.3	2.9	3.9
35-39	-6.6	4.4	4.7	5.2
40-44	+0.6	2.2	2.7	3.7
45-49	-3.5	3.2	3.5	4.4
50-54	+1.2	1.8	2.2	2.8
55-59	-1.6	1.7	2.1	2.7
60-64	-1.9	1.9	2.1	2.6
65-69	-5.1	3.6	3.9	4.2
70-74	-1.2	1.4	1.7	2.2
75-79	-2.1	2.1	2.3	2.9
80-84	+1.4	0.1	0.2	0.2
85-89	+0.8	0.0	0.0	0.0
90-94	+0.8	0.0	0.0	0.0
95-100	0.0	0.0	0.0	0.0

Figure 6 (100% of the PBM-definition national line):

Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	-0.7	67.2	74.1	90.9
4	-0.1	37.8	45.0	54.7
8	0.0	26.3	31.5	40.3
16	-0.3	19.7	22.6	29.7
32	-0.8	13.8	16.4	21.0
64	-0.8	9.7	11.4	14.7
128	-0.8	6.8	8.3	10.7
256	-0.9	4.9	5.7	7.7
512	-0.9	3.4	3.9	5.2
1,024	-0.9	2.6	2.9	3.6
2,048	-0.9	1.8	2.1	2.8
4,096	-1.0	1.2	1.5	1.8
8,192	-0.9	0.9	1.1	1.5
16,384	-1.0	0.6	0.7	1.0

Figure 7 (PBM-definition national poverty lines and the line that marks the poorest half of people below 100% of the PBM-definition national poverty line): Average differences between estimates and true values for poverty rates of a group of households at a point in time, precision, and the α factor for precision, 2010/11 scorecard applied to the 2010/11 validation sample

	Poverty line				
	Food	<u>National</u>		200%	Poorest half below 100% natl.
		100%	150%		
Estimate minus true value	-0.5	-1.0	+0.3	+0.3	+0.1
Precision of difference	0.5	0.6	0.6	0.6	0.5
α factor for precision	1.01	1.03	0.97	0.97	0.97

Results pertain to the 2010/11 scorecard applied to the 2010/11 validation sample.

Differences between estimates and true values are displayed in units of percentage points.

Precision is measured as 90-percent confidence intervals in units of \pm percentage points.

Differences and precision estimated from 1,000 bootstraps with $n = 16,384$.

α is estimated from 1,000 bootstrap samples of $n = 256, 512, 1,024, 2,048, 4,096, 8,192, \text{ and } 16,384$.

Figure 7 (PBM-definition international 2005 and 2011 PPP poverty lines): Average differences between estimates and true values for poverty rates of a group of households at a point in time, precision, and the α factor for precision, 2010/11 scorecard applied to the 2010/11 validation sample

	Poverty lines						
	Intl. 2005 PPP lines					Intl. 2011 PPP lines	
	\$1.25	\$2.00	\$2.50	\$5.00	\$8.44	\$1.90	\$3.10
Error (estimate minus observed value)	-0.1	+1.0	+0.8	0.0	0.0	+0.5	+0.8
Precision of difference	0.6	0.6	0.5	0.3	0.2	0.6	0.5
Alpha factor for precision	0.98	1.01	1.05	1.22	1.04	0.96	1.05

Results pertain to the 2014/15 scorecard applied to the 2014/15 validation sample.

Differences between estimates and observed values are displayed in units of percentage points.

Precision is measured as 90-percent confidence intervals in units of \pm percentage points.

Differences and precision estimated from 1,000 bootstraps with $n = 16,384$.

Alpha is estimated from 1,000 bootstrap samples of $n = 256, 512, 1,024, 2,048, 4,096, 8,192,$ and $16,384$.

Figure 7 (Government-definition national poverty lines and the line that marks the poorest half of people below 100% of the government-definition national poverty line): Average differences between estimates and true values for poverty rates of a group of households at a point in time, precision, and the α factor for precision, 2010/11 scorecard applied to the 2010/11 validation sample

	Poverty line				
	Food	<u>National</u>		200%	Poorest half below 100% natl.
		100%	150%		
Estimate minus true value	-0.3	+1.2	+0.8	+1.0	+0.2
Precision of difference	0.5	0.6	0.6	0.5	0.5
α factor for precision	0.94	0.96	0.94	0.93	0.94

Results pertain to the 2010/11 scorecard applied to the 2010/11 validation sample.

Differences between estimates and true values are displayed in units of percentage points.

Precision is measured as 90-percent confidence intervals in units of \pm percentage points.

Differences and precision estimated from 1,000 bootstraps with $n = 16,384$.

α is estimated from 1,000 bootstrap samples of $n = 256, 512, 1,024, 2,048, 4,096, 8,192, \text{ and } 16,384$.

Figure 7 (Government-definition international 2005 and 2011 PPP poverty lines): Average differences between estimates and true values for poverty rates of a group of households at a point in time, precision, and the α factor for precision, 2010/11 scorecard applied to the 2010/11 validation sample

	Poverty lines						
	Intl. 2005 PPP lines					Intl. 2011 PPP lines	
	\$1.25	\$2.00	\$2.50	\$5.00	\$8.44	\$1.90	\$3.10
Error (estimate minus observed value)	+0.5	+1.1	+0.9	+1.4	+0.6	+0.7	+0.9
Precision of difference	0.6	0.5	0.5	0.4	0.3	0.6	0.5
Alpha factor for precision	0.91	0.97	1.09	1.47	1.69	0.94	1.09

Results pertain to the 2014/15 scorecard applied to the 2014/15 validation sample.

Differences between estimates and observed values are displayed in units of percentage points.

Precision is measured as 90-percent confidence intervals in units of \pm percentage points.

Differences and precision estimated from 1,000 bootstraps with $n = 16,384$.

Alpha is estimated from 1,000 bootstrap samples of $n = 256, 512, 1,024, 2,048, 4,096, 8,192, \text{ and } 16,384$.

**Figure 7 (Old-definition international 2005 PPP poverty lines):
Average differences between estimates and true values for
poverty rates of a group of households at a point in time,
precision, and the α factor for precision, 2010/11 scorecard
applied to the 2010/11 validation sample**

	Intl. 2005 PPP poverty lines	
	\$1.25	\$2.00
Estimate minus true value	+0.5	+0.9
Precision of difference	0.6	0.5
α factor for precision	0.91	1.09

Results pertain to the 2010/11 scorecard applied to the 2010/11 validation sample.

Differences between estimates and true values are displayed in units of percentage points.

Precision is measured as 90-percent confidence intervals in units of \pm percentage points.

Differences and precision estimated from 1,000 bootstraps with $n = 16,384$.

α is estimated from 1,000 bootstrap samples of $n = 256, 512, 1,024, 2,048, 4,096, 8,192,$ and $16,384$.

Figure 8 (BMP-definition national poverty lines): Average differences between estimates and true values for changes in poverty rates between two points in time, precision, and the α factor for precision, 2010/11 scorecard applied to the 2010/11 validation sample (baseline) and all of the 2004/5 data (follow-up)

	Poverty line					
	National					
	Food	100%	150%	200%		
Estimated change minus true change	+6.0	0.0	-4.0	-4.3		
Precision of estimated change minus true change	0.6	0.9	0.8	0.8		
α factor for precision of estimated change	0.98	0.99	0.94	0.92		
2010/11 scorecard is applied to 2010/11 validation sample (baseline) and all 2004/5 data (follow-up).						
Differences between estimates of changes and true changes are displayed in units of percentage points.						
Precision is measured as 90-percent confidence intervals in units of \pm percentage points.						
Differences and precision estimated from 1,000 bootstraps with $n = 16,384$.						
α is estimated from 1,000 bootstrap samples of $n = 256, 512, 1,024, 2,048, 4,096, 8,192, \text{ and } 16,384$.						

**Figure 8 (BMP-definition international 2005 and 2011 PPP poverty lines):
Average differences between estimates and true values for changes in
poverty rates between two points in time, precision, and the α factor for
precision, 2010/11 scorecard applied to the 2010/11 validation sample
(baseline) and all of the 2004/5 data (follow-up)**

	Poverty lines						
	Intl. 2005 PPP lines					Intl. 2011 PPP lines	
	\$1.25	\$2.00	\$2.50	\$5.00	\$8.44	\$1.90	\$3.10
Error (estimate minus observed value)	-3.6	-4.2	-2.5	-0.4	-0.1	-3.9	-2.5
Precision of difference	0.9	0.7	0.6	0.4	0.2	0.8	0.6
Alpha factor for precision	0.96	0.89	0.91	1.01	0.93	0.91	0.91

New 2014 scorecard applied to the 2014 validation sample (baseline) and 2008 validation sample (follow-up).

Differences between estimates and observed values are displayed in units of percentage points.

Precision is measured as 90-percent confidence intervals in units of \pm percentage points.

Differences and precision estimated from 1,000 bootstraps with $n = 16,384$.

Alpha is estimated from 1,000 bootstrap samples of $n = 256, 512, 1,024, 2,048, 4,096, 8,192,$ and $16,384$.

Figure 8 (Government-definition national poverty lines): Average differences between estimates and true values for changes in poverty rates between two points in time, precision, and the α factor for precision, 2010/11 scorecard applied to the 2010/11 validation sample (baseline) and all of the 2004/5 data (follow-up)

	Poverty line			
	National			
	Food	100%	150%	200%
Estimated change minus true change	+7.6	+5.3	+1.4	+0.7
Precision of estimated change minus true change	0.7	0.8	0.8	0.7
α factor for precision of estimated change	0.94	0.94	0.89	0.87
2010/11 scorecard is applied to 2010/11 validation sample (baseline) and all 2004/5 data (follow-up).				
Differences between estimates of changes and true changes are displayed in units of percentage points.				
Precision is measured as 90-percent confidence intervals in units of \pm percentage points.				
Differences and precision estimated from 1,000 bootstraps with $n = 16,384$.				
α is estimated from 1,000 bootstrap samples of $n = 256, 512, 1,024, 2,048, 4,096, 8,192, \text{ and } 16,384$.				

Figure 8 (Government-definition international 2005 and 2011 PPP poverty lines):
Average differences between estimates and true values for changes in poverty rates between two points in time, precision, and the α factor for precision, 2010/11 scorecard applied to the 2010/11 validation sample (baseline) and all of the 2004/5 data (follow-up)

	Poverty lines						
	Intl. 2005 PPP lines					Intl. 2011 PPP lines	
	\$1.25	\$2.00	\$2.50	\$5.00	\$8.44	\$1.90	\$3.10
Error (estimate minus observed value)	+2.0	+0.7	-0.7	-1.3	-0.5	+1.0	-0.7
Precision of difference	0.8	0.6	0.6	0.4	0.3	0.7	0.5
Alpha factor for precision	0.87	0.88	0.92	1.13	1.31	0.86	0.90

New 2014 scorecard applied to the 2014 validation sample (baseline) and 2008 validation sample (follow-up).

Differences between estimates and observed values are displayed in units of percentage points.

Precision is measured as 90-percent confidence intervals in units of \pm percentage points.

Differences and precision estimated from 1,000 bootstraps with $n = 16,384$.

Alpha is estimated from 1,000 bootstrap samples of $n = 256, 512, 1,024, 2,048, 4,096, 8,192, \text{ and } 16,384$.

Figure 8 (Old-definition international 2005 PPP poverty lines): Average differences between estimates and true values for changes in poverty rates between two points in time, precision, and the α factor for precision, 2010/11 scorecard applied to the 2010/11 validation sample (baseline) and all of the 2004/5 data (follow-up)

	Poverty line	
	International 2005 PPP poverty lines	
	\$1.25	\$2.50
Estimated change minus true change	+2.1	-0.5
Precision of estimated change minus true change	0.8	0.5
α factor for precision of estimated change	0.88	0.90
2010/11 scorecard is applied to 2010/11 validation sample (baseline) and all 2004/5 data (follow-up).		
Differences between estimates of changes and true changes are displayed in units of percentage points.		
Precision is measured as 90-percent confidence intervals in units of \pm percentage points.		
Differences and precision estimated from 1,000 bootstraps with $n = 16,384$.		
α is estimated from 1,000 bootstrap samples of $n = 256, 512, 1,024, 2,048, 4,096, 8,192, \text{ and } 16,384$.		

Figure 9 (All poverty lines): Possible targeting outcomes

		Targeting segment	
		<u>Targeted</u>	<u>Non-targeted</u>
<u>True poverty status</u>	Below poverty line	<u>Inclusion</u> Below poverty line correctly targeted	<u>Undercoverage</u> Below poverty line mistakenly non-targeted
	<u>Above poverty line</u>	Leakage Above poverty line mistakenly targeted	Exclusion Above poverty line correctly non-targeted

Figure 10 (100% of the PBM-definition national line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	32.3	0.0	67.6	67.7	-99.2
≤9	0.3	32.1	0.1	67.6	67.8	-98.1
≤14	1.1	31.3	0.2	67.4	68.6	-92.5
≤19	2.6	29.8	0.6	67.1	69.7	-82.1
≤24	5.3	27.0	1.9	65.7	71.1	-61.2
≤29	9.7	22.7	4.5	63.1	72.8	-26.6
≤34	14.5	17.9	8.1	59.5	74.0	+14.4
≤39	19.2	13.1	12.8	54.8	74.1	+58.4
≤44	23.3	9.1	19.1	48.5	71.8	+41.0
≤49	26.2	6.2	25.9	41.7	67.9	+19.9
≤54	28.2	4.1	33.3	34.3	62.6	-2.8
≤59	30.0	2.4	41.5	26.1	56.1	-28.2
≤64	31.3	1.1	49.0	18.6	49.8	-51.4
≤69	31.9	0.5	54.8	12.8	44.7	-69.2
≤74	32.2	0.2	59.5	8.2	40.3	-83.6
≤79	32.4	0.0	62.4	5.2	37.6	-92.6
≤84	32.4	0.0	64.3	3.3	35.7	-98.6
≤89	32.4	0.0	65.6	2.0	34.4	-102.7
≤94	32.4	0.0	66.6	1.1	33.4	-105.6
≤100	32.4	0.0	67.6	0.0	32.4	-108.8

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

Figure 11 (100% of the PBM-definition national line): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	100.0	0.4	Only poor targeted
≤9	0.3	80.6	0.8	4.1:1
≤14	1.3	86.7	3.5	6.5:1
≤19	3.2	82.4	8.1	4.7:1
≤24	7.2	74.0	16.5	2.8:1
≤29	14.1	68.3	29.8	2.2:1
≤34	22.6	64.0	44.7	1.8:1
≤39	32.0	60.1	59.4	1.5:1
≤44	42.4	54.9	71.9	1.2:1
≤49	52.2	50.3	81.0	1.0:1
≤54	61.5	45.9	87.2	0.8:1
≤59	71.5	42.0	92.7	0.7:1
≤64	80.3	38.9	96.5	0.6:1
≤69	86.7	36.8	98.4	0.6:1
≤74	91.6	35.1	99.3	0.5:1
≤79	94.7	34.2	100.0	0.5:1
≤84	96.7	33.5	100.0	0.5:1
≤89	98.0	33.0	100.0	0.5:1
≤94	98.9	32.7	100.0	0.5:1
≤100	100.0	32.4	100.0	0.5:1

**Figures for
the PBM-Definition Food Poverty Line**

Figure 3 (PBM-definition food line): Estimated poverty likelihoods associated with scores

If a household's score is then the likelihood (%) of being below the poverty line is:
0-4	100.0
5-9	68.3
10-14	59.7
15-19	58.6
20-24	46.5
25-29	35.8
30-34	25.7
35-39	20.0
40-44	14.7
45-49	10.5
50-54	5.6
55-59	3.6
60-64	2.1
65-69	0.9
70-74	0.6
75-79	0.6
80-84	0.4
85-89	0.0
90-94	0.0
95-100	0.0

Figure 5 (PBM-definition food line): For each score range, average differences between estimated and true poverty likelihoods for households, with confidence intervals, from 1,000 bootstraps of $n = 16,384$, 2010/11 scorecard applied to the 2010/11 validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	+15.5	12.4	16.5	20.1
5-9	+11.1	15.0	17.8	24.8
10-14	-9.1	7.6	8.2	9.9
15-19	+1.0	5.0	5.8	7.5
20-24	-1.4	3.6	4.3	5.5
25-29	+4.8	2.4	2.9	3.7
30-34	-6.2	4.2	4.4	4.7
35-39	-0.4	1.9	2.3	2.9
40-44	-1.2	1.8	2.0	2.5
45-49	+3.7	1.0	1.1	1.5
50-54	-1.5	1.4	1.5	1.9
55-59	-0.8	0.9	1.1	1.5
60-64	-1.5	1.2	1.3	1.5
65-69	-2.1	1.6	1.7	1.9
70-74	-0.6	0.8	0.9	1.1
75-79	+0.6	0.0	0.0	0.0
80-84	+0.3	0.1	0.2	0.2
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 6 (PBM-definition food line): Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+0.3	57.9	68.0	78.1
4	-1.0	29.3	35.3	48.0
8	-1.1	20.4	24.7	33.0
16	-1.1	15.0	18.0	23.1
32	-0.9	10.2	12.3	16.1
64	-0.6	7.6	9.0	10.8
128	-0.5	5.1	5.9	7.8
256	-0.6	3.6	4.0	5.9
512	-0.6	2.5	3.1	3.8
1,024	-0.5	1.8	2.1	2.8
2,048	-0.6	1.3	1.5	2.0
4,096	-0.5	0.9	1.0	1.3
8,192	-0.5	0.6	0.7	1.0
16,384	-0.5	0.5	0.6	0.7

Figure 10 (PBM-definition food line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	14.0	0.0	85.8	85.9	-98.3
≤9	0.2	13.9	0.1	85.8	86.0	-96.0
≤14	0.9	13.3	0.4	85.4	86.3	-84.6
≤19	1.9	12.3	1.3	84.6	86.5	-64.2
≤24	3.7	10.5	3.5	82.3	86.0	-22.9
≤29	5.8	8.4	8.3	77.5	83.3	+40.8
≤34	8.3	5.8	14.2	71.6	80.0	-0.6
≤39	10.2	4.0	21.9	64.0	74.1	-54.5
≤44	11.8	2.4	30.6	55.2	67.0	-116.5
≤49	12.6	1.5	39.5	46.3	58.9	-179.5
≤54	13.2	0.9	48.3	37.6	50.8	-241.2
≤59	13.7	0.5	57.9	28.0	41.6	-309.1
≤64	13.9	0.2	66.4	19.5	33.4	-368.9
≤69	14.1	0.1	72.6	13.3	27.4	-412.8
≤74	14.1	0.0	77.5	8.4	22.5	-447.5
≤79	14.1	0.0	80.6	5.2	19.4	-469.5
≤84	14.2	0.0	82.5	3.3	17.5	-483.3
≤89	14.2	0.0	83.9	2.0	16.1	-492.6
≤94	14.2	0.0	84.8	1.1	15.2	-499.2
≤100	14.2	0.0	85.8	0.0	14.2	-506.6

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

Figure 11 (PBM-definition food line): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	83.0	0.8	4.9:1
≤9	0.3	70.6	1.7	2.4:1
≤14	1.3	67.5	6.2	2.1:1
≤19	3.2	59.7	13.4	1.5:1
≤24	7.2	51.3	26.1	1.1:1
≤29	14.1	41.0	41.0	0.7:1
≤34	22.6	37.0	59.0	0.6:1
≤39	32.0	31.7	71.8	0.5:1
≤44	42.4	27.7	83.1	0.4:1
≤49	52.2	24.2	89.2	0.3:1
≤54	61.5	21.5	93.5	0.3:1
≤59	71.5	19.1	96.5	0.2:1
≤64	80.3	17.4	98.5	0.2:1
≤69	86.7	16.3	99.6	0.2:1
≤74	91.6	15.4	99.9	0.2:1
≤79	94.7	14.9	99.9	0.2:1
≤84	96.7	14.6	100.0	0.2:1
≤89	98.0	14.4	100.0	0.2:1
≤94	98.9	14.3	100.0	0.2:1
≤100	100.0	14.2	100.0	0.2:1

**Figures for
150% of the PBM-Definition National Poverty Line**

**Figure 3 (150% of the PBM-definition national line):
 Estimated poverty likelihoods associated with scores**

If a household's score is then the likelihood (%) of being below the poverty line is:
0-4	100.0
5-9	100.0
10-14	97.1
15-19	94.8
20-24	91.3
25-29	84.2
30-34	80.0
35-39	77.0
40-44	68.1
45-49	60.1
50-54	43.8
55-59	38.1
60-64	34.5
65-69	27.3
70-74	15.1
75-79	11.7
80-84	7.1
85-89	4.2
90-94	2.0
95-100	0.0

Figure 5 (150% of the PBM-definition national line): For each score range, average differences between estimated and true poverty likelihoods for households, with confidence intervals, from 1,000 bootstraps of $n = 16,384$, 2010/11 scorecard applied to the 2010/11 validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	0.0	0.0	0.0	0.0
5-9	+11.0	9.0	11.5	13.9
10-14	-2.9	1.5	1.5	1.5
15-19	+3.9	2.9	3.5	4.6
20-24	+6.5	2.6	3.1	4.2
25-29	+0.7	2.1	2.4	3.1
30-34	+1.8	2.0	2.4	3.4
35-39	+2.2	2.1	2.5	3.1
40-44	+5.2	2.1	2.6	3.4
45-49	+0.9	2.4	2.9	3.9
50-54	-7.2	4.8	5.1	5.5
55-59	-6.4	4.3	4.5	5.0
60-64	+3.5	2.2	2.6	3.4
65-69	+3.2	2.5	2.9	3.9
70-74	-3.2	2.9	3.2	4.0
75-79	-8.1	6.0	6.5	7.4
80-84	+4.2	1.3	1.5	1.9
85-89	+2.7	1.0	1.2	1.6
90-94	+1.4	0.6	0.7	0.9
95-100	0.0	0.0	0.0	0.0

Figure 6 (150% of the PBM-definition national line):

Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	-0.3	69.4	76.4	88.0
4	+0.3	38.3	44.8	57.8
8	+0.8	27.6	32.2	42.5
16	+0.4	20.1	23.6	33.4
32	+0.2	15.1	18.3	23.8
64	+0.3	10.6	12.5	16.8
128	+0.3	7.5	8.7	11.2
256	+0.3	5.1	5.9	7.4
512	+0.3	3.7	4.5	5.7
1,024	+0.3	2.6	3.0	4.0
2,048	+0.3	1.7	2.1	2.7
4,096	+0.3	1.2	1.5	1.8
8,192	+0.3	0.9	1.1	1.4
16,384	+0.3	0.6	0.7	1.0

Figure 10 (150% of the PBM-definition national line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	52.9	0.0	47.0	47.1	-99.5
≤9	0.3	52.7	0.0	47.0	47.3	-98.8
≤14	1.3	51.7	0.0	47.0	48.2	-95.1
≤19	3.0	50.0	0.2	46.8	49.8	-88.4
≤24	6.5	46.6	0.8	46.2	52.7	-74.2
≤29	12.3	40.7	1.8	45.1	57.4	-50.2
≤34	19.0	34.0	3.6	43.4	62.5	-21.5
≤39	25.9	27.1	6.1	40.9	66.8	+9.3
≤44	32.7	20.4	9.7	37.2	69.9	+41.6
≤49	38.1	15.0	14.1	32.9	70.9	+70.2
≤54	42.9	10.2	18.6	28.3	71.2	+64.8
≤59	47.2	5.9	24.4	22.6	69.7	+54.0
≤64	49.9	3.1	30.4	16.6	66.4	+42.6
≤69	51.4	1.6	35.2	11.7	63.2	+33.5
≤74	52.3	0.7	39.3	7.7	60.0	+25.9
≤79	52.9	0.2	41.9	5.1	58.0	+21.0
≤84	53.0	0.1	43.7	3.2	56.2	+17.5
≤89	53.0	0.0	45.0	2.0	55.0	+15.1
≤94	53.0	0.0	45.9	1.1	54.1	+13.4
≤100	53.0	0.0	47.0	0.0	53.0	+11.4

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

Figure 11 (150% of the PBM-definition national line): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	100.0	0.2	Only poor targeted
≤9	0.3	93.8	0.6	15.1:1
≤14	1.3	98.4	2.4	61.9:1
≤19	3.2	94.3	5.6	16.5:1
≤24	7.2	89.5	12.2	8.6:1
≤29	14.1	87.0	23.2	6.7:1
≤34	22.6	84.3	35.9	5.4:1
≤39	32.0	81.0	48.9	4.3:1
≤44	42.4	77.0	61.6	3.4:1
≤49	52.2	72.9	71.8	2.7:1
≤54	61.5	69.7	80.8	2.3:1
≤59	71.5	65.9	88.9	1.9:1
≤64	80.3	62.1	94.1	1.6:1
≤69	86.7	59.3	97.0	1.5:1
≤74	91.6	57.1	98.7	1.3:1
≤79	94.7	55.8	99.7	1.3:1
≤84	96.7	54.8	99.9	1.2:1
≤89	98.0	54.1	100.0	1.2:1
≤94	98.9	53.6	100.0	1.2:1
≤100	100.0	53.0	100.0	1.1:1

**Figures for
200% of the PBM-Definition National Poverty Line**

**Figure 3 (200% of the PBM-definition national line):
 Estimated poverty likelihoods associated with scores**

If a household's score is then the likelihood (%) of being below the poverty line is:
0-4	100.0
5-9	100.0
10-14	98.5
15-19	98.3
20-24	94.6
25-29	90.5
30-34	90.5
35-39	89.5
40-44	83.3
45-49	78.4
50-54	64.4
55-59	58.2
60-64	53.4
65-69	45.3
70-74	34.4
75-79	23.3
80-84	18.9
85-89	13.7
90-94	10.7
95-100	0.0

Figure 5 (200% of the PBM-definition national line): For each score range, average differences between estimated and true poverty likelihoods for households, with confidence intervals, from 1,000 bootstraps of $n = 16,384$, 2010/11 scorecard applied to the 2010/11 validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	0.0	0.0	0.0	0.0
5-9	0.0	0.0	0.0	0.0
10-14	-1.5	0.7	0.7	0.7
15-19	+1.2	1.7	2.1	2.6
20-24	+2.5	2.0	2.5	3.3
25-29	-2.2	1.8	2.0	2.3
30-34	+1.6	1.5	1.8	2.4
35-39	+2.9	1.6	1.9	2.8
40-44	+5.4	2.0	2.4	3.2
45-49	+2.5	2.0	2.3	3.0
50-54	-9.9	6.0	6.1	6.6
55-59	-9.3	5.7	5.9	6.3
60-64	+2.1	2.3	2.9	3.8
65-69	+7.3	2.9	3.4	4.5
70-74	+6.1	3.2	3.7	4.5
75-79	-10.3	7.4	7.8	9.1
80-84	+11.2	2.1	2.6	3.5
85-89	+2.7	3.1	3.6	4.7
90-94	-0.9	4.7	5.6	7.3
95-100	-0.5	0.6	0.7	0.8

Figure 6 (200% of the PBM-definition national line):

Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	-1.0	66.6	77.5	87.1
4	+0.2	35.0	41.2	54.5
8	+0.3	26.7	31.3	39.9
16	+0.6	19.5	22.6	29.9
32	+0.4	13.7	16.2	21.2
64	+0.4	9.4	11.3	13.8
128	+0.4	6.4	7.7	9.9
256	+0.4	4.4	5.4	7.0
512	+0.3	3.4	3.9	5.1
1,024	+0.3	2.3	2.8	3.8
2,048	+0.3	1.7	1.9	2.6
4,096	+0.3	1.2	1.4	1.8
8,192	+0.3	0.8	1.0	1.2
16,384	+0.3	0.6	0.7	0.9

Figure 10 (200% of the PBM-definition national line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	67.8	0.0	32.1	32.2	-99.6
≤9	0.3	67.6	0.0	32.1	32.4	-99.0
≤14	1.3	66.6	0.0	32.1	33.4	-96.2
≤19	3.1	64.8	0.1	32.0	35.2	-90.7
≤24	6.9	61.0	0.3	31.8	38.6	-79.3
≤29	13.3	54.6	0.8	31.3	44.6	-59.6
≤34	20.9	47.0	1.7	30.4	51.3	-36.0
≤39	29.1	38.8	3.0	29.1	58.2	-10.0
≤44	37.3	30.6	5.1	27.0	64.3	+17.4
≤49	44.5	23.4	7.7	24.4	68.9	+42.4
≤54	51.3	16.6	10.2	21.9	73.1	+66.1
≤59	57.9	10.1	13.7	18.4	76.3	+79.8
≤64	62.3	5.6	18.0	14.1	76.4	+73.5
≤69	64.8	3.1	21.9	10.2	75.0	+67.8
≤74	66.3	1.6	25.3	6.7	73.0	+62.7
≤79	67.3	0.6	27.5	4.6	71.9	+59.5
≤84	67.5	0.4	29.2	2.9	70.5	+57.1
≤89	67.8	0.1	30.2	1.9	69.6	+55.5
≤94	67.9	0.0	31.1	1.0	68.9	+54.3
≤100	67.9	0.0	32.1	0.0	67.9	+52.7

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

Figure 11 (200% of the PBM-definition national line): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	100.0	0.2	Only poor targeted
≤9	0.3	100.0	0.5	Only poor targeted
≤14	1.3	100.0	1.9	Only poor targeted
≤19	3.2	98.3	4.6	58.5:1
≤24	7.2	95.3	10.1	20.2:1
≤29	14.1	94.3	19.6	16.5:1
≤34	22.6	92.5	30.8	12.4:1
≤39	32.0	90.8	42.8	9.9:1
≤44	42.4	88.0	54.9	7.3:1
≤49	52.2	85.3	65.5	5.8:1
≤54	61.5	83.4	75.5	5.0:1
≤59	71.5	80.9	85.2	4.2:1
≤64	80.3	77.6	91.8	3.5:1
≤69	86.7	74.8	95.4	3.0:1
≤74	91.6	72.3	97.6	2.6:1
≤79	94.7	71.0	99.1	2.4:1
≤84	96.7	69.8	99.5	2.3:1
≤89	98.0	69.1	99.8	2.2:1
≤94	98.9	68.6	100.0	2.2:1
≤100	100.0	67.9	100.0	2.1:1

**Figures for
the Line Marking the Poorest Half of People below
100% of the PBM-Definition National Poverty Line**

Figure 3 (Poorest half below 100% of the PBM-definition 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	68.3
10-14	60.9
15-19	59.9
20-24	50.0
25-29	38.6
30-34	26.8
35-39	21.1
40-44	17.1
45-49	13.6
50-54	6.5
55-59	5.3
60-64	2.5
65-69	1.0
70-74	0.7
75-79	0.7
80-84	0.7
85-89	0.7
90-94	0.7
95-100	0.0

Figure 5 (Poorest half below 100% of the PBM-definition national line): For each score range, average differences between estimated and true poverty likelihoods for households, with confidence intervals, from 1,000 bootstraps of $n = 16,384$, 2010/11 scorecard applied to the 2010/11 validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	+15.5	12.4	16.5	20.1
5-9	+11.1	15.0	17.8	24.8
10-14	-12.5	9.1	9.7	10.7
15-19	+1.9	5.0	5.8	7.6
20-24	+1.3	3.6	4.3	5.5
25-29	+6.2	2.4	2.9	3.7
30-34	-7.8	5.0	5.2	5.8
35-39	-1.1	2.0	2.3	3.1
40-44	+1.8	1.6	2.0	2.6
45-49	+5.8	1.1	1.3	1.7
50-54	-1.4	1.3	1.5	1.9
55-59	0.0	1.1	1.2	1.5
60-64	-1.1	1.0	1.1	1.4
65-69	-1.7	1.4	1.4	1.7
70-74	-0.7	0.8	0.9	1.2
75-79	+0.6	0.2	0.2	0.3
80-84	+0.6	0.1	0.2	0.2
85-89	+0.7	0.0	0.0	0.0
90-94	+0.7	0.0	0.0	0.0
95-100	0.0	0.0	0.0	0.0

Figure 6 (Poorest half below 100% of the PBM-definition national line): Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+0.5	58.7	68.2	78.0
4	-0.3	29.4	35.0	47.9
8	-0.4	20.8	25.0	31.6
16	-0.4	14.9	17.7	22.9
32	-0.2	10.1	12.4	16.8
64	0.0	7.6	9.0	11.1
128	+0.2	5.2	6.0	8.2
256	+0.1	3.5	4.4	5.7
512	+0.1	2.6	3.1	3.9
1,024	+0.1	1.9	2.2	3.0
2,048	+0.1	1.3	1.5	2.0
4,096	+0.1	0.8	1.0	1.3
8,192	+0.1	0.6	0.8	1.0
16,384	+0.1	0.5	0.5	0.7

**Figure 10 (Poorest half below 100% of the PBM-definition national line):
Percentages of households by cut-off score and targeting classification,
along with the hit rate and BPAC, 2010/11 scorecard applied to the
2010/11 validation sample**

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	15.0	0.0	84.8	84.9	-98.4
≤9	0.2	14.9	0.1	84.7	85.0	-96.2
≤14	0.9	14.2	0.4	84.4	85.4	-85.4
≤19	1.9	13.2	1.2	83.6	85.5	-66.2
≤24	3.8	11.3	3.4	81.4	85.2	-27.1
≤29	6.0	9.1	8.1	76.7	82.7	+33.1
≤34	8.7	6.4	13.9	71.0	79.7	+8.4
≤39	10.7	4.4	21.3	63.5	74.3	-41.0
≤44	12.4	2.7	30.0	54.8	67.2	-98.7
≤49	13.4	1.7	38.8	46.0	59.4	-156.7
≤54	14.1	1.1	47.4	37.4	51.5	-213.7
≤59	14.6	0.5	56.9	27.9	42.5	-276.7
≤64	14.9	0.2	65.4	19.5	34.3	-332.6
≤69	15.0	0.1	71.6	13.2	28.3	-373.8
≤74	15.1	0.0	76.5	8.4	23.4	-406.1
≤79	15.1	0.0	79.6	5.2	20.4	-426.6
≤84	15.1	0.0	81.5	3.3	18.4	-439.5
≤89	15.1	0.0	82.9	2.0	17.1	-448.2
≤94	15.1	0.0	83.8	1.1	16.2	-454.4
≤100	15.1	0.0	84.8	0.0	15.1	-461.4

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

Figure 11 (Poorest half below 100% of the PBM-definition national line): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	83.0	0.7	4.9:1
≤9	0.3	70.6	1.6	2.4:1
≤14	1.3	70.1	6.0	2.3:1
≤19	3.2	61.1	12.8	1.6:1
≤24	7.2	52.6	25.1	1.1:1
≤29	14.1	42.3	39.6	0.7:1
≤34	22.6	38.7	57.8	0.6:1
≤39	32.0	33.5	71.0	0.5:1
≤44	42.4	29.2	81.9	0.4:1
≤49	52.2	25.6	88.5	0.3:1
≤54	61.5	22.9	93.0	0.3:1
≤59	71.5	20.4	96.4	0.3:1
≤64	80.3	18.5	98.4	0.2:1
≤69	86.7	17.3	99.3	0.2:1
≤74	91.6	16.5	99.9	0.2:1
≤79	94.7	15.9	99.9	0.2:1
≤84	96.7	15.6	100.0	0.2:1
≤89	98.0	15.4	100.0	0.2:1
≤94	98.9	15.3	100.0	0.2:1
≤100	100.0	15.1	100.0	0.2:1

**Figures for
the PBM-Definition \$1.25/day 2005 PPP Poverty Line**

Figure 3 (PBM-definition \$1.25/day 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	97.4
10-14	95.9
15-19	94.6
20-24	90.5
25-29	83.4
30-34	77.6
35-39	73.8
40-44	65.5
45-49	58.0
50-54	41.6
55-59	35.2
60-64	30.9
65-69	24.4
70-74	13.3
75-79	10.3
80-84	6.4
85-89	2.4
90-94	1.3
95-100	0.0

Figure 5 (PBM-definition \$1.25/day line): For each score range, average differences between estimated and true poverty likelihoods for households, with confidence intervals, from 1,000 bootstraps of $n = 16,384$, 2010/11 scorecard applied to the 2010/11 validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	0.0	0.0	0.0	0.0
5-9	+8.4	9.0	11.5	13.9
10-14	-4.0	2.0	2.0	2.0
15-19	+4.2	3.0	3.6	4.5
20-24	+6.5	2.7	3.2	4.2
25-29	+0.5	2.1	2.5	3.1
30-34	-0.1	2.0	2.3	3.4
35-39	+1.9	2.1	2.5	3.2
40-44	+5.0	2.1	2.7	3.5
45-49	+0.6	2.4	3.0	3.9
50-54	-4.9	3.6	3.9	4.3
55-59	-5.8	4.0	4.2	4.8
60-64	+1.8	2.2	2.6	3.3
65-69	+1.4	2.4	2.9	3.9
70-74	-4.3	3.4	3.7	4.1
75-79	-8.5	6.3	6.7	7.6
80-84	+4.6	1.0	1.2	1.5
85-89	+0.9	1.0	1.2	1.6
90-94	+0.7	0.6	0.7	0.9
95-100	0.0	0.0	0.0	0.0

Figure 6 (PBM-definition \$1.25/day line): Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	-0.7	69.3	76.6	88.3
4	-0.1	38.7	45.3	57.9
8	+0.4	27.6	33.1	41.0
16	+0.2	19.9	23.7	32.1
32	-0.1	15.0	17.6	23.1
64	0.0	10.3	12.1	17.0
128	-0.1	7.5	8.8	11.5
256	-0.1	5.0	6.0	7.6
512	-0.1	3.8	4.6	5.9
1,024	0.0	2.6	3.1	4.1
2,048	-0.1	1.8	2.1	2.7
4,096	-0.1	1.3	1.5	1.8
8,192	-0.1	0.9	1.1	1.3
16,384	-0.1	0.6	0.7	0.9

Figure 10 (PBM-definition \$1.25/day line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	50.9	0.0	49.0	49.1	-99.5
≤9	0.3	50.7	0.0	49.0	49.3	-98.7
≤14	1.3	49.7	0.0	49.0	50.2	-94.9
≤19	3.0	48.0	0.2	48.8	51.8	-88.0
≤24	6.4	44.6	0.8	48.2	54.6	-73.3
≤29	12.2	38.8	2.0	47.0	59.2	-48.4
≤34	18.9	32.1	3.7	45.3	64.1	-18.7
≤39	25.5	25.5	6.5	42.5	68.0	+12.8
≤44	32.0	19.0	10.4	38.6	70.5	+45.8
≤49	37.1	13.9	15.0	34.0	71.1	+70.5
≤54	41.5	9.5	20.0	29.0	70.5	+60.8
≤59	45.5	5.5	26.0	23.0	68.5	+48.9
≤64	48.1	2.9	32.2	16.8	64.8	+36.8
≤69	49.5	1.5	37.2	11.8	61.3	+27.1
≤74	50.4	0.6	41.3	7.7	58.1	+19.1
≤79	50.9	0.1	43.9	5.1	56.0	+14.0
≤84	50.9	0.1	45.8	3.2	54.2	+10.3
≤89	51.0	0.0	47.0	2.0	53.0	+7.8
≤94	51.0	0.0	47.9	1.1	52.1	+6.0
≤100	51.0	0.0	49.0	0.0	51.0	+3.9

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

Figure 11 (PBM-definition \$1.25/day line): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	100.0	0.3	Only poor targeted
≤9	0.3	93.8	0.6	15.1:1
≤14	1.3	98.1	2.5	50.8:1
≤19	3.2	93.7	5.8	14.8:1
≤24	7.2	88.9	12.6	8.0:1
≤29	14.1	86.2	23.9	6.2:1
≤34	22.6	83.5	37.0	5.1:1
≤39	32.0	79.6	50.0	3.9:1
≤44	42.4	75.4	62.7	3.1:1
≤49	52.2	71.2	72.8	2.5:1
≤54	61.5	67.5	81.4	2.1:1
≤59	71.5	63.6	89.2	1.7:1
≤64	80.3	59.9	94.3	1.5:1
≤69	86.7	57.1	97.1	1.3:1
≤74	91.6	55.0	98.8	1.2:1
≤79	94.7	53.7	99.7	1.2:1
≤84	96.7	52.7	99.9	1.1:1
≤89	98.0	52.0	100.0	1.1:1
≤94	98.9	51.5	100.0	1.1:1
≤100	100.0	51.0	100.0	1.0:1

**Figures for
the PBM-Definition \$2.00/day 2005 PPP Poverty Line**

Figure 3 (PBM-definition \$2.00/day line): Estimated poverty likelihoods associated with scores

If a household's score is then the likelihood (%) of being below the poverty line is:
0-4	100.0
5-9	100.0
10-14	100.0
15-19	99.5
20-24	97.4
25-29	95.0
30-34	95.0
35-39	93.7
40-44	88.9
45-49	83.8
50-54	72.3
55-59	68.6
60-64	64.8
65-69	54.3
70-74	41.6
75-79	30.2
80-84	27.9
85-89	18.1
90-94	11.2
95-100	7.8

Figure 5 (PBM-definition \$2.00/day line): For each score range, average differences between estimated and true poverty likelihoods for households, with confidence intervals, from 1,000 bootstraps of $n = 16,384$, 2010/11 scorecard applied to the 2010/11 validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	0.0	0.0	0.0	0.0
5-9	0.0	0.0	0.0	0.0
10-14	0.0	0.0	0.0	0.0
15-19	+1.8	1.6	1.9	2.5
20-24	-0.1	1.0	1.1	1.5
25-29	-0.4	1.1	1.3	1.7
30-34	+2.3	1.2	1.5	2.0
35-39	+4.0	1.5	1.8	2.3
40-44	+3.2	1.7	2.0	2.7
45-49	+1.1	1.7	2.1	2.6
50-54	-6.9	4.4	4.6	4.9
55-59	-5.0	3.5	3.7	3.9
60-64	+7.2	2.4	2.8	3.9
65-69	+5.2	2.9	3.5	4.5
70-74	+6.6	3.3	4.0	5.0
75-79	-6.1	5.2	5.7	7.5
80-84	+13.1	3.2	4.0	5.3
85-89	-16.0	11.3	12.0	13.2
90-94	-13.7	10.2	10.8	12.2
95-100	+4.7	1.8	2.2	2.8

Figure 6 (PBM-definition \$2.00/day line): Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+0.7	59.0	79.4	90.2
4	+0.9	34.5	41.4	56.5
8	+1.2	23.9	29.3	38.0
16	+1.3	18.7	21.7	28.2
32	+1.1	12.8	15.3	20.5
64	+1.1	9.0	10.8	13.6
128	+1.1	6.5	7.5	9.9
256	+1.1	4.5	5.2	6.9
512	+1.1	3.2	3.8	5.0
1,024	+1.0	2.3	2.7	3.4
2,048	+1.1	1.7	1.9	2.5
4,096	+1.0	1.2	1.4	1.8
8,192	+1.0	0.8	1.0	1.3
16,384	+1.0	0.6	0.7	1.0

Figure 10 (PBM-definition \$2.00/day line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	73.8	0.0	26.0	26.2	-99.7
≤9	0.3	73.6	0.0	26.0	26.4	-99.1
≤14	1.3	72.7	0.0	26.0	27.3	-96.5
≤19	3.1	70.8	0.0	26.0	29.1	-91.5
≤24	7.1	66.9	0.2	25.9	32.9	-80.7
≤29	13.7	60.3	0.5	25.6	39.2	-62.4
≤34	21.6	52.4	1.0	25.0	46.6	-40.3
≤39	30.0	43.9	2.0	24.1	54.1	-16.1
≤44	39.0	34.9	3.4	22.7	61.7	+10.1
≤49	47.0	27.0	5.2	20.8	67.8	+34.0
≤54	54.3	19.7	7.2	18.8	73.1	+56.5
≤59	61.6	12.4	9.9	16.1	77.7	+80.0
≤64	66.7	7.3	13.6	12.4	79.1	+81.6
≤69	69.8	4.2	16.9	9.2	79.0	+77.2
≤74	71.7	2.3	19.9	6.1	77.8	+73.0
≤79	72.8	1.1	21.9	4.1	77.0	+70.4
≤84	73.3	0.7	23.4	2.6	75.8	+68.3
≤89	73.7	0.3	24.3	1.7	75.4	+67.1
≤94	73.9	0.1	25.0	1.0	74.9	+66.1
≤100	74.0	0.0	26.0	0.0	74.0	+64.8

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

Figure 11 (PBM-definition \$2.00/day line): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	100.0	0.2	Only poor targeted
≤9	0.3	100.0	0.5	Only poor targeted
≤14	1.3	100.0	1.8	Only poor targeted
≤19	3.2	98.8	4.2	82.6:1
≤24	7.2	97.8	9.5	44.7:1
≤29	14.1	96.7	18.5	29.0:1
≤34	22.6	95.4	29.1	20.8:1
≤39	32.0	93.8	40.6	15.1:1
≤44	42.4	92.0	52.8	11.5:1
≤49	52.2	90.0	63.5	9.0:1
≤54	61.5	88.2	73.4	7.5:1
≤59	71.5	86.1	83.3	6.2:1
≤64	80.3	83.0	90.2	4.9:1
≤69	86.7	80.5	94.4	4.1:1
≤74	91.6	78.2	96.9	3.6:1
≤79	94.7	76.9	98.5	3.3:1
≤84	96.7	75.8	99.0	3.1:1
≤89	98.0	75.2	99.6	3.0:1
≤94	98.9	74.7	99.9	3.0:1
≤100	100.0	74.0	100.0	2.8:1

**Figures for
the PBM-Definition \$2.50/day 2005 PPP Poverty Line**

Figure 3 (PBM-definition \$2.50/day 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	99.6
25-29	98.1
30-34	97.1
35-39	96.7
40-44	94.4
45-49	92.5
50-54	86.7
55-59	82.7
60-64	78.1
65-69	66.7
70-74	56.6
75-79	45.8
80-84	43.3
85-89	30.4
90-94	18.9
95-100	12.0

Figure 5 (PBM-definition \$2.50/day line): For each score range, average differences between estimated and true poverty likelihoods for households, with confidence intervals, from 1,000 bootstraps of $n = 16,384$, 2010/11 scorecard applied to the 2010/11 validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	0.0	0.0	0.0	0.0
5-9	0.0	0.0	0.0	0.0
10-14	0.0	0.0	0.0	0.0
15-19	+2.3	1.6	1.9	2.5
20-24	+1.3	0.9	1.1	1.3
25-29	+1.0	1.0	1.2	1.6
30-34	-0.2	0.8	1.0	1.2
35-39	+0.8	0.8	1.0	1.3
40-44	+2.2	1.3	1.5	2.1
45-49	+0.7	1.2	1.4	2.0
50-54	-0.2	1.6	1.9	2.5
55-59	+0.3	1.8	2.1	3.0
60-64	+9.5	2.3	2.7	3.7
65-69	-4.0	3.3	3.5	3.9
70-74	+5.9	3.8	4.4	5.8
75-79	-11.5	7.9	8.5	9.3
80-84	+8.1	5.3	6.3	8.0
85-89	-17.6	12.3	12.9	14.0
90-94	-31.8	19.6	20.4	22.1
95-100	+4.3	2.9	3.4	4.3

Figure 6 (PBM-definition \$2.50/day line): Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+0.2	60.8	70.4	88.9
4	+0.6	30.6	37.5	51.3
8	+1.0	21.9	26.7	34.6
16	+1.2	16.7	19.6	26.2
32	+1.0	11.7	14.0	18.2
64	+0.9	8.1	9.5	13.4
128	+0.9	5.5	6.6	8.2
256	+0.9	3.9	4.6	6.0
512	+0.9	2.8	3.5	4.5
1,024	+0.8	2.0	2.4	3.1
2,048	+0.8	1.5	1.8	2.5
4,096	+0.8	1.1	1.3	1.8
8,192	+0.8	0.7	0.9	1.1
16,384	+0.8	0.5	0.6	0.8

Figure 10 (PBM-definition \$2.50/day line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	82.8	0.0	17.0	17.2	-99.7
≤9	0.3	82.6	0.0	17.0	17.4	-99.2
≤14	1.3	81.7	0.0	17.0	18.3	-96.9
≤19	3.1	79.8	0.0	17.0	20.1	-92.4
≤24	7.1	75.9	0.1	16.9	24.0	-82.7
≤29	13.8	69.1	0.3	16.8	30.6	-66.3
≤34	22.1	60.9	0.5	16.6	38.6	-46.1
≤39	31.1	51.9	0.9	16.1	47.2	-23.9
≤44	40.7	42.2	1.7	15.4	56.1	+0.2
≤49	49.6	33.3	2.5	14.5	64.1	+22.7
≤54	57.8	25.2	3.7	13.3	71.1	+43.8
≤59	66.1	16.9	5.5	11.6	77.6	+65.9
≤64	72.2	10.8	8.1	8.9	81.1	+83.8
≤69	76.6	6.4	10.1	7.0	83.5	+87.8
≤74	79.3	3.7	12.3	4.7	84.0	+85.2
≤79	81.0	2.0	13.7	3.3	84.3	+83.4
≤84	81.8	1.1	14.9	2.2	84.0	+82.1
≤89	82.4	0.5	15.6	1.5	83.9	+81.2
≤94	82.8	0.1	16.1	0.9	83.8	+80.6
≤100	83.0	0.0	17.0	0.0	83.0	+79.5

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

Figure 11 (PBM-definition \$2.50/day line): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	100.0	0.2	Only poor targeted
≤9	0.3	100.0	0.4	Only poor targeted
≤14	1.3	100.0	1.6	Only poor targeted
≤19	3.2	98.8	3.8	82.6:1
≤24	7.2	98.4	8.6	62.5:1
≤29	14.1	98.0	16.7	48.1:1
≤34	22.6	97.8	26.6	45.1:1
≤39	32.0	97.1	37.5	33.4:1
≤44	42.4	96.1	49.1	24.4:1
≤49	52.2	95.1	59.8	19.5:1
≤54	61.5	93.9	69.6	15.4:1
≤59	71.5	92.4	79.6	12.1:1
≤64	80.3	89.9	87.0	8.9:1
≤69	86.7	88.4	92.3	7.6:1
≤74	91.6	86.6	95.6	6.4:1
≤79	94.7	85.5	97.6	5.9:1
≤84	96.7	84.6	98.7	5.5:1
≤89	98.0	84.1	99.4	5.3:1
≤94	98.9	83.7	99.8	5.1:1
≤100	100.0	83.0	100.0	4.9:1

**Figures for
the PBM-Definition \$5.00/day 2005 PPP Poverty Line**

Figure 3 (PBM-definition \$5.00/day 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	99.9
25-29	99.3
30-34	99.3
35-39	99.2
40-44	99.1
45-49	99.1
50-54	99.1
55-59	96.7
60-64	96.1
65-69	93.4
70-74	89.4
75-79	84.8
80-84	76.9
85-89	70.5
90-94	69.1
95-100	46.3

Figure 5 (PBM-definition \$5.00/day line): For each score range, average differences between estimated and true poverty likelihoods for households, with confidence intervals, from 1,000 bootstraps of $n = 16,384$, 2010/11 scorecard applied to the 2010/11 validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	0.0	0.0	0.0	0.0
5-9	0.0	0.0	0.0	0.0
10-14	0.0	0.0	0.0	0.0
15-19	0.0	0.0	0.0	0.0
20-24	+0.2	0.3	0.4	0.5
25-29	-0.4	0.3	0.3	0.3
30-34	-0.7	0.3	0.3	0.3
35-39	+0.7	0.5	0.7	0.8
40-44	-0.4	0.3	0.3	0.5
45-49	+0.2	0.4	0.5	0.6
50-54	0.0	0.4	0.5	0.6
55-59	-1.1	0.9	0.9	1.1
60-64	+4.4	1.4	1.5	2.1
65-69	-1.3	1.2	1.5	1.9
70-74	-2.8	2.1	2.2	2.4
75-79	+0.1	3.1	3.6	4.3
80-84	+9.4	7.2	8.2	10.5
85-89	-12.9	8.4	8.7	9.4
90-94	-8.3	7.1	7.5	9.7
95-100	-1.3	7.1	8.4	10.8

Figure 6 (PBM-definition \$5.00/day line): Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+0.4	7.6	59.5	63.7
4	-0.4	16.1	22.9	34.0
8	-0.2	10.8	14.3	22.3
16	0.0	8.9	11.4	15.2
32	-0.2	6.3	8.1	11.4
64	-0.2	4.7	5.8	7.8
128	-0.1	3.5	4.1	5.6
256	0.0	2.5	2.9	3.7
512	0.0	1.7	2.1	2.8
1,024	0.0	1.3	1.6	2.0
2,048	0.0	0.9	1.1	1.3
4,096	0.0	0.7	0.8	1.0
8,192	0.0	0.5	0.6	0.7
16,384	0.0	0.3	0.4	0.5

Figure 10 (PBM-definition \$5.00/day line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	95.6	0.0	4.3	4.4	-99.7
≤9	0.3	95.4	0.0	4.3	4.6	-99.3
≤14	1.3	94.4	0.0	4.3	5.6	-97.3
≤19	3.2	92.5	0.0	4.3	7.5	-93.4
≤24	7.2	88.5	0.0	4.3	11.5	-84.9
≤29	14.1	81.6	0.0	4.3	18.4	-70.5
≤34	22.6	73.1	0.0	4.3	26.8	-52.8
≤39	31.9	63.8	0.2	4.1	36.0	-33.2
≤44	42.2	53.5	0.2	4.1	46.3	-11.6
≤49	51.8	43.9	0.4	3.9	55.7	+8.7
≤54	61.0	34.6	0.5	3.8	64.9	+28.1
≤59	70.9	24.8	0.7	3.6	74.5	+48.8
≤64	78.9	16.8	1.4	2.9	81.9	+66.4
≤69	84.9	10.8	1.8	2.6	87.5	+79.3
≤74	89.3	6.4	2.3	2.0	91.3	+89.1
≤79	91.9	3.8	2.9	1.4	93.3	+95.0
≤84	93.5	2.2	3.2	1.1	94.6	+96.7
≤89	94.5	1.2	3.5	0.8	95.3	+96.3
≤94	95.2	0.5	3.7	0.6	95.8	+96.1
≤100	95.7	0.0	4.3	0.0	95.7	+95.5

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

Figure 11 (PBM-definition \$5.00/day line): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	100.0	0.1	Only poor targeted
≤9	0.3	100.0	0.3	Only poor targeted
≤14	1.3	100.0	1.4	Only poor targeted
≤19	3.2	100.0	3.3	Only poor targeted
≤24	7.2	99.7	7.5	382.6:1
≤29	14.1	99.7	14.7	374.6:1
≤34	22.6	99.8	23.6	599.4:1
≤39	32.0	99.5	33.3	182.1:1
≤44	42.4	99.5	44.1	185.2:1
≤49	52.2	99.3	54.1	142.5:1
≤54	61.5	99.2	63.8	131.5:1
≤59	71.5	99.1	74.1	105.3:1
≤64	80.3	98.3	82.5	57.7:1
≤69	86.7	98.0	88.7	48.4:1
≤74	91.6	97.5	93.3	38.6:1
≤79	94.7	97.0	96.0	32.0:1
≤84	96.7	96.7	97.7	29.2:1
≤89	98.0	96.4	98.7	26.8:1
≤94	98.9	96.2	99.5	25.4:1
≤100	100.0	95.7	100.0	22.2:1

**Figures for
the PBM-Definition \$8.44/day 2005 PPP Poverty Line**

Figure 3 (PBM-definition \$8.44/day 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	100.0
25-29	99.9
30-34	99.8
35-39	99.8
40-44	99.8
45-49	99.8
50-54	99.8
55-59	98.9
60-64	98.7
65-69	98.0
70-74	95.8
75-79	94.1
80-84	93.4
85-89	91.8
90-94	83.0
95-100	76.0

Figure 5 (PBM-definition \$8.44/day line): For each score range, average differences between estimated and true poverty likelihoods for households, with confidence intervals, from 1,000 bootstraps of $n = 16,384$, 2010/11 scorecard applied to the 2010/11 validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	0.0	0.0	0.0	0.0
5-9	0.0	0.0	0.0	0.0
10-14	0.0	0.0	0.0	0.0
15-19	0.0	0.0	0.0	0.0
20-24	0.0	0.0	0.0	0.0
25-29	+0.1	0.2	0.3	0.3
30-34	-0.2	0.1	0.1	0.1
35-39	+0.8	0.5	0.5	0.7
40-44	+0.3	0.3	0.3	0.5
45-49	+0.4	0.3	0.4	0.5
50-54	+0.3	0.3	0.4	0.5
55-59	-0.3	0.4	0.5	0.7
60-64	+1.5	0.9	1.0	1.4
65-69	-1.7	1.0	1.0	1.0
70-74	-1.6	1.2	1.2	1.3
75-79	+4.1	2.6	3.2	4.0
80-84	-4.9	2.9	3.0	3.1
85-89	-1.6	2.4	3.0	3.7
90-94	-14.4	7.9	8.1	8.4
95-100	+5.7	6.6	8.0	10.6

Figure 6 (PBM-definition \$8.44/day line): Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+0.2	2.9	4.1	58.2
4	+0.2	9.2	15.8	25.9
8	+0.2	7.6	10.0	14.9
16	+0.1	5.0	6.5	9.0
32	-0.1	3.5	4.1	6.0
64	-0.1	2.4	2.9	3.7
128	0.0	1.8	2.1	3.0
256	0.0	1.3	1.6	2.0
512	0.0	0.9	1.1	1.4
1,024	0.0	0.6	0.8	1.1
2,048	0.0	0.5	0.6	0.8
4,096	0.0	0.3	0.4	0.5
8,192	0.0	0.2	0.3	0.4
16,384	0.0	0.2	0.2	0.3

Figure 10 (PBM-definition \$8.44/day line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	98.2	0.0	1.6	1.8	-99.7
≤9	0.3	98.0	0.0	1.6	2.0	-99.3
≤14	1.3	97.1	0.0	1.6	2.9	-97.4
≤19	3.2	95.2	0.0	1.6	4.8	-93.6
≤24	7.2	91.1	0.0	1.6	8.9	-85.3
≤29	14.1	84.3	0.0	1.6	15.7	-71.3
≤34	22.6	75.8	0.0	1.6	24.2	-54.1
≤39	31.9	66.4	0.1	1.5	33.5	-35.0
≤44	42.3	56.1	0.2	1.5	43.7	-13.9
≤49	52.0	46.4	0.2	1.4	53.4	+5.9
≤54	61.3	37.1	0.3	1.4	62.6	+24.8
≤59	71.2	27.1	0.3	1.3	72.5	+45.1
≤64	79.8	18.6	0.5	1.1	80.9	+62.7
≤69	86.1	12.3	0.6	1.1	87.1	+75.6
≤74	90.9	7.5	0.8	0.9	91.7	+85.5
≤79	93.7	4.7	1.1	0.6	94.2	+91.5
≤84	95.6	2.8	1.1	0.5	96.1	+95.5
≤89	96.8	1.6	1.3	0.4	97.1	+98.0
≤94	97.6	0.7	1.3	0.3	98.0	+98.7
≤100	98.4	0.0	1.6	0.0	98.4	+98.3

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

Figure 11 (PBM-definition \$8.44/day line): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	100.0	0.1	Only poor targeted
≤9	0.3	100.0	0.3	Only poor targeted
≤14	1.3	100.0	1.3	Only poor targeted
≤19	3.2	100.0	3.2	Only poor targeted
≤24	7.2	100.0	7.3	Only poor targeted
≤29	14.1	99.9	14.3	750.2:1
≤34	22.6	99.9	22.9	1,199.9:1
≤39	32.0	99.7	32.5	326.2:1
≤44	42.4	99.6	43.0	280.3:1
≤49	52.2	99.6	52.8	249.4:1
≤54	61.5	99.6	62.3	243.2:1
≤59	71.5	99.6	72.4	221.8:1
≤64	80.3	99.3	81.1	147.3:1
≤69	86.7	99.3	87.5	149.2:1
≤74	91.6	99.2	92.4	118.2:1
≤79	94.7	98.9	95.2	86.7:1
≤84	96.7	98.8	97.2	84.6:1
≤89	98.0	98.7	98.4	76.9:1
≤94	98.9	98.7	99.3	74.9:1
≤100	100.0	98.4	100.0	60.2:1

**Figures for
the Government-Definition Food Poverty Line**

Figure 3 (Government-definition food line): Estimated poverty likelihoods associated with scores

If a household's score is then the likelihood (%) of being below the poverty line is:
0-4	100.0
5-9	81.6
10-14	71.8
15-19	70.3
20-24	55.7
25-29	49.9
30-34	36.4
35-39	27.6
40-44	21.5
45-49	16.5
50-54	9.4
55-59	5.6
60-64	4.0
65-69	2.2
70-74	1.0
75-79	0.6
80-84	0.4
85-89	0.0
90-94	0.0
95-100	0.0

Figure 5 (Government-definition food line): For each score range, average differences between estimated and true poverty likelihoods for households, with confidence intervals, from 1,000 bootstraps of $n = 16,384$, 2010/11 scorecard applied to the 2010/11 validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	+15.5	12.4	16.5	20.1
5-9	-9.5	8.0	8.4	11.6
10-14	-5.8	5.7	6.9	9.0
15-19	+7.2	4.8	6.0	8.1
20-24	-3.5	3.5	4.1	5.1
25-29	+6.1	2.6	3.2	4.2
30-34	-6.6	4.5	4.7	5.1
35-39	-3.1	2.6	2.8	3.4
40-44	-0.2	1.9	2.3	3.0
45-49	+6.3	1.3	1.5	2.1
50-54	-0.5	1.4	1.7	2.1
55-59	-2.0	1.6	1.7	2.0
60-64	-0.1	0.9	1.1	1.4
65-69	-3.6	2.6	2.8	3.2
70-74	+0.5	0.3	0.4	0.5
75-79	-0.6	0.7	0.8	1.1
80-84	+0.3	0.1	0.2	0.2
85-89	0.0	0.0	0.0	0.0
90-94	-0.4	0.5	0.5	0.7
95-100	0.0	0.0	0.0	0.0

Figure 6 (Government-definition food line): Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	-0.5	64.2	71.4	84.8
4	-0.1	31.8	39.1	50.4
8	-0.6	22.1	25.6	34.3
16	-0.7	15.7	19.0	25.7
32	-0.6	10.8	13.0	17.8
64	-0.3	8.2	9.7	12.1
128	-0.3	5.8	6.5	8.8
256	-0.4	3.7	4.5	5.9
512	-0.4	2.8	3.3	4.2
1,024	-0.3	1.9	2.2	3.0
2,048	-0.3	1.3	1.6	2.1
4,096	-0.3	1.0	1.1	1.5
8,192	-0.3	0.7	0.8	1.1
16,384	-0.3	0.5	0.6	0.8

Figure 10 (Government-definition food line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	19.8	0.0	80.0	80.1	-98.8
≤9	0.3	19.7	0.0	80.0	80.3	-96.9
≤14	1.0	18.9	0.3	79.8	80.8	-88.4
≤19	2.2	17.8	1.0	79.1	81.3	-73.1
≤24	4.6	15.4	2.6	77.4	82.0	-40.9
≤29	7.7	12.2	6.4	73.6	81.4	+9.5
≤34	11.3	8.7	11.3	68.7	80.0	+43.3
≤39	14.1	5.9	18.0	62.1	76.1	+9.9
≤44	16.3	3.6	26.1	54.0	70.3	-30.6
≤49	17.6	2.4	34.6	45.5	63.0	-73.4
≤54	18.5	1.4	43.0	37.0	55.6	-115.6
≤59	19.2	0.7	52.3	27.7	47.0	-162.2
≤64	19.6	0.4	60.7	19.3	38.9	-204.4
≤69	19.8	0.1	66.8	13.2	33.0	-235.0
≤74	19.9	0.1	71.8	8.3	28.2	-259.7
≤79	19.9	0.0	74.8	5.2	25.2	-275.0
≤84	19.9	0.0	76.8	3.3	23.2	-284.8
≤89	19.9	0.0	78.1	2.0	21.9	-291.4
≤94	19.9	0.0	79.0	1.1	21.0	-296.0
≤100	19.9	0.0	80.1	0.0	19.9	-301.3

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

Figure 11 (Government-definition food line): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	83.0	0.5	4.9:1
≤9	0.3	85.7	1.4	6.0:1
≤14	1.3	78.4	5.1	3.6:1
≤19	3.2	69.2	11.0	2.3:1
≤24	7.2	63.5	23.0	1.7:1
≤29	14.1	54.6	38.7	1.2:1
≤34	22.6	49.9	56.5	1.0:1
≤39	32.0	43.9	70.5	0.8:1
≤44	42.4	38.5	81.9	0.6:1
≤49	52.2	33.7	88.1	0.5:1
≤54	61.5	30.1	92.8	0.4:1
≤59	71.5	26.9	96.4	0.4:1
≤64	80.3	24.4	98.2	0.3:1
≤69	86.7	22.9	99.4	0.3:1
≤74	91.6	21.7	99.6	0.3:1
≤79	94.7	21.0	99.9	0.3:1
≤84	96.7	20.6	100.0	0.3:1
≤89	98.0	20.3	100.0	0.3:1
≤94	98.9	20.2	100.0	0.3:1
≤100	100.0	19.9	100.0	0.2:1

**Figures for
100% of the Government-Definition National Poverty Line**

**Figure 3 (100% of the Government-definition national):
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	97.4
10-14	95.7
15-19	94.4
20-24	87.4
25-29	80.0
30-34	72.0
35-39	70.2
40-44	57.4
45-49	47.9
50-54	30.5
55-59	24.9
60-64	20.0
65-69	12.4
70-74	6.5
75-79	5.3
80-84	2.7
85-89	1.1
90-94	1.1
95-100	0.0

Figure 5 (100% of the Government-definition national):
For each score range, average differences between
estimated and true poverty likelihoods for
households, with confidence intervals, from 1,000
bootstraps of $n = 16,384$, 2010/11 scorecard applied
to the 2010/11 validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	0.0	0.0	0.0	0.0
5-9	-2.6	1.3	1.3	1.3
10-14	-2.9	2.1	2.1	2.1
15-19	+5.6	3.2	3.8	4.8
20-24	+2.9	2.5	3.0	3.8
25-29	+1.9	2.3	2.7	3.6
30-34	-2.9	2.4	2.6	3.2
35-39	+5.5	2.2	2.7	3.5
40-44	+5.9	2.2	2.6	3.6
45-49	+8.2	2.5	3.1	4.1
50-54	-5.1	3.7	3.9	4.3
55-59	-4.9	3.5	3.7	4.1
60-64	0.0	1.9	2.3	3.0
65-69	+0.9	1.9	2.3	3.1
70-74	-0.4	1.6	1.8	2.3
75-79	-1.2	2.1	2.4	3.0
80-84	+2.6	0.1	0.2	0.2
85-89	+0.3	0.7	0.8	1.1
90-94	+0.7	0.5	0.5	0.7
95-100	0.0	0.0	0.0	0.0

**Figure 6 (100% of the Government-definition national):
Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample**

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+0.7	69.9	78.8	91.1
4	+1.1	37.7	43.7	56.1
8	+1.1	27.2	31.5	38.9
16	+1.3	19.5	23.9	31.4
32	+1.1	14.3	17.2	23.1
64	+1.2	10.3	12.1	16.3
128	+1.3	6.9	8.2	11.4
256	+1.2	4.8	5.6	7.6
512	+1.2	3.6	4.2	5.1
1,024	+1.2	2.5	2.9	4.0
2,048	+1.2	1.7	1.9	2.6
4,096	+1.2	1.2	1.4	1.8
8,192	+1.2	0.8	1.0	1.3
16,384	+1.2	0.6	0.7	0.9

Figure 10 (100% of the Government-definition national): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	43.3	0.0	56.6	56.7	-99.4
≤9	0.3	43.1	0.0	56.6	56.9	-98.5
≤14	1.3	42.1	0.0	56.6	57.9	-94.0
≤19	2.9	40.4	0.2	56.4	59.3	-85.9
≤24	6.4	37.0	0.8	55.8	62.2	-68.6
≤29	11.9	31.5	2.2	54.4	66.3	-40.0
≤34	18.2	25.2	4.4	52.2	70.4	-6.0
≤39	24.4	19.0	7.6	49.0	73.4	+30.0
≤44	29.9	13.5	12.5	44.1	74.1	+66.7
≤49	34.0	9.4	18.2	38.4	72.4	+58.0
≤54	37.3	6.0	24.2	32.4	69.8	+44.3
≤59	40.3	3.1	31.3	25.3	65.6	+27.9
≤64	42.1	1.3	38.2	18.4	60.5	+11.9
≤69	42.7	0.7	43.9	12.7	55.4	-1.3
≤74	43.1	0.3	48.5	8.1	51.2	-11.8
≤79	43.3	0.0	51.4	5.2	48.6	-18.5
≤84	43.4	0.0	53.3	3.3	46.6	-22.9
≤89	43.4	0.0	54.6	2.0	45.4	-25.9
≤94	43.4	0.0	55.6	1.1	44.4	-28.1
≤100	43.4	0.0	56.6	0.0	43.4	-30.5

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

Figure 11 (100% of the Government-definition national): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	100.0	0.3	Only poor targeted
≤9	0.3	100.0	0.8	Only poor targeted
≤14	1.3	98.6	3.0	71.8:1
≤19	3.2	92.9	6.8	13.1:1
≤24	7.2	88.6	14.7	7.7:1
≤29	14.1	84.2	27.4	5.3:1
≤34	22.6	80.6	41.9	4.1:1
≤39	32.0	76.1	56.2	3.2:1
≤44	42.4	70.6	69.0	2.4:1
≤49	52.2	65.1	78.3	1.9:1
≤54	61.5	60.7	86.1	1.5:1
≤59	71.5	56.3	92.8	1.3:1
≤64	80.3	52.4	97.0	1.1:1
≤69	86.7	49.3	98.5	1.0:1
≤74	91.6	47.0	99.4	0.9:1
≤79	94.7	45.7	99.9	0.8:1
≤84	96.7	44.8	99.9	0.8:1
≤89	98.0	44.3	100.0	0.8:1
≤94	98.9	43.8	100.0	0.8:1
≤100	100.0	43.4	100.0	0.8:1

**Figures for
150% of the Government-Definition National Poverty Line**

**Figure 3 (150% of the Government-definition national):
 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	98.7
15-19	98.5
20-24	96.3
25-29	94.7
30-34	94.7
35-39	90.1
40-44	82.7
45-49	76.8
50-54	57.1
55-59	48.9
60-64	44.1
65-69	34.1
70-74	23.3
75-79	17.7
80-84	11.5
85-89	3.7
90-94	2.4
95-100	0.0

Figure 5 (150% of the Government-definition national):
For each score range, average differences between
estimated and true poverty likelihoods for
households, with confidence intervals, from 1,000
bootstraps of $n = 16,384$, 2010/11 scorecard applied
to the 2010/11 validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	0.0	0.0	0.0	0.0
5-9	0.0	0.0	0.0	0.0
10-14	-1.3	0.7	0.7	0.7
15-19	+0.7	1.5	1.9	2.5
20-24	+0.8	1.5	1.8	2.3
25-29	+0.3	1.3	1.5	2.1
30-34	+4.0	1.4	1.6	2.0
35-39	+3.5	1.7	2.0	2.8
40-44	+3.4	1.9	2.2	2.9
45-49	+2.5	2.0	2.4	3.3
50-54	-7.0	4.7	5.0	5.4
55-59	-9.8	6.0	6.2	6.5
60-64	+5.1	2.3	2.8	4.0
65-69	+6.3	2.8	3.2	4.2
70-74	+3.0	2.6	3.0	4.0
75-79	-0.1	3.3	4.0	5.2
80-84	+5.9	1.8	2.3	3.2
85-89	-1.5	2.3	2.8	3.7
90-94	-5.6	4.9	5.4	6.1
95-100	0.0	0.0	0.0	0.0

Figure 6 (150% of the Government-definition national):
Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	-1.4	66.3	74.3	91.6
4	+0.7	35.9	41.6	53.7
8	+0.6	26.0	31.4	39.6
16	+0.7	18.9	22.7	27.7
32	+0.9	13.3	16.0	20.7
64	+0.9	9.5	11.2	14.7
128	+0.9	6.7	7.9	10.5
256	+0.9	4.5	5.3	6.9
512	+0.8	3.5	4.0	5.0
1,024	+0.8	2.4	2.9	3.4
2,048	+0.8	1.6	2.0	2.6
4,096	+0.8	1.1	1.3	1.7
8,192	+0.8	0.8	1.0	1.2
16,384	+0.8	0.6	0.7	0.9

Figure 10 (150% of the Government-definition national): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	64.1	0.0	35.7	35.9	-99.6
≤9	0.3	63.9	0.0	35.7	36.1	-99.0
≤14	1.3	63.0	0.0	35.7	37.0	-96.0
≤19	3.1	61.1	0.0	35.7	38.8	-90.2
≤24	7.0	57.2	0.2	35.5	42.6	-77.8
≤29	13.6	50.7	0.6	35.2	48.7	-56.9
≤34	21.3	43.0	1.3	34.4	55.7	-31.7
≤39	29.6	34.7	2.5	33.3	62.8	-4.2
≤44	37.8	26.4	4.6	31.2	69.0	+24.9
≤49	45.0	19.3	7.2	28.5	73.5	+51.1
≤54	51.1	13.2	10.5	25.3	76.3	+75.2
≤59	56.9	7.3	14.6	21.1	78.0	+77.2
≤64	60.5	3.8	19.8	15.9	76.4	+69.2
≤69	62.2	2.0	24.4	11.3	73.5	+62.0
≤74	63.3	0.9	28.3	7.4	70.8	+56.0
≤79	63.9	0.4	30.8	4.9	68.8	+52.0
≤84	64.1	0.2	32.6	3.1	67.2	+49.3
≤89	64.2	0.1	33.8	1.9	66.1	+47.3
≤94	64.3	0.0	34.7	1.1	65.3	+46.0
≤100	64.3	0.0	35.7	0.0	64.3	+44.4

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

Figure 11 (150% of the Government-definition national): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	100.0	0.2	Only poor targeted
≤9	0.3	100.0	0.5	Only poor targeted
≤14	1.3	100.0	2.0	Only poor targeted
≤19	3.2	99.1	4.9	105.3:1
≤24	7.2	97.3	10.9	35.8:1
≤29	14.1	96.0	21.1	24.2:1
≤34	22.6	94.2	33.1	16.3:1
≤39	32.0	92.3	46.0	12.0:1
≤44	42.4	89.3	58.9	8.3:1
≤49	52.2	86.2	70.0	6.2:1
≤54	61.5	83.0	79.4	4.9:1
≤59	71.5	79.6	88.6	3.9:1
≤64	80.3	75.3	94.1	3.1:1
≤69	86.7	71.8	96.8	2.5:1
≤74	91.6	69.1	98.5	2.2:1
≤79	94.7	67.5	99.5	2.1:1
≤84	96.7	66.3	99.7	2.0:1
≤89	98.0	65.5	99.9	1.9:1
≤94	98.9	64.9	100.0	1.9:1
≤100	100.0	64.3	100.0	1.8:1

**Figures for
200% of the Government-Definition National Poverty Line**

Figure 3 (200% of the Government-definition national line): Estimated poverty likelihoods associated with scores

If a household's score is then the likelihood (%) of being below the poverty line is:
0-4	100.0
5-9	100.0
10-14	100.0
15-19	100.0
20-24	98.6
25-29	98.1
30-34	97.6
35-39	96.4
40-44	93.2
45-49	90.3
50-54	80.5
55-59	74.4
60-64	67.1
65-69	53.0
70-74	38.3
75-79	31.2
80-84	24.5
85-89	13.3
90-94	8.8
95-100	0.6

Figure 5 (200% of the Government-definition national line): For each score range, average differences between estimated and true poverty likelihoods for households, with confidence intervals, from 1,000 bootstraps of $n = 16,384$, 2010/11 scorecard applied to the 2010/11 validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	0.0	0.0	0.0	0.0
5-9	0.0	0.0	0.0	0.0
10-14	0.0	0.0	0.0	0.0
15-19	+2.2	1.5	1.9	2.5
20-24	-0.8	0.6	0.6	0.7
25-29	+0.2	0.8	1.0	1.3
30-34	+2.7	1.0	1.2	1.7
35-39	+0.8	0.8	1.0	1.3
40-44	+1.3	1.3	1.5	2.0
45-49	+0.7	1.3	1.6	1.9
50-54	-2.0	2.0	2.2	2.7
55-59	-1.8	2.0	2.3	3.0
60-64	+6.2	2.3	2.7	3.7
65-69	+1.7	3.0	3.6	4.9
70-74	-0.1	3.4	4.1	5.2
75-79	-2.1	4.5	5.5	6.7
80-84	+12.9	2.8	3.3	4.5
85-89	+2.7	3.3	3.7	5.3
90-94	-7.7	6.7	7.4	8.8
95-100	-1.9	2.0	2.3	3.0

Figure 6 (200% of the Government-definition national line): Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	-0.2	60.7	76.0	89.1
4	+0.9	30.4	36.3	54.5
8	+0.9	20.9	25.4	33.6
16	+1.1	16.0	18.9	24.7
32	+1.3	11.3	13.5	18.7
64	+1.2	8.2	9.8	13.6
128	+1.1	5.9	6.8	8.5
256	+1.1	3.9	4.6	6.0
512	+1.1	3.0	3.5	4.4
1,024	+1.1	2.0	2.4	3.2
2,048	+1.0	1.5	1.8	2.3
4,096	+1.0	1.0	1.2	1.5
8,192	+1.0	0.7	0.9	1.1
16,384	+1.0	0.5	0.6	0.9

Figure 10 (200% of the Government-definition national line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	76.8	0.0	23.0	23.2	-99.7
≤9	0.3	76.6	0.0	23.0	23.4	-99.1
≤14	1.3	75.6	0.0	23.0	24.4	-96.6
≤19	3.1	73.8	0.0	23.0	26.2	-91.8
≤24	7.2	69.8	0.1	23.0	30.1	-81.3
≤29	13.9	63.0	0.2	22.8	36.8	-63.5
≤34	22.0	54.9	0.6	22.5	44.5	-42.1
≤39	31.0	46.0	1.0	22.0	53.0	-18.1
≤44	40.6	36.4	1.8	21.2	61.8	+7.8
≤49	49.2	27.8	3.0	20.1	69.3	+31.7
≤54	56.9	20.0	4.6	18.4	75.3	+53.9
≤59	64.6	12.4	7.0	16.1	80.6	+76.9
≤64	70.0	7.0	10.4	12.7	82.6	+86.5
≤69	73.2	3.7	13.5	9.6	82.8	+82.5
≤74	75.2	1.8	16.5	6.6	81.8	+78.6
≤79	76.2	0.8	18.5	4.5	80.7	+75.9
≤84	76.6	0.4	20.1	2.9	79.5	+73.8
≤89	76.8	0.2	21.3	1.8	78.6	+72.4
≤94	76.9	0.0	22.0	1.0	78.0	+71.4
≤100	77.0	0.0	23.0	0.0	77.0	+70.0

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

Figure 11 (200% of the Government-definition national): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	100.0	0.2	Only poor targeted
≤9	0.3	100.0	0.4	Only poor targeted
≤14	1.3	100.0	1.7	Only poor targeted
≤19	3.2	99.1	4.1	105.3:1
≤24	7.2	99.2	9.3	118.0:1
≤29	14.1	98.6	18.1	69.7:1
≤34	22.6	97.4	28.6	37.7:1
≤39	32.0	96.8	40.3	29.9:1
≤44	42.4	95.7	52.7	22.0:1
≤49	52.2	94.3	63.9	16.5:1
≤54	61.5	92.5	74.0	12.4:1
≤59	71.5	90.3	83.9	9.3:1
≤64	80.3	87.1	90.9	6.8:1
≤69	86.7	84.5	95.1	5.4:1
≤74	91.6	82.0	97.7	4.6:1
≤79	94.7	80.4	99.0	4.1:1
≤84	96.7	79.2	99.5	3.8:1
≤89	98.0	78.3	99.8	3.6:1
≤94	98.9	77.7	100.0	3.5:1
≤100	100.0	77.0	100.0	3.3:1

**Figures for
the Line Marking the Poorest Half of People below
100% of the Government-Definition National Poverty Line**

Figure 3 (Poorest half below 100% of the Government-definition 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	81.6
10-14	73.3
15-19	72.3
20-24	60.3
25-29	51.3
30-34	37.4
35-39	29.0
40-44	23.6
45-49	18.0
50-54	9.1
55-59	6.1
60-64	3.7
65-69	2.2
70-74	0.7
75-79	0.6
80-84	0.4
85-89	0.0
90-94	0.0
95-100	0.0

Figure 5 (Poorest half below 100% of the Government-definition national line): For each score range, average differences between estimated and true poverty likelihoods for households, with confidence intervals, from 1,000 bootstraps of $n = 16,384$, 2010/11 scorecard applied to the 2010/11 validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	+15.5	12.4	16.5	20.1
5-9	-9.5	8.0	8.4	11.6
10-14	+1.0	6.4	7.6	10.0
15-19	+12.9	4.8	5.9	8.1
20-24	-1.3	3.3	4.0	5.0
25-29	+3.2	2.8	3.2	4.2
30-34	-6.1	4.2	4.5	4.9
35-39	-2.1	2.2	2.6	3.3
40-44	+1.8	1.9	2.4	3.1
45-49	+7.4	1.3	1.5	2.2
50-54	-1.2	1.4	1.7	2.2
55-59	-1.6	1.4	1.5	2.0
60-64	-0.1	0.9	1.1	1.3
65-69	-4.0	2.8	3.0	3.4
70-74	+0.4	0.3	0.3	0.4
75-79	-0.4	0.6	0.7	1.0
80-84	+0.3	0.1	0.2	0.2
85-89	0.0	0.0	0.0	0.0
90-94	-0.4	0.5	0.5	0.7
95-100	0.0	0.0	0.0	0.0

Figure 6 (Poorest half below 100% of the Government-definition national line): Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+1.0	63.9	71.2	85.0
4	-0.1	32.9	39.2	51.4
8	-0.1	22.0	26.2	37.2
16	-0.1	16.2	18.9	28.0
32	-0.1	11.0	13.4	17.8
64	+0.1	8.0	9.7	12.6
128	+0.2	5.5	6.8	9.1
256	+0.1	3.8	4.5	6.3
512	+0.1	2.7	3.3	4.5
1,024	+0.1	2.0	2.3	3.1
2,048	+0.2	1.4	1.6	2.1
4,096	+0.2	1.0	1.1	1.6
8,192	+0.2	0.7	0.8	1.1
16,384	+0.2	0.5	0.6	0.8

Figure 10 (Poorest half below 100% of the Government-definition national line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	20.2	0.0	79.6	79.7	-98.8
≤9	0.3	20.0	0.0	79.6	79.8	-97.0
≤14	1.0	19.3	0.3	79.3	80.3	-88.7
≤19	2.1	18.2	1.0	78.6	80.7	-74.2
≤24	4.6	15.8	2.6	77.0	81.6	-42.2
≤29	8.0	12.3	6.1	73.5	81.5	+8.7
≤34	11.5	8.8	11.0	68.6	80.1	+45.7
≤39	14.3	6.0	17.7	62.0	76.3	+13.2
≤44	16.6	3.7	25.7	53.9	70.5	-26.6
≤49	17.9	2.4	34.2	45.4	63.3	-68.3
≤54	18.9	1.5	42.6	37.0	55.9	-109.5
≤59	19.6	0.7	51.9	27.7	47.3	-155.2
≤64	19.9	0.4	60.3	19.3	39.3	-196.6
≤69	20.2	0.1	66.4	13.2	33.4	-226.5
≤74	20.3	0.1	71.3	8.3	28.6	-250.7
≤79	20.3	0.0	74.4	5.2	25.6	-265.8
≤84	20.3	0.0	76.3	3.3	23.6	-275.4
≤89	20.3	0.0	77.6	2.0	22.3	-281.8
≤94	20.3	0.0	78.6	1.1	21.4	-286.4
≤100	20.3	0.0	79.6	0.0	20.3	-291.6

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

Figure 11 (Poorest half below 100% of the Government-definition national line): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	83.0	0.5	4.9:1
≤9	0.3	85.7	1.4	6.0:1
≤14	1.3	76.0	4.9	3.2:1
≤19	3.2	66.7	10.4	2.0:1
≤24	7.2	63.4	22.5	1.7:1
≤29	14.1	56.7	39.4	1.3:1
≤34	22.6	51.0	56.6	1.0:1
≤39	32.0	44.7	70.4	0.8:1
≤44	42.4	39.1	81.6	0.6:1
≤49	52.2	34.3	88.0	0.5:1
≤54	61.5	30.6	92.7	0.4:1
≤59	71.5	27.4	96.4	0.4:1
≤64	80.3	24.8	98.1	0.3:1
≤69	86.7	23.3	99.5	0.3:1
≤74	91.6	22.1	99.6	0.3:1
≤79	94.7	21.4	99.9	0.3:1
≤84	96.7	21.0	100.0	0.3:1
≤89	98.0	20.7	100.0	0.3:1
≤94	98.9	20.5	100.0	0.3:1
≤100	100.0	20.3	100.0	0.3:1

**Figures for
the Government-Definition \$1.25/day 2005 PPP Line**

Figure 3 (Government-definition \$1.25/day 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	98.7
15-19	98.5
20-24	96.3
25-29	95.1
30-34	95.1
35-39	91.0
40-44	82.9
45-49	77.9
50-54	58.6
55-59	49.6
60-64	44.4
65-69	34.8
70-74	23.3
75-79	17.7
80-84	11.9
85-89	4.8
90-94	2.4
95-100	0.0

Figure 5 (Government-definition \$1.25/day line): For each score range, average differences between estimated and true poverty likelihoods for households, with confidence intervals, from 1,000 bootstraps of $n = 16,384$, 2010/11 scorecard applied to the 2010/11 validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	0.0	0.0	0.0	0.0
5-9	0.0	0.0	0.0	0.0
10-14	-1.3	0.7	0.7	0.7
15-19	+0.7	1.5	1.9	2.5
20-24	+0.8	1.5	1.8	2.3
25-29	+0.7	1.3	1.5	2.1
30-34	+4.1	1.4	1.6	2.1
35-39	+3.4	1.6	1.9	2.5
40-44	+3.4	1.9	2.2	2.9
45-49	+3.6	2.0	2.5	3.3
50-54	-5.9	4.1	4.4	4.8
55-59	-9.4	5.8	6.1	6.3
60-64	+5.1	2.4	2.8	4.0
65-69	+5.8	2.7	3.4	4.3
70-74	+2.7	2.5	3.1	3.9
75-79	-0.4	3.4	4.1	5.1
80-84	+6.3	1.8	2.3	3.2
85-89	-0.4	2.3	2.8	3.7
90-94	-5.6	4.9	5.4	6.1
95-100	0.0	0.0	0.0	0.0

Figure 6 (Government-definition \$1.25/day line):

Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	-1.1	66.7	74.1	91.6
4	+1.0	35.6	41.8	53.9
8	+0.9	26.1	31.4	38.6
16	+1.0	18.9	22.8	28.1
32	+1.1	13.2	16.0	21.0
64	+1.2	9.5	11.5	14.3
128	+1.2	6.6	7.8	10.3
256	+1.2	4.4	5.4	6.7
512	+1.1	3.5	4.1	4.9
1,024	+1.0	2.4	2.8	3.4
2,048	+1.1	1.6	1.9	2.5
4,096	+1.1	1.1	1.3	1.8
8,192	+1.1	0.8	0.9	1.2
16,384	+1.1	0.6	0.7	0.9

Figure 10 (Government-definition \$1.25/day line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	64.5	0.0	35.4	35.5	-99.6
≤9	0.3	64.3	0.0	35.4	35.7	-99.0
≤14	1.3	63.3	0.0	35.4	36.7	-96.0
≤19	3.1	61.5	0.0	35.3	38.5	-90.2
≤24	7.0	57.6	0.2	35.2	42.2	-78.0
≤29	13.6	51.1	0.6	34.8	48.4	-57.2
≤34	21.3	43.3	1.3	34.1	55.4	-32.1
≤39	29.6	35.0	2.4	33.0	62.6	-4.6
≤44	38.0	26.7	4.4	30.9	68.9	+24.3
≤49	45.1	19.6	7.1	28.3	73.4	+50.5
≤54	51.2	13.4	10.3	25.1	76.3	+74.4
≤59	57.1	7.5	14.4	21.0	78.1	+77.7
≤64	60.7	3.9	19.6	15.8	76.6	+69.7
≤69	62.6	2.1	24.1	11.3	73.8	+62.7
≤74	63.7	1.0	27.9	7.4	71.1	+56.8
≤79	64.3	0.4	30.5	4.9	69.2	+52.9
≤84	64.5	0.2	32.2	3.1	67.6	+50.1
≤89	64.5	0.1	33.5	1.9	66.4	+48.2
≤94	64.6	0.0	34.3	1.1	65.7	+46.9
≤100	64.6	0.0	35.4	0.0	64.6	+45.3

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

Figure 11 (Government-definition \$1.25/day line): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	100.0	0.2	Only poor targeted
≤9	0.3	100.0	0.5	Only poor targeted
≤14	1.3	100.0	2.0	Only poor targeted
≤19	3.2	99.1	4.9	105.3:1
≤24	7.2	97.3	10.9	35.8:1
≤29	14.1	96.0	21.0	24.2:1
≤34	22.6	94.3	33.0	16.6:1
≤39	32.0	92.5	45.8	12.3:1
≤44	42.4	89.5	58.7	8.6:1
≤49	52.2	86.4	69.7	6.4:1
≤54	61.5	83.3	79.3	5.0:1
≤59	71.5	79.9	88.4	4.0:1
≤64	80.3	75.6	94.0	3.1:1
≤69	86.7	72.2	96.8	2.6:1
≤74	91.6	69.5	98.5	2.3:1
≤79	94.7	67.9	99.5	2.1:1
≤84	96.7	66.7	99.7	2.0:1
≤89	98.0	65.9	99.9	1.9:1
≤94	98.9	65.3	100.0	1.9:1
≤100	100.0	64.6	100.0	1.8:1

**Figures for
the Government-Definition \$2.00/day 2005 PPP Line**

Figure 3 (Government-definition \$2.00/day 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	99.4
25-29	99.4
30-34	98.9
35-39	98.0
40-44	97.2
45-49	95.0
50-54	87.8
55-59	83.4
60-64	75.4
65-69	67.6
70-74	51.4
75-79	44.2
80-84	38.6
85-89	16.5
90-94	12.7
95-100	3.4

Figure 5 (Government-definition \$2.00/day line): For each score range, average differences between estimated and true poverty likelihoods for households, with confidence intervals, from 1,000 bootstraps of $n = 16,384$, 2010/11 scorecard applied to the 2010/11 validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	0.0	0.0	0.0	0.0
5-9	0.0	0.0	0.0	0.0
10-14	0.0	0.0	0.0	0.0
15-19	+2.2	1.5	1.9	2.5
20-24	-0.5	0.3	0.3	0.3
25-29	+0.6	0.7	0.8	1.1
30-34	+1.4	0.8	0.9	1.2
35-39	-0.6	0.5	0.6	0.7
40-44	+2.3	1.1	1.3	1.7
45-49	+1.4	1.1	1.3	1.7
50-54	-1.5	1.5	1.7	2.1
55-59	-3.2	2.4	2.5	2.7
60-64	+4.4	2.2	2.6	3.4
65-69	+2.9	2.9	3.4	4.3
70-74	+3.4	3.6	4.4	5.9
75-79	-2.8	4.6	5.5	7.2
80-84	+18.9	3.7	4.4	5.6
85-89	-1.3	4.2	5.0	6.3
90-94	-4.3	5.6	6.7	8.7
95-100	-0.6	2.3	2.8	3.7

Figure 6 (Government-definition \$2.00/day line):

Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+1.2	53.9	69.6	85.0
4	+1.0	27.3	34.2	46.6
8	+1.0	19.3	23.5	30.1
16	+1.2	14.4	17.6	22.1
32	+1.4	10.6	12.3	15.9
64	+1.4	7.6	9.0	11.8
128	+1.3	5.2	6.2	8.1
256	+1.2	3.5	4.4	5.5
512	+1.2	2.6	3.1	4.2
1,024	+1.2	1.8	2.2	2.8
2,048	+1.2	1.4	1.7	2.1
4,096	+1.2	1.0	1.1	1.4
8,192	+1.2	0.7	0.8	1.1
16,384	+1.1	0.5	0.6	0.8

Figure 10 (Government-definition \$2.00/day line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	83.1	0.0	16.8	16.9	-99.7
≤9	0.3	82.8	0.0	16.8	17.2	-99.2
≤14	1.3	81.9	0.0	16.8	18.1	-96.9
≤19	3.1	80.0	0.0	16.8	19.9	-92.4
≤24	7.2	76.0	0.0	16.8	24.0	-82.7
≤29	14.0	69.2	0.1	16.7	30.7	-66.1
≤34	22.3	60.9	0.3	16.5	38.9	-46.0
≤39	31.6	51.6	0.4	16.4	48.0	-23.5
≤44	41.5	41.7	0.9	15.9	57.4	+0.9
≤49	50.6	32.6	1.6	15.2	65.8	+23.5
≤54	59.0	24.2	2.6	14.3	73.2	+44.8
≤59	67.6	15.6	4.0	12.8	80.4	+67.2
≤64	73.8	9.4	6.5	10.3	84.1	+85.3
≤69	78.0	5.2	8.6	8.2	86.2	+89.6
≤74	80.6	2.6	11.1	5.8	86.3	+86.7
≤79	82.1	1.1	12.7	4.1	86.2	+84.7
≤84	82.6	0.6	14.1	2.7	85.3	+83.1
≤89	82.9	0.2	15.1	1.8	84.7	+81.9
≤94	83.1	0.1	15.8	1.0	84.1	+81.0
≤100	83.2	0.0	16.8	0.0	83.2	+79.8

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

Figure 11 (Government-definition \$2.00/day line): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	100.0	0.2	Only poor targeted
≤9	0.3	100.0	0.4	Only poor targeted
≤14	1.3	100.0	1.6	Only poor targeted
≤19	3.2	99.1	3.8	105.3:1
≤24	7.2	99.5	8.6	184.7:1
≤29	14.1	99.3	16.9	142.1:1
≤34	22.6	98.8	26.8	82.0:1
≤39	32.0	98.6	38.0	71.9:1
≤44	42.4	97.9	49.9	45.9:1
≤49	52.2	97.0	60.8	31.9:1
≤54	61.5	95.8	70.9	23.0:1
≤59	71.5	94.4	81.2	17.0:1
≤64	80.3	91.9	88.7	11.3:1
≤69	86.7	90.0	93.8	9.0:1
≤74	91.6	87.9	96.8	7.3:1
≤79	94.7	86.6	98.6	6.5:1
≤84	96.7	85.4	99.3	5.9:1
≤89	98.0	84.6	99.7	5.5:1
≤94	98.9	84.0	99.9	5.3:1
≤100	100.0	83.2	100.0	4.9:1

**Figures for
the Government-Definition \$2.50/day 2005 PPP Line**

Figure 3 (Government-definition \$2.50/day 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	100.0
25-29	99.8
30-34	99.8
35-39	99.7
40-44	99.2
45-49	97.5
50-54	93.5
55-59	90.4
60-64	85.4
65-69	79.8
70-74	65.0
75-79	56.7
80-84	50.3
85-89	27.1
90-94	18.0
95-100	13.0

Figure 5 (Government-definition \$2.50/day line): For each score range, average differences between estimated and true poverty likelihoods for households, with confidence intervals, from 1,000 bootstraps of $n = 16,384$, 2010/11 scorecard applied to the 2010/11 validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	0.0	0.0	0.0	0.0
5-9	0.0	0.0	0.0	0.0
10-14	0.0	0.0	0.0	0.0
15-19	0.0	0.0	0.0	0.0
20-24	0.0	0.0	0.0	0.0
25-29	-0.1	0.1	0.1	0.1
30-34	+0.2	0.3	0.4	0.5
35-39	0.0	0.2	0.3	0.4
40-44	+2.1	1.0	1.2	1.5
45-49	+0.8	0.8	1.0	1.3
50-54	-1.4	1.2	1.3	1.5
55-59	-1.0	1.4	1.6	2.1
60-64	+4.9	2.0	2.4	3.1
65-69	+4.1	2.6	3.0	4.0
70-74	+4.4	4.0	4.8	6.0
75-79	+0.8	4.5	5.4	7.4
80-84	+8.5	5.6	6.7	8.8
85-89	-21.4	14.2	14.9	15.6
90-94	-12.2	9.6	10.3	11.6
95-100	+3.9	3.3	4.0	5.4

Figure 6 (Government-definition \$2.50/day line):

Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+1.1	50.0	64.3	85.5
4	+1.3	25.1	32.5	47.7
8	+1.0	18.4	22.3	30.3
16	+1.0	13.8	16.7	22.4
32	+1.1	9.9	11.6	14.9
64	+1.0	7.1	8.5	11.2
128	+1.0	5.1	6.2	8.0
256	+0.9	3.4	4.0	5.3
512	+0.9	2.4	2.8	3.7
1,024	+0.9	1.7	2.1	2.6
2,048	+0.9	1.3	1.6	2.2
4,096	+0.9	0.9	1.1	1.5
8,192	+0.9	0.6	0.8	1.0
16,384	+0.9	0.5	0.5	0.7

Figure 10 (Government-definition \$2.50/day line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	88.5	0.0	11.4	11.5	-99.7
≤9	0.3	88.3	0.0	11.4	11.7	-99.2
≤14	1.3	87.3	0.0	11.4	12.7	-97.1
≤19	3.2	85.5	0.0	11.4	14.5	-92.8
≤24	7.2	81.4	0.0	11.4	18.6	-83.7
≤29	14.1	74.5	0.0	11.3	25.5	-68.1
≤34	22.5	66.1	0.0	11.3	33.9	-49.1
≤39	31.9	56.7	0.1	11.3	43.2	-27.8
≤44	42.1	46.5	0.3	11.1	53.2	-4.6
≤49	51.6	37.1	0.6	10.8	62.4	+17.0
≤54	60.4	28.2	1.1	10.2	70.6	+37.5
≤59	69.6	19.0	1.9	9.5	79.1	+59.3
≤64	76.7	11.9	3.6	7.8	84.6	+77.2
≤69	81.6	7.0	5.0	6.3	88.0	+89.9
≤74	84.8	3.8	6.8	4.5	89.3	+92.3
≤79	86.6	2.0	8.1	3.2	89.8	+90.8
≤84	87.6	1.0	9.1	2.3	89.9	+89.7
≤89	88.2	0.4	9.8	1.6	89.8	+88.9
≤94	88.5	0.1	10.4	0.9	89.5	+88.2
≤100	88.6	0.0	11.4	0.0	88.6	+87.2

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

Figure 11 (Government-definition \$2.50/day line): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	100.0	0.1	Only poor targeted
≤9	0.3	100.0	0.4	Only poor targeted
≤14	1.3	100.0	1.5	Only poor targeted
≤19	3.2	100.0	3.6	Only poor targeted
≤24	7.2	100.0	8.1	Only poor targeted
≤29	14.1	99.9	15.9	1,289.9:1
≤34	22.6	99.8	25.4	567.9:1
≤39	32.0	99.7	36.0	374.4:1
≤44	42.4	99.4	47.5	154.8:1
≤49	52.2	98.9	58.2	87.9:1
≤54	61.5	98.2	68.1	53.9:1
≤59	71.5	97.3	78.6	36.7:1
≤64	80.3	95.6	86.6	21.6:1
≤69	86.7	94.2	92.1	16.2:1
≤74	91.6	92.6	95.7	12.4:1
≤79	94.7	91.4	97.7	10.7:1
≤84	96.7	90.6	98.8	9.6:1
≤89	98.0	90.0	99.5	9.0:1
≤94	98.9	89.5	99.9	8.5:1
≤100	100.0	88.6	100.0	7.8:1

**Figures for
the Government-Definition \$5.00/day 2005 PPP Line**

Figure 3 (Government-definition \$5.00/day 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	100.0
25-29	100.0
30-34	100.0
35-39	100.0
40-44	100.0
45-49	100.0
50-54	99.8
55-59	98.8
60-64	96.9
65-69	95.9
70-74	92.0
75-79	87.9
80-84	83.8
85-89	76.4
90-94	60.6
95-100	53.9

Figure 5 (Government-definition \$5.00/day line): For each score range, average differences between estimated and true poverty likelihoods for households, with confidence intervals, from 1,000 bootstraps of $n = 16,384$, 2010/11 scorecard applied to the 2010/11 validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	0.0	0.0	0.0	0.0
5-9	0.0	0.0	0.0	0.0
10-14	0.0	0.0	0.0	0.0
15-19	0.0	0.0	0.0	0.0
20-24	0.0	0.0	0.0	0.0
25-29	0.0	0.0	0.0	0.0
30-34	0.0	0.0	0.0	0.0
35-39	0.0	0.0	0.0	0.0
40-44	0.0	0.0	0.0	0.0
45-49	0.0	0.1	0.1	0.1
50-54	-0.2	0.1	0.1	0.1
55-59	-0.5	0.4	0.4	0.5
60-64	+3.0	1.2	1.5	1.9
65-69	0.0	1.1	1.3	1.7
70-74	+8.5	4.1	4.7	6.1
75-79	+8.3	3.9	4.5	6.0
80-84	+19.4	6.9	7.8	10.0
85-89	-8.3	6.1	6.4	6.9
90-94	-18.1	11.9	12.6	13.4
95-100	+17.6	6.7	8.0	10.3

Figure 6 (Government-definition \$5.00/day line):

Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+0.6	6.0	46.3	67.9
4	+0.5	14.9	20.0	36.7
8	+0.7	10.6	15.4	25.2
16	+0.9	9.0	13.2	18.6
32	+1.1	7.7	9.4	12.1
64	+1.1	5.6	6.4	8.9
128	+1.3	4.2	4.8	5.9
256	+1.3	2.8	3.3	4.2
512	+1.3	2.0	2.4	3.2
1,024	+1.3	1.5	1.8	2.2
2,048	+1.4	1.1	1.3	1.6
4,096	+1.3	0.7	0.9	1.2
8,192	+1.3	0.5	0.6	0.8
16,384	+1.3	0.4	0.5	0.6

Figure 10 (Government-definition \$5.00/day line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	96.3	0.0	3.5	3.7	-99.7
≤9	0.3	96.1	0.0	3.5	3.9	-99.3
≤14	1.3	95.2	0.0	3.5	4.8	-97.3
≤19	3.2	93.3	0.0	3.5	6.7	-93.4
≤24	7.2	89.3	0.0	3.5	10.7	-85.0
≤29	14.1	82.3	0.0	3.5	17.7	-70.7
≤34	22.6	73.9	0.0	3.5	26.1	-53.2
≤39	32.0	64.4	0.0	3.5	35.6	-33.6
≤44	42.4	54.1	0.0	3.5	45.9	-12.1
≤49	52.1	44.3	0.0	3.5	55.7	+8.1
≤54	61.5	35.0	0.0	3.5	65.0	+27.5
≤59	71.4	25.0	0.1	3.4	74.9	+48.2
≤64	79.6	16.8	0.7	2.9	82.5	+65.8
≤69	85.7	10.8	0.9	2.6	88.3	+78.7
≤74	90.2	6.2	1.4	2.1	92.4	+88.5
≤79	92.8	3.7	1.9	1.6	94.4	+94.4
≤84	94.3	2.1	2.4	1.2	95.5	+97.6
≤89	95.4	1.1	2.6	0.9	96.2	+97.3
≤94	96.1	0.4	2.9	0.6	96.7	+97.0
≤100	96.5	0.0	3.5	0.0	96.5	+96.3

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

Figure 11 (Government-definition \$5.00/day line): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	100.0	0.1	Only poor targeted
≤9	0.3	100.0	0.3	Only poor targeted
≤14	1.3	100.0	1.3	Only poor targeted
≤19	3.2	100.0	3.3	Only poor targeted
≤24	7.2	100.0	7.5	Only poor targeted
≤29	14.1	100.0	14.6	Only poor targeted
≤34	22.6	100.0	23.4	Only poor targeted
≤39	32.0	100.0	33.2	Only poor targeted
≤44	42.4	100.0	44.0	Only poor targeted
≤49	52.2	100.0	54.1	2,834.1:1
≤54	61.5	100.0	63.7	2,545.9:1
≤59	71.5	99.8	74.0	653.2:1
≤64	80.3	99.2	82.6	122.0:1
≤69	86.7	98.9	88.9	90.6:1
≤74	91.6	98.5	93.5	64.8:1
≤79	94.7	97.9	96.2	47.7:1
≤84	96.7	97.6	97.8	40.0:1
≤89	98.0	97.3	98.9	36.0:1
≤94	98.9	97.1	99.6	33.3:1
≤100	100.0	96.5	100.0	27.3:1

**Figures for
the Government-Definition \$8.44/day 2005 PPP Line**

Figure 3 (Government-definition \$8.44/day 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	100.0
25-29	100.0
30-34	100.0
35-39	100.0
40-44	100.0
45-49	100.0
50-54	100.0
55-59	99.8
60-64	98.9
65-69	98.5
70-74	97.6
75-79	95.9
80-84	93.9
85-89	92.0
90-94	84.0
95-100	69.7

Figure 5 (Government-definition \$8.44/day line): For each score range, average differences between estimated and true poverty likelihoods for households, with confidence intervals, from 1,000 bootstraps of $n = 16,384$, 2010/11 scorecard applied to the 2010/11 validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	0.0	0.0	0.0	0.0
5-9	0.0	0.0	0.0	0.0
10-14	0.0	0.0	0.0	0.0
15-19	0.0	0.0	0.0	0.0
20-24	0.0	0.0	0.0	0.0
25-29	0.0	0.0	0.0	0.0
30-34	0.0	0.0	0.0	0.0
35-39	0.0	0.0	0.0	0.0
40-44	0.0	0.0	0.0	0.0
45-49	0.0	0.0	0.0	0.0
50-54	0.0	0.0	0.0	0.0
55-59	-0.2	0.1	0.1	0.1
60-64	+0.2	0.4	0.5	0.7
65-69	-1.0	0.6	0.7	0.7
70-74	-0.4	0.7	0.9	1.2
75-79	+2.4	1.7	2.1	2.7
80-84	+21.7	7.4	8.5	10.8
85-89	-1.7	2.6	3.2	4.4
90-94	-8.6	6.1	6.4	7.0
95-100	-2.2	6.5	7.7	9.9

Figure 6 (Government-definition \$8.44/day line):

Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+0.1	2.1	4.0	59.9
4	0.0	5.8	13.5	21.4
8	0.0	5.8	7.9	13.8
16	+0.3	4.4	5.6	15.3
32	+0.3	3.6	7.8	10.0
64	+0.4	4.1	4.9	6.5
128	+0.4	2.7	3.3	4.7
256	+0.4	1.9	2.3	3.1
512	+0.4	1.4	1.7	2.2
1,024	+0.5	1.0	1.2	1.8
2,048	+0.5	0.7	0.9	1.2
4,096	+0.5	0.5	0.6	0.8
8,192	+0.5	0.4	0.4	0.6
16,384	+0.4	0.3	0.3	0.4

Figure 10 (Government-definition \$8.44/day line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	98.6	0.0	1.3	1.4	-99.7
≤9	0.3	98.4	0.0	1.3	1.6	-99.3
≤14	1.3	97.4	0.0	1.3	2.6	-97.4
≤19	3.2	95.6	0.0	1.3	4.4	-93.6
≤24	7.2	91.5	0.0	1.3	8.5	-85.4
≤29	14.1	84.6	0.0	1.3	15.4	-71.4
≤34	22.6	76.1	0.0	1.3	23.9	-54.2
≤39	32.0	66.7	0.0	1.3	33.3	-35.1
≤44	42.4	56.3	0.0	1.3	43.7	-14.1
≤49	52.2	46.6	0.0	1.3	53.4	+5.7
≤54	61.5	37.2	0.0	1.3	62.8	+24.6
≤59	71.5	27.2	0.0	1.3	72.8	+44.9
≤64	80.1	18.6	0.2	1.1	81.2	+62.5
≤69	86.4	12.3	0.2	1.1	87.5	+75.3
≤74	91.3	7.5	0.4	0.9	92.2	+85.3
≤79	94.1	4.6	0.6	0.7	94.8	+91.3
≤84	95.9	2.8	0.8	0.5	96.4	+95.1
≤89	97.1	1.6	0.9	0.4	97.5	+97.7
≤94	98.0	0.7	1.0	0.3	98.3	+99.0
≤100	98.7	0.0	1.3	0.0	98.7	+98.7

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

Figure 11 (Government-definition \$8.44/day line): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	100.0	0.1	Only poor targeted
≤9	0.3	100.0	0.3	Only poor targeted
≤14	1.3	100.0	1.3	Only poor targeted
≤19	3.2	100.0	3.2	Only poor targeted
≤24	7.2	100.0	7.3	Only poor targeted
≤29	14.1	100.0	14.3	Only poor targeted
≤34	22.6	100.0	22.9	Only poor targeted
≤39	32.0	100.0	32.4	Only poor targeted
≤44	42.4	100.0	42.9	Only poor targeted
≤49	52.2	100.0	52.8	Only poor targeted
≤54	61.5	100.0	62.3	Only poor targeted
≤59	71.5	100.0	72.5	29,073.4:1
≤64	80.3	99.8	81.2	470.5:1
≤69	86.7	99.8	87.6	399.2:1
≤74	91.6	99.6	92.4	253.4:1
≤79	94.7	99.4	95.3	154.1:1
≤84	96.7	99.2	97.2	121.9:1
≤89	98.0	99.1	98.4	108.6:1
≤94	98.9	99.0	99.2	101.9:1
≤100	100.0	98.7	100.0	77.4:1

**Figures for
the Old-Definition \$1.25/day 2005 PPP Poverty Line**

Figure 3 (Old-definition \$1.25/day 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	98.7
15-19	98.5
20-24	96.4
25-29	95.5
30-34	95.5
35-39	91.3
40-44	83.4
45-49	78.7
50-54	60.8
55-59	52.4
60-64	46.6
65-69	36.6
70-74	23.6
75-79	18.6
80-84	13.2
85-89	4.8
90-94	2.4
95-100	0.6

Figure 5 (Old-definition \$1.25/day line): For each score range, average differences between estimated and true poverty likelihoods for households, with confidence intervals, from 1,000 bootstraps of $n = 16,384$, 2010/11 scorecard applied to the 2010/11 validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	0.0	0.0	0.0	0.0
5-9	0.0	0.0	0.0	0.0
10-14	-1.3	0.7	0.7	0.7
15-19	+0.7	1.5	1.9	2.5
20-24	-0.9	1.1	1.3	1.7
25-29	+0.5	1.2	1.4	1.9
30-34	+3.8	1.3	1.5	2.0
35-39	+2.5	1.6	1.9	2.4
40-44	+2.9	1.8	2.2	2.9
45-49	+4.1	2.0	2.5	3.3
50-54	-5.2	3.8	4.0	4.3
55-59	-9.0	5.6	5.9	6.3
60-64	+4.9	2.3	2.8	4.0
65-69	+6.3	2.8	3.3	4.4
70-74	+2.4	2.6	3.1	4.1
75-79	-0.8	3.5	4.2	5.3
80-84	+7.6	1.8	2.3	3.2
85-89	-0.5	2.3	2.8	3.7
90-94	-8.9	6.9	7.4	8.8
95-100	+0.6	0.0	0.0	0.0

Figure 6 (Old-definition \$1.25/day line): Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	-1.3	66.0	73.4	91.2
4	+0.8	34.8	41.0	53.3
8	+0.8	25.5	30.6	38.1
16	+0.9	18.7	22.4	27.9
32	+1.0	13.0	15.2	21.3
64	+1.0	9.2	11.0	13.8
128	+1.1	6.4	7.7	9.4
256	+1.1	4.4	5.2	6.7
512	+1.0	3.3	4.0	4.6
1,024	+1.0	2.3	2.7	3.6
2,048	+1.0	1.6	1.9	2.5
4,096	+1.0	1.1	1.3	1.8
8,192	+1.0	0.8	0.9	1.2
16,384	+1.0	0.6	0.7	0.9

Figure 10 (Old-definition \$1.25/day line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	65.7	0.0	34.2	34.3	-99.6
≤9	0.3	65.5	0.0	34.2	34.5	-99.0
≤14	1.3	64.5	0.0	34.2	35.5	-96.0
≤19	3.1	62.7	0.0	34.2	37.3	-90.4
≤24	7.1	58.7	0.2	34.1	41.1	-78.3
≤29	13.6	52.2	0.5	33.7	47.4	-57.8
≤34	21.4	44.4	1.2	33.0	54.4	-33.1
≤39	29.9	35.9	2.2	32.0	61.9	-5.9
≤44	38.3	27.5	4.1	30.1	68.4	+22.7
≤49	45.5	20.3	6.7	27.5	73.0	+48.4
≤54	51.8	14.0	9.7	24.5	76.2	+72.2
≤59	57.9	7.9	13.7	20.5	78.4	+79.2
≤64	61.7	4.1	18.6	15.6	77.2	+71.7
≤69	63.6	2.2	23.1	11.1	74.7	+64.9
≤74	64.7	1.0	26.9	7.3	72.1	+59.1
≤79	65.4	0.4	29.3	4.9	70.3	+55.4
≤84	65.6	0.2	31.1	3.1	68.7	+52.7
≤89	65.7	0.1	32.3	1.9	67.6	+50.9
≤94	65.8	0.0	33.2	1.1	66.8	+49.6
≤100	65.8	0.0	34.2	0.0	65.8	+48.0

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

Figure 11 (Old-definition \$1.25/day line): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	100.0	0.2	Only poor targeted
≤9	0.3	100.0	0.5	Only poor targeted
≤14	1.3	100.0	2.0	Only poor targeted
≤19	3.2	99.1	4.8	105.3:1
≤24	7.2	97.9	10.7	45.5:1
≤29	14.1	96.5	20.7	27.8:1
≤34	22.6	94.8	32.5	18.2:1
≤39	32.0	93.2	45.4	13.8:1
≤44	42.4	90.3	58.2	9.4:1
≤49	52.2	87.2	69.1	6.8:1
≤54	61.5	84.2	78.7	5.3:1
≤59	71.5	80.9	88.0	4.2:1
≤64	80.3	76.8	93.7	3.3:1
≤69	86.7	73.4	96.6	2.8:1
≤74	91.6	70.7	98.4	2.4:1
≤79	94.7	69.0	99.4	2.2:1
≤84	96.7	67.8	99.7	2.1:1
≤89	98.0	67.0	99.8	2.0:1
≤94	98.9	66.5	100.0	2.0:1
≤100	100.0	65.8	100.0	1.9:1

**Figures for
the Old-Definition \$2.50/day 2005 PPP Poverty Line**

Figure 3 (Old-definition \$2.50/day 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	100.0
25-29	99.8
30-34	99.8
35-39	99.8
40-44	99.5
45-49	98.8
50-54	95.4
55-59	90.6
60-64	86.0
65-69	80.9
70-74	68.2
75-79	59.2
80-84	52.5
85-89	28.5
90-94	21.4
95-100	15.7

Figure 5 (Old-definition \$2.50/day line): For each score range, average differences between estimated and true poverty likelihoods for households, with confidence intervals, from 1,000 bootstraps of $n = 16,384$, 2010/11 scorecard applied to the 2010/11 validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	0.0	0.0	0.0	0.0
5-9	0.0	0.0	0.0	0.0
10-14	0.0	0.0	0.0	0.0
15-19	0.0	0.0	0.0	0.0
20-24	0.0	0.0	0.0	0.0
25-29	-0.1	0.1	0.1	0.1
30-34	+0.3	0.3	0.4	0.5
35-39	+0.1	0.2	0.3	0.4
40-44	+2.2	0.9	1.2	1.5
45-49	+2.0	0.8	1.0	1.3
50-54	-0.6	0.8	1.0	1.3
55-59	-0.9	1.4	1.7	2.1
60-64	+4.6	1.9	2.3	3.0
65-69	+3.0	2.5	3.0	4.0
70-74	+7.2	4.0	4.7	6.1
75-79	+2.7	4.4	5.5	7.5
80-84	+9.2	5.7	6.8	8.9
85-89	-25.2	15.9	16.8	17.6
90-94	-8.8	8.0	8.6	10.7
95-100	+6.3	3.3	4.1	5.4

Figure 6 (Old-definition \$2.50/day line): Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+1.5	50.0	63.4	85.2
4	+1.5	24.5	32.0	47.8
8	+1.3	18.5	22.1	30.4
16	+1.3	13.5	16.6	22.5
32	+1.4	9.9	11.7	15.4
64	+1.4	7.2	8.5	11.2
128	+1.3	5.2	6.0	7.9
256	+1.3	3.4	4.0	5.5
512	+1.3	2.4	2.9	3.7
1,024	+1.3	1.7	2.1	2.7
2,048	+1.3	1.2	1.6	2.1
4,096	+1.3	0.9	1.0	1.5
8,192	+1.3	0.6	0.8	1.0
16,384	+1.3	0.5	0.5	0.7

Figure 10 (Old-definition \$2.50/day line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	89.1	0.0	10.8	10.9	-99.7
≤9	0.3	88.9	0.0	10.8	11.1	-99.3
≤14	1.3	87.9	0.0	10.8	12.1	-97.1
≤19	3.2	86.0	0.0	10.8	14.0	-92.9
≤24	7.2	82.0	0.0	10.8	18.0	-83.8
≤29	14.1	75.1	0.0	10.8	24.9	-68.3
≤34	22.5	66.6	0.0	10.8	33.3	-49.4
≤39	31.9	57.2	0.1	10.7	42.7	-28.3
≤44	42.2	47.0	0.2	10.6	52.7	-5.2
≤49	51.6	37.6	0.6	10.3	61.9	+16.4
≤54	60.5	28.7	1.0	9.8	70.3	+36.8
≤59	69.8	19.4	1.8	9.1	78.8	+58.5
≤64	77.0	12.2	3.3	7.5	84.5	+76.4
≤69	82.0	7.2	4.7	6.2	88.2	+89.1
≤74	85.2	4.0	6.4	4.4	89.6	+92.8
≤79	87.0	2.1	7.7	3.1	90.2	+91.4
≤84	88.1	1.1	8.6	2.2	90.3	+90.3
≤89	88.8	0.4	9.3	1.6	90.3	+89.6
≤94	89.1	0.1	9.9	0.9	90.0	+88.9
≤100	89.2	0.0	10.8	0.0	89.2	+87.9

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

Figure 11 (Old-definition \$2.50/day line): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	100.0	0.1	Only poor targeted
≤9	0.3	100.0	0.4	Only poor targeted
≤14	1.3	100.0	1.5	Only poor targeted
≤19	3.2	100.0	3.6	Only poor targeted
≤24	7.2	100.0	8.1	Only poor targeted
≤29	14.1	99.9	15.8	1,289.9:1
≤34	22.6	99.8	25.3	567.9:1
≤39	32.0	99.7	35.8	374.4:1
≤44	42.4	99.4	47.3	169.2:1
≤49	52.2	98.9	57.9	93.2:1
≤54	61.5	98.4	67.9	60.8:1
≤59	71.5	97.5	78.2	39.8:1
≤64	80.3	95.9	86.3	23.3:1
≤69	86.7	94.6	91.9	17.6:1
≤74	91.6	93.0	95.5	13.3:1
≤79	94.7	91.9	97.6	11.3:1
≤84	96.7	91.1	98.7	10.2:1
≤89	98.0	90.6	99.5	9.6:1
≤94	98.9	90.0	99.9	9.0:1
≤100	100.0	89.2	100.0	8.2:1

**Figures for
the Govt.-definition \$1.90/day 2011 PPP Poverty Line**

Figure 3 (Govt.-def. \$1.90/day 2011 PPP): 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	97.9
25-29	96.4
30-34	96.4
35-39	94.5
40-44	88.7
45-49	84.9
50-54	69.1
55-59	64.3
60-64	56.5
65-69	43.6
70-74	31.1
75-79	24.5
80-84	18.0
85-89	9.4
90-94	6.1
95-100	0.6

Figure 5 (Govt.-def. \$1.90/day 2011 PPP): For each score range, average differences between estimated and true poverty likelihoods for households, with confidence intervals, from 1,000 bootstraps of $n = 16,384$, 2010/11 scorecard applied to the 2010/11 validation sample

Score	Difference between estimate and observed value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	0.0	0.0	0.0	0.0
5-9	0.0	0.0	0.0	0.0
10-14	0.0	0.0	0.0	0.0
15-19	+2.1	1.5	1.9	2.5
20-24	0.0	1.0	1.1	1.6
25-29	+1.1	1.2	1.4	1.9
30-34	+2.4	1.2	1.3	1.8
35-39	+1.5	1.2	1.5	2.0
40-44	+1.8	1.6	1.9	2.4
45-49	+2.6	1.8	2.1	2.8
50-54	-5.9	4.0	4.2	4.6
55-59	-6.9	4.5	4.7	5.1
60-64	+7.0	2.3	2.9	4.3
65-69	+2.5	3.0	3.5	4.5
70-74	+1.4	3.1	3.8	4.8
75-79	-5.3	4.9	5.4	6.7
80-84	+11.5	2.0	2.5	3.1
85-89	+3.5	2.4	3.0	3.9
90-94	-6.2	5.5	6.1	8.0
95-100	+0.6	0.0	0.0	0.0

Figure 6 (Govt.-def. \$1.90/day 2011 PPP): Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and observed value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	-0.7	62.7	78.8	89.2
4	+0.6	32.9	39.4	53.3
8	+0.3	23.4	28.0	37.5
16	+0.6	17.6	20.9	27.5
32	+0.9	12.4	14.6	19.8
64	+0.9	8.7	10.4	14.8
128	+0.8	5.9	7.0	9.4
256	+0.8	4.1	4.8	6.3
512	+0.7	3.2	3.8	4.6
1,024	+0.7	2.2	2.6	3.4
2,048	+0.7	1.5	1.8	2.4
4,096	+0.7	1.1	1.3	1.7
8,192	+0.7	0.8	0.9	1.1
16,384	+0.7	0.6	0.7	0.9

Figure 10 (Govt.-def. \$1.90/day 2011 PPP): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	<u>Inclusion:</u> Poor correctly targeted	<u>Undercoverage:</u> Poor mistakenly not targeted	<u>Leakage:</u> Non-poor mistakenly targeted	<u>Exclusion:</u> Non-poor correctly not targeted	<u>Hit rate</u> Inclusion + Exclusion	<u>BPAC</u> See text
<=4	0.1	71.3	0.0	28.6	28.7	-99.6
<=9	0.3	71.1	0.0	28.6	28.9	-99.1
<=14	1.3	70.1	0.0	28.6	29.9	-96.4
<=19	3.1	68.3	0.0	28.5	31.7	-91.2
<=24	7.1	64.3	0.1	28.4	35.5	-80.0
<=29	13.7	57.7	0.4	28.1	41.9	-61.0
<=34	21.7	49.7	0.9	27.7	49.4	-38.0
<=39	30.5	41.0	1.6	27.0	57.5	-12.5
<=44	39.5	31.9	2.9	25.7	65.2	+14.7
<=49	47.4	24.0	4.7	23.8	71.2	+39.4
<=54	54.5	17.0	7.0	21.5	76.0	+62.4
<=59	61.5	10.0	10.1	18.5	80.0	+85.9
<=64	66.0	5.5	14.3	14.2	80.2	+80.0
<=69	68.6	2.9	18.1	10.5	79.0	+74.7
<=74	70.1	1.3	21.5	7.0	77.2	+69.9
<=79	71.0	0.4	23.7	4.8	75.8	+66.8
<=84	71.2	0.2	25.5	3.1	74.3	+64.3
<=89	71.3	0.1	26.7	1.9	73.2	+62.6
<=94	71.4	0.0	27.5	1.1	72.5	+61.5
<=100	71.4	0.0	28.6	0.0	71.4	+60.0

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

Figure 11 (Govt.-def. \$1.90/day 2011 PPP): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
<=4	0.1	100.0	0.2	Only poor targeted
<=9	0.3	100.0	0.5	Only poor targeted
<=14	1.3	100.0	1.8	Only poor targeted
<=19	3.2	99.1	4.4	105.3:1
<=24	7.2	98.3	9.9	59.5:1
<=29	14.1	97.1	19.2	33.0:1
<=34	22.6	96.1	30.4	24.6:1
<=39	32.0	95.2	42.7	19.7:1
<=44	42.4	93.3	55.4	13.8:1
<=49	52.2	90.9	66.4	10.0:1
<=54	61.5	88.6	76.3	7.7:1
<=59	71.5	85.9	86.1	6.1:1
<=64	80.3	82.2	92.4	4.6:1
<=69	86.7	79.1	96.0	3.8:1
<=74	91.6	76.5	98.1	3.3:1
<=79	94.7	74.9	99.4	3.0:1
<=84	96.7	73.6	99.7	2.8:1
<=89	98.0	72.8	99.8	2.7:1
<=94	98.9	72.2	100.0	2.6:1
<=100	100.0	71.4	100.0	2.5:1

**Figures for
the Govt.-definition \$3.10/day 2011 PPP Poverty Line**

Figure 3 (Govt.-def. \$3.10/day 2011 PPP): 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	99.9
25-29	99.6
30-34	99.6
35-39	99.6
40-44	99.2
45-49	97.2
50-54	92.2
55-59	88.3
60-64	82.7
65-69	77.3
70-74	62.5
75-79	52.6
80-84	47.1
85-89	23.6
90-94	15.3
95-100	10.1

Figure 5 (Govt.-def. \$3.10/day 2011 PPP): For each score range, average differences between estimated and true poverty likelihoods for households, with confidence intervals, from 1,000 bootstraps of $n = 16,384$, 2010/11 scorecard applied to the 2010/11 validation sample

Score	Difference between estimate and observed value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	0.0	0.0	0.0	0.0
5-9	0.0	0.0	0.0	0.0
10-14	0.0	0.0	0.0	0.0
15-19	0.0	0.0	0.0	0.0
20-24	-0.1	0.0	0.0	0.0
25-29	-0.3	0.2	0.2	0.2
30-34	+1.5	0.7	0.8	1.0
35-39	+0.2	0.3	0.4	0.5
40-44	+2.7	1.0	1.2	1.6
45-49	+1.0	0.8	1.0	1.4
50-54	-1.7	1.4	1.5	1.7
55-59	-2.2	1.8	2.0	2.2
60-64	+4.9	2.0	2.4	3.4
65-69	+4.3	2.7	3.2	4.1
70-74	+3.1	4.0	4.7	6.3
75-79	-0.8	4.4	5.3	7.2
80-84	+6.7	5.5	6.6	8.8
85-89	-11.9	9.0	9.6	10.6
90-94	-12.5	9.6	10.4	11.9
95-100	+1.0	3.3	4.0	5.4

Figure 6 (Govt.-def. \$3.10/day 2011 PPP): Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and observed value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+0.5	57.4	65.1	86.4
4	+1.0	26.4	31.8	48.7
8	+0.8	19.0	23.3	31.2
16	+1.0	14.3	17.1	22.0
32	+1.1	10.1	11.8	15.0
64	+1.0	7.3	8.6	11.1
128	+1.0	5.1	6.2	7.8
256	+0.9	3.5	4.2	5.4
512	+0.9	2.5	2.9	4.1
1,024	+0.9	1.8	2.1	2.8
2,048	+0.9	1.3	1.6	2.2
4,096	+0.9	0.9	1.1	1.5
8,192	+0.9	0.7	0.8	1.0
16,384	+0.9	0.5	0.6	0.7

Figure 10 (Govt.-def. \$3.10/day 2011 PPP): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	Inclusion: Poor correctly targeted	Undercoverage: Poor mistakenly not targeted	Leakage: Non-poor mistakenly targeted	Exclusion: Non-poor correctly not targeted	Hit rate Inclusion + Exclusion	BPAC See text
<=4	0.1	87.3	0.0	12.5	12.7	-99.7
<=9	0.3	87.1	0.0	12.5	12.9	-99.2
<=14	1.3	86.2	0.0	12.5	13.8	-97.0
<=19	3.2	84.3	0.0	12.5	15.7	-92.7
<=24	7.2	80.2	0.0	12.5	19.8	-83.5
<=29	14.1	73.3	0.0	12.5	26.6	-67.7
<=34	22.4	65.0	0.2	12.4	34.8	-48.5
<=39	31.8	55.7	0.2	12.3	44.1	-27.0
<=44	41.9	45.5	0.5	12.1	54.0	-3.6
<=49	51.3	36.1	0.8	11.7	63.0	+18.3
<=54	60.1	27.4	1.5	11.1	71.1	+39.0
<=59	69.2	18.3	2.4	10.2	79.4	+60.9
<=64	76.1	11.4	4.2	8.3	84.4	+78.8
<=69	80.8	6.6	5.9	6.7	87.5	+91.5
<=74	83.9	3.6	7.7	4.8	88.7	+91.2
<=79	85.6	1.8	9.1	3.4	89.0	+89.6
<=84	86.6	0.9	10.1	2.4	89.0	+88.4
<=89	87.1	0.4	11.0	1.6	88.6	+87.5
<=94	87.3	0.1	11.6	0.9	88.3	+86.7
<=100	87.5	0.0	12.5	0.0	87.5	+85.7

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

Figure 11 (Govt.-def. \$3.10/day 2011 PPP): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
<=4	0.1	100.0	0.1	Only poor targeted
<=9	0.3	100.0	0.4	Only poor targeted
<=14	1.3	100.0	1.5	Only poor targeted
<=19	3.2	100.0	3.6	Only poor targeted
<=24	7.2	100.0	8.2	Only poor targeted
<=29	14.1	99.9	16.1	850.8:1
<=34	22.6	99.3	25.7	146.7:1
<=39	32.0	99.3	36.4	142.7:1
<=44	42.4	98.9	47.9	90.1:1
<=49	52.2	98.4	58.7	61.4:1
<=54	61.5	97.6	68.7	41.2:1
<=59	71.5	96.7	79.1	29.4:1
<=64	80.3	94.7	87.0	18.0:1
<=69	86.7	93.2	92.4	13.8:1
<=74	91.6	91.6	95.9	10.9:1
<=79	94.7	90.4	97.9	9.4:1
<=84	96.7	89.5	99.0	8.5:1
<=89	98.0	88.8	99.5	7.9:1
<=94	98.9	88.3	99.9	7.5:1
<=100	100.0	87.5	100.0	7.0:1

**Figures for
the PBM-definition \$1.90/day 2011 PPP Poverty Line**

Figure 3 (PBM-def. \$1.90/day 2011 PPP): Estimated poverty likelihoods associated with scores

If a household's score is then the likelihood (%) of being below the poverty line is:
0-4	100.0
5-9	99.7
10-14	98.4
15-19	97.1
20-24	94.0
25-29	88.5
30-34	88.1
35-39	84.1
40-44	78.0
45-49	72.0
50-54	56.5
55-59	50.9
60-64	46.6
65-69	38.5
70-74	27.5
75-79	17.8
80-84	13.4
85-89	6.7
90-94	4.3
95-100	0.0

Figure 5 (PBM-def. \$1.90/day 2011 PPP): For each score range, average differences between estimated and true poverty likelihoods for households, with confidence intervals, from 1,000 bootstraps of $n = 16,384$, 2010/11 scorecard applied to the 2010/11 validation sample

Score	Difference between estimate and observed value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	0.0	0.0	0.0	0.0
5-9	+10.8	9.0	11.5	13.9
10-14	-1.6	0.8	0.8	0.8
15-19	+0.4	1.8	2.1	2.6
20-24	+3.5	2.2	2.7	3.5
25-29	-2.0	1.8	2.0	2.6
30-34	+2.6	1.7	2.1	2.6
35-39	+1.1	1.8	2.1	2.9
40-44	+5.4	2.1	2.5	3.4
45-49	+2.7	2.2	2.5	3.5
50-54	-6.2	4.2	4.4	4.9
55-59	-7.1	4.7	4.9	5.2
60-64	+2.4	2.4	2.8	3.6
65-69	+8.3	2.7	3.1	3.9
70-74	+2.8	3.0	3.4	4.4
75-79	-13.1	8.7	9.3	10.5
80-84	+6.9	2.1	2.5	3.3
85-89	-0.9	2.8	3.3	4.3
90-94	-7.0	6.0	6.5	7.8
95-100	-0.5	0.6	0.7	0.8

Figure 6 (PBM-def. \$1.90/day 2011 PPP): Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and observed value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	-0.2	65.7	72.8	89.2
4	+0.5	37.3	43.8	58.2
8	+0.8	28.2	32.6	43.0
16	+0.7	19.9	23.7	31.8
32	+0.5	14.0	16.2	22.6
64	+0.5	10.0	12.0	15.4
128	+0.5	6.9	8.2	10.9
256	+0.5	4.8	5.5	7.3
512	+0.4	3.5	4.3	5.8
1,024	+0.5	2.5	3.0	3.8
2,048	+0.5	1.7	2.0	2.6
4,096	+0.5	1.2	1.4	1.9
8,192	+0.5	0.9	1.0	1.3
16,384	+0.5	0.6	0.7	0.9

Figure 10 (PBM-def. \$1.90/day 2011 PPP): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	Inclusion: Poor correctly targeted	Undercoverage: Poor mistakenly not targeted	Leakage: Non-poor mistakenly targeted	Exclusion: Non-poor correctly not targeted	Hit rate Inclusion + Exclusion	BPAC See text
<=4	0.1	62.3	0.0	37.6	37.7	-99.6
<=9	0.3	62.1	0.0	37.5	37.9	-99.0
<=14	1.3	61.2	0.0	37.5	38.8	-95.9
<=19	3.1	59.4	0.1	37.5	40.6	-90.0
<=24	6.8	55.7	0.4	37.1	43.9	-77.6
<=29	13.0	49.4	1.1	36.5	49.5	-56.5
<=34	20.4	42.1	2.2	35.4	55.7	-31.2
<=39	28.2	34.3	3.9	33.7	61.9	-3.6
<=44	35.9	26.5	6.5	31.1	66.9	+25.4
<=49	42.4	20.0	9.8	27.8	70.2	+51.5
<=54	48.3	14.1	13.2	24.3	72.6	+75.9
<=59	54.0	8.5	17.6	20.0	73.9	+71.8
<=64	57.8	4.6	22.5	15.1	72.9	+64.0
<=69	59.8	2.6	26.9	10.7	70.5	+57.0
<=74	61.1	1.4	30.5	7.0	68.1	+51.1
<=79	61.9	0.5	32.8	4.8	66.7	+47.5
<=84	62.1	0.3	34.6	3.0	65.2	+44.7
<=89	62.3	0.1	35.7	1.9	64.2	+42.8
<=94	62.4	0.0	36.5	1.0	63.5	+41.5
<=100	62.4	0.0	37.6	0.0	62.4	+39.8

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

Figure 11 (PBM-def. \$1.90/day 2011 PPP): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
<=4	0.1	100.0	0.2	Only poor targeted
<=9	0.3	93.8	0.5	15.1:1
<=14	1.3	98.4	2.1	61.9:1
<=19	3.2	97.1	4.9	33.7:1
<=24	7.2	94.0	10.9	15.6:1
<=29	14.1	92.3	20.9	12.1:1
<=34	22.6	90.2	32.6	9.2:1
<=39	32.0	87.9	45.1	7.3:1
<=44	42.4	84.6	57.5	5.5:1
<=49	52.2	81.3	67.9	4.3:1
<=54	61.5	78.5	77.3	3.7:1
<=59	71.5	75.4	86.4	3.1:1
<=64	80.3	72.0	92.6	2.6:1
<=69	86.7	69.0	95.8	2.2:1
<=74	91.6	66.7	97.8	2.0:1
<=79	94.7	65.4	99.2	1.9:1
<=84	96.7	64.3	99.5	1.8:1
<=89	98.0	63.6	99.8	1.7:1
<=94	98.9	63.1	100.0	1.7:1
<=100	100.0	62.4	100.0	1.7:1

**Figures for
the PBM-definition \$1.90/day 2011 PPP Poverty Line**

Figure 3 (PBM-def. \$3.10/day 2011 PPP): 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	99.6
25-29	98.1
30-34	97.1
35-39	96.9
40-44	94.4
45-49	92.6
50-54	86.9
55-59	83.2
60-64	78.6
65-69	66.8
70-74	57.0
75-79	46.5
80-84	43.8
85-89	30.4
90-94	18.9
95-100	12.0

Figure 5 (PBM-def. \$3.10/day 2011 PPP): For each score range, average differences between estimated and true poverty likelihoods for households, with confidence intervals, from 1,000 bootstraps of $n = 16,384$, 2010/11 scorecard applied to the 2010/11 validation sample

Score	Difference between estimate and observed value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	0.0	0.0	0.0	0.0
5-9	0.0	0.0	0.0	0.0
10-14	0.0	0.0	0.0	0.0
15-19	+2.3	1.6	1.9	2.5
20-24	+1.3	0.9	1.1	1.3
25-29	+1.0	1.0	1.2	1.6
30-34	-0.2	0.8	0.9	1.2
35-39	+0.9	0.8	1.0	1.3
40-44	+2.0	1.3	1.5	2.0
45-49	+0.5	1.1	1.3	1.8
50-54	0.0	1.6	1.9	2.5
55-59	+0.7	1.8	2.1	3.0
60-64	+9.3	2.3	2.7	3.7
65-69	-3.9	3.2	3.4	3.8
70-74	+6.1	3.8	4.4	6.0
75-79	-12.4	8.4	8.9	10.0
80-84	+8.5	5.3	6.3	8.0
85-89	-17.6	12.3	12.9	14.0
90-94	-31.8	19.6	20.4	22.1
95-100	+4.3	2.9	3.4	4.3

Figure 6 (PBM-def. \$3.10/day 2011 PPP): Average differences between estimated poverty rates and true values for a group at a point in time, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2010/11 scorecard applied to the 2010/11 validation sample

Sample Size <i>n</i>	Difference between estimate and observed value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+0.2	60.8	69.3	89.0
4	+0.7	30.5	37.1	51.3
8	+1.0	21.9	26.7	34.7
16	+1.2	16.5	19.5	26.6
32	+1.0	11.7	14.0	18.2
64	+0.9	8.0	9.5	13.4
128	+0.9	5.5	6.8	8.3
256	+0.9	3.9	4.6	6.0
512	+0.8	2.8	3.5	4.6
1,024	+0.8	2.1	2.5	3.2
2,048	+0.8	1.5	1.8	2.4
4,096	+0.8	1.1	1.3	1.8
8,192	+0.8	0.7	0.9	1.1
16,384	+0.8	0.5	0.6	0.8

Figure 10 (PBM-def. \$3.10/day 2011 PPP): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	Inclusion: Poor correctly targeted	Undercoverage: Poor mistakenly not targeted	Leakage: Non-poor mistakenly targeted	Exclusion: Non-poor correctly not targeted	Hit rate Inclusion + Exclusion	BPAC See text
<=4	0.1	83.1	0.0	16.8	16.9	-99.7
<=9	0.3	82.9	0.0	16.8	17.1	-99.2
<=14	1.3	81.9	0.0	16.8	18.1	-96.9
<=19	3.1	80.1	0.0	16.7	19.9	-92.4
<=24	7.1	76.1	0.1	16.7	23.8	-82.8
<=29	13.8	69.4	0.3	16.5	30.3	-66.4
<=34	22.1	61.1	0.5	16.3	38.4	-46.3
<=39	31.1	52.1	0.9	15.9	47.0	-24.1
<=44	40.8	42.4	1.6	15.2	55.9	0.0
<=49	49.7	33.5	2.5	14.3	64.0	+22.4
<=54	57.9	25.3	3.6	13.1	71.0	+43.5
<=59	66.2	17.0	5.3	11.4	77.6	+65.5
<=64	72.4	10.8	7.9	8.8	81.2	+83.5
<=69	76.8	6.5	9.9	6.9	83.6	+88.1
<=74	79.5	3.7	12.1	4.7	84.2	+85.4
<=79	81.3	2.0	13.5	3.3	84.6	+83.8
<=84	82.1	1.1	14.6	2.2	84.3	+82.5
<=89	82.7	0.5	15.3	1.5	84.1	+81.6
<=94	83.1	0.1	15.9	0.9	84.0	+80.9
<=100	83.2	0.0	16.8	0.0	83.2	+79.8

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

Figure 11 (PBM-def. \$3.10/day 2011 PPP): Share of all households who are targeted (that is, score at or below a cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), 2010/11 scorecard applied to the 2010/11 validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
<=4	0.1	100.0	0.2	Only poor targeted
<=9	0.3	100.0	0.4	Only poor targeted
<=14	1.3	100.0	1.6	Only poor targeted
<=19	3.2	98.8	3.8	82.6:1
<=24	7.2	98.4	8.5	62.5:1
<=29	14.1	98.0	16.6	48.1:1
<=34	22.6	97.9	26.6	45.7:1
<=39	32.0	97.2	37.4	34.4:1
<=44	42.4	96.2	49.0	25.1:1
<=49	52.2	95.3	59.7	20.2:1
<=54	61.5	94.1	69.5	15.9:1
<=59	71.5	92.5	79.5	12.4:1
<=64	80.3	90.1	87.0	9.1:1
<=69	86.7	88.6	92.2	7.8:1
<=74	91.6	86.8	95.6	6.6:1
<=79	94.7	85.8	97.6	6.0:1
<=84	96.7	84.9	98.7	5.6:1
<=89	98.0	84.4	99.4	5.4:1
<=94	98.9	84.0	99.8	5.2:1
<=100	100.0	83.2	100.0	5.0:1