Simple Poverty Scorecard® Poverty-Assessment Tool Tanzania

Mark Schreiner

27 June 2016

Hati hii na zana husika zimo kwa lugha ya Kiswahili kwenye `SimplePovertyScorecard.com`
This document is in English at `SimplePovertyScorecard.com`

Abstract

The Simple Poverty Scorecard-brand poverty-assessment tool uses ten low-cost indicators from Tanzania's 2011/12 Household Budget Survey to estimate the likelihood that a household has consumption below a given poverty line. Field workers can collect responses in about ten minutes. Accuracy is reported for a range of poverty lines. The scorecard is a practical way for pro-poor programs in Tanzania to measure poverty rates, to track changes in poverty rates over time, and to segment clients for targeted services.

Version note

This paper uses 2011/12 data, replacing Schreiner (2013a), which uses 2007 data. The new 2011/12 scorecard here should be used from now on. The two scorecards do *not* use the same definition of *poverty*, so there is no valid way for existing users of Schreiner (2013a) to measure change over time with a baseline from the old 2007 scorecard and a follow-up from the new 2011/12 scorecard.

Acknowledgements

This paper was funded by Grameen Foundation (GF). Data are from Tanzania's National Bureau of Statistics. Thanks go to Julie Peachey, Sharada Ramanathan, and Sasun Tsirunyan. This scorecard was re-branded by GF as a Progress out of Poverty Index® tool. The PPI® is a performance-management tool that GF promotes to help organizations achieve their social objectives more effectively. PPI learning materials and resources are at `progressoutofpoverty.org`. "Progress out of Poverty Index" and "PPI" are Registered Trademarks of Innovations for Poverty Action. "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

			Name	Identifier
Interview ID:				
Interview date:		Participant:		
Country:	TZA	Field agent:		
Scorecard:	002	Service point:		
Sampling wgt.:		Number of household members:		

Indicator	Response	Points	Score
1. How many household members are 18-years-old or younger?	A. Six or more	0	
	B. Five	2	
	C. Four	5	
	D. Three	11	
	E. Two	14	
	F. One	17	
	G. None	28	
2. Are all household members ages 6 to 18 currently in school?	A. No	0	
	B. Yes	3	
	C. No members ages 6 to 18	5	
3. What is the main building material used for the walls of the main building?	A. Baked bricks	0	
	B. Poles and mud, grass, sun-dried bricks, or other	6	
	C. Stones, cement bricks, or timber	13	
4. What is the main building material used for the roof of the main building?	A. Grass/leaves, mud and leaves, or other	0	
	B. Iron sheets, tiles, concrete, or asbestos	6	
5. What is the main fuel used for cooking?	A. Firewood, coal, solar, gas (biogas), wood/farm residuals, or animal residuals	0	
	B. Charcoal, paraffin, gas (industrial), electricity, generator/private source, or other	9	
6. Does your household have any televisions?	A. No	0	
	B. Yes	15	
7. Does your household have any radios, cassette/tape recorders, or hi-fi systems?	A. No	0	
	B. Yes	4	
8. Does your household have any lanterns?	A. No	0	
	B. Yes	4	
9. Does your household have any tables?	A. No	0	
	B. Yes	4	
10. If the household cultivated any crops in the last 12 months, does it currently own any bulls, cows, steers, heifers, male calves, female calves, or oxen?	A. No crops, and no cattle	0	
	B. No crops, and cattle	0	
	C. Crops, but no cattle	5	
	D. Crops, and cattle	12	

Score:

Back-page Worksheet: Household Membership, Age, and School Attendance

In the scorecard header, write the interview's unique identifier (if known), the interview date, and the sampling weight of the participant (if known). Then record the names and the unique identification numbers of the participant (who may differ from the respondent), of yourself as the field agent, and of the service point that the participant uses.

Ask the respondent: *Please tell me the name and age of each member of this household. A* household *is one person or a group of people—regardless of blood or marital relationship—who normally live together in the same residence, eat together, and pool resources. Please start with the head of the household.*

In the header by "Number of household members:", record the number of members (regardless of age).

In the table below, mark whether each household member is 18-years-old or younger. Count how many are 18-years-old or younger and mark the response to the first scorecard indicator.

Then mark whether each household member is ages 6 to 18. For each member in this age range, ask: "Is <name> currently in school?" and mark the corresponding response. Then check whether all members ages 6 to 18 are currently in school (or whether there are any members in this age range at all) and mark the response for the second scorecard indicator.

Keep in mind the full definition of *household* in the "Guidelines for the Interpretation of Scorecard Indicators".

Name	**Age**	**≤ 18?**	**≥ 6 and ≤ 18?**	**Is <name> currently in school?**
1.		No Yes	No Yes	Not 6 to 18 No Yes
2.		No Yes	No Yes	Not 6 to 18 No Yes
3.		No Yes	No Yes	Not 6 to 18 No Yes
4.		No Yes	No Yes	Not 6 to 18 No Yes
5.		No Yes	No Yes	Not 6 to 18 No Yes
6.		No Yes	No Yes	Not 6 to 18 No Yes
7.		No Yes	No Yes	Not 6 to 18 No Yes
8.		No Yes	No Yes	Not 6 to 18 No Yes
9.		No Yes	No Yes	Not 6 to 18 No Yes
10.		No Yes	No Yes	Not 6 to 18 No Yes
11.		No Yes	No Yes	Not 6 to 18 No Yes
12.		No Yes	No Yes	Not 6 to 18 No Yes
13.		No Yes	No Yes	Not 6 to 18 No Yes
# HH members:		**#≤18:**	—	**#No: #Yes:**

Look-up table to convert scores to poverty likelihoods: National poverty lines

	Poverty likelihood (%)			
	National poverty lines			
Score	**Food**	**100%**	**150%**	**200%**
0–4	100.0	100.0	100.0	100.0
5–9	39.5	100.0	100.0	100.0
10–14	36.6	82.3	93.1	100.0
15–19	29.9	62.1	89.2	98.6
20–24	21.0	51.2	84.5	93.8
25–29	13.3	40.3	77.4	93.8
30–34	10.4	32.9	68.0	87.4
35–39	4.4	20.2	58.1	79.4
40–44	2.8	14.0	42.5	67.9
45–49	1.5	10.9	40.2	63.7
50–54	1.3	6.6	29.2	51.2
55–59	0.6	4.1	24.2	43.8
60–64	0.6	2.2	13.5	31.8
65–69	0.4	1.3	8.6	28.1
70–74	0.0	1.0	5.9	19.5
75–79	0.0	1.0	5.9	16.8
80–84	0.0	1.0	2.6	7.3
85–89	0.0	0.0	2.2	7.3
90–94	0.0	0.0	0.0	7.3
95–100	0.0	0.0	0.0	0.0

Look-up table to convert scores to poverty likelihoods: International 2005 and 2011 PPP poverty lines

	Poverty likelihood (%)							
	2005 PPP poverty lines				2011 PPP poverty lines			
Score	**$1.25**	**$2.00**	**$2.50**	**$5.00**	**$1.90**	**$3.10**	**$3.80**	**$4.00**
0–4	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
5–9	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
10–14	86.7	100.0	100.0	100.0	89.0	100.0	100.0	100.0
15–19	77.5	99.9	99.9	100.0	85.0	99.9	99.9	99.9
20–24	71.9	93.0	97.0	99.7	78.0	96.3	97.4	98.3
25–29	59.6	92.4	96.7	99.5	70.6	94.5	97.4	98.0
30–34	48.0	84.3	95.0	99.5	57.7	91.7	96.7	97.7
35–39	35.2	77.2	89.5	99.5	47.2	84.4	92.5	94.1
40–44	23.5	64.3	79.6	98.5	31.2	71.8	85.4	88.9
45–49	20.2	54.3	75.7	97.3	28.5	66.1	81.8	84.1
50–54	13.8	45.4	61.2	94.2	18.8	54.0	69.7	72.2
55–59	7.3	38.7	56.9	93.4	11.9	45.5	63.1	65.3
60–64	4.2	25.1	45.6	86.1	5.8	33.7	50.2	53.8
65–69	1.7	21.3	36.8	79.9	3.7	28.9	44.7	46.8
70–74	1.7	13.8	27.1	78.7	3.1	21.2	34.4	37.2
75–79	1.7	12.8	23.2	71.3	2.9	20.1	33.0	35.7
80–84	1.3	3.1	8.1	44.2	1.3	6.3	11.4	13.4
85–89	0.1	2.8	8.1	38.3	1.2	6.3	11.4	13.4
90–94	0.0	0.0	7.6	30.6	0.0	6.3	7.6	7.6
95–100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Look-up table to convert scores to poverty likelihoods: Relative and percentile-based poverty lines

	Poverty likelihood (%)					
	Poorest half	Percentile-based lines				
Score	<100% Natl.	20th	40th	50th	60th	80th
0–4	100.0	100.0	100.0	100.0	100.0	100.0
5–9	80.6	100.0	100.0	100.0	100.0	100.0
10–14	49.2	76.6	89.0	90.1	98.6	100.0
15–19	42.6	53.2	78.8	88.1	95.0	99.9
20–24	25.6	38.3	71.8	80.1	87.6	97.1
25–29	20.6	28.0	58.0	73.7	83.9	96.9
30–34	16.9	23.2	48.4	60.8	74.2	95.4
35–39	8.1	12.5	36.3	48.9	62.4	88.0
40–44	5.4	7.9	23.2	33.0	45.4	80.4
45–49	3.1	5.4	17.9	27.4	39.4	72.4
50–54	1.7	3.4	12.1	17.6	26.8	54.6
55–59	0.5	0.6	6.0	8.6	18.8	46.4
60–64	0.5	0.6	1.8	4.4	10.3	33.0
65–69	0.3	0.4	1.1	1.3	5.5	22.7
70–74	0.0	0.0	1.1	1.3	2.6	13.9
75–79	0.0	0.0	1.0	1.3	2.6	12.8
80–84	0.0	0.0	0.0	1.3	1.4	5.3
85–89	0.0	0.0	0.0	0.1	0.1	5.3
90–94	0.0	0.0	0.0	0.0	0.0	0.0
95–100	0.0	0.0	0.0	0.0	0.0	0.0

Simple Poverty Scorecard® Poverty-Assessment Tool Tanzania

1. Introduction

Pro-poor programs in Tanzania can use the Simple Poverty Scorecard poverty-assessment tool to estimate the likelihood that a household has consumption below a given poverty line, to estimate a population's poverty rate at a point in time, to track changes in a population's poverty rate over time, and to segment participants for differentiated services.

The direct approach to poverty measurement via consumption surveys is difficult and costly. Tanzania's 2011/12 Household Budget Survey (HBS) is an unusually lengthy case in point. It has 68 pages and includes about 700 questions, many of which may be asked multiple times (for example, for each household member, each crop, or each consumption item). In addition, all members 5-years-old or older of an interviewed household kept an individual consumption diary, and enumerators visited households "for 28 days on a regular basis" (every two or three days for households with a literate member, and daily for other households).[1]

In comparison, the indirect approach of the scorecard is quick and low-cost. It uses 10 verifiable indicators (such as "What is the main building material for the roof of

[1] NBS (2014, p. xiii).

the main dwelling?" and "Does the household have any tables?") to get a score that is correlated with poverty status as measured by the exhaustive HBS 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 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, Tanzania's national line). USAID microenterprise partners in Tanzania can use scoring with the $1.25/day 2005 PPP poverty 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-

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

[3] USAID defines a household as *very poor* if its daily per-capita consumption is less than the highest of the $1.25/day 2005 PPP line (TZS1,139, Table 1) or the line (TZS693) that marks the poorest half of people below 100% of the national line. USAID (2014, p. 8) has approved the scorecard—when re-branded as a Progress Out of Poverty Index®—for use by its microenterprise partners.

poor organizations may be able to implement a low-cost scorecard to help with monitoring poverty and (if desired) segmenting clients for differentiated services.

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 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 approaches are usually about as accurate as complex, opaque ones (Schreiner, 2012a; Caire and Schreiner, 2012).

Beyond its low cost and transparency, the technical approach of the scorecard 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 poverty measurement via scorecards.

The scorecard is based on data from the 2011/12 HBS by Tanzania's National Bureau of Statistics (NBS). 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 in Tanzania

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-adult-equivalent or per-capita consumption below a given poverty line.

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

Third, the scorecard can estimate the annual rate of change in the poverty rate. With two independent samples from the same population, this is the difference in the average poverty likelihood in the baseline sample versus the average likelihood in the follow-up sample, divided by the difference (in years) between the average interview date in the baseline sample and the average interview date in the follow-up sample. With one sample in which each household is scored twice, the estimate is the sum of the changes in each household's poverty likelihood from baseline to follow-up, divided by

the sum of the years that passed between each household's two interviews (Schreiner, 2014a).

The scorecard can also be used to segment participants for differentiated services. To help managers choose appropriate targeting cut-offs for their purposes, several measures of targeting accuracy are reported for a range of possible cut-offs.

This paper presents a single scorecard whose indicators and points are derived with Tanzania's national poverty line applied to data from the 2011/12 HBS. Scores from this one scorecard are calibrated with data from the 2011/12 HBS to poverty likelihoods for 18 poverty lines.

The scorecard is constructed using half of the data from the 2011/12 HBS. That same half of the 2011/12 data is also used to calibrate scores to poverty likelihoods for 18 poverty lines. The other half of the data from the 2011/12 HBS is used to validate the scorecard's accuracy for estimating households' poverty likelihoods, for estimating populations' poverty rates at a point in time, and for segmenting participants.

All three scoring-based estimators (the poverty likelihood of a household, the poverty rate of a population at a point in time, and the annual rate of change in a population's poverty rate) 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 is constructed from a single sample and so misses the mark to some unknown extent when applied (as in this paper)

to a validation sample. Furthermore, it makes errors when applied (in practice) to a different population or when applied before or after 2011/12 (because the relationships between indicators and poverty change over time).[4]

Thus, while the indirect scoring approach is less costly than the direct survey approach, it makes errors when applied in practice. (Estimates from the direct survey approach are correct by definition.) There are errors because scoring necessarily assumes that future relationships between indicators and poverty in all populations 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 validation sample, the error (difference between scorecard estimates of groups' poverty rates versus the true rates) at a point in time for 100% of the national poverty line is 0.0 percentage points. Across all 18 poverty lines, the average absolute error is about 0.5 percentage points, and the maximum average absolute error is 1.8 percentage points. These estimation errors are due to sampling variation, not bias; the average difference would be zero if the whole 2011/12 HBS survey were to be repeatedly re-fielded and divided into sub-samples before repeating the entire process of constructing and validating scorecards.

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

With $n = 16{,}384$, the 90-percent confidence intervals are ± 0.7 percentage points or less. For $n = 1{,}024$, the 90-percent intervals are ± 2.6 percentage points or less.

Section 2 below documents data and poverty lines. Sections 3 and 4 describe scorecard construction and offer guidelines for implementation. Sections 5 and 6 tell how to estimate households' poverty likelihoods and populations' 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 Tanzania. The last section is a summary.

The "Guidelines for the Interpretation of Scorecard Indicators" tells how to ask questions (and how to interpret responses) so as to mimic practice in Tanzania's 2011/12 HBS as closely as possible. These "Guidelines" (and the "Back-page Worksheet") are integral parts of the Simple Poverty Scorecard tool.

2. Data and poverty lines

This section presents the data used to construct and validate the scorecard. It also documents the 18 definitions of *poverty* to which scores are calibrated.

2.1 Data

Indicators and points for the scorecard are selected (*constructed*) based on a random half of the data from the 10,186 households in the 2011/12 HBS, Tanzania's most-recent available national consumption survey.

The half of the 2011/12 data that is used to construct the scorecard is also used to associate (*calibrate*) scores to poverty likelihoods for all poverty lines.

The other half of the 2011/12 data is used to test (*validate*) scorecard accuracy *out-of-sample*, that is, with data that is not used in construction/calibration.

Interviews for the 2011/12 HBS took place from 1 October 2011 to 12 October 2012. Consumption is in units of TZS per adult equivalent or per person per day in average prices for Tanzania during the 2011/12 HBS fieldwork.

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 adult equivalents or 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.[5] By assumption, each household member has the same poverty status (or estimated poverty likelihood) as the other members in that same household.

To illustrate, suppose a program serves two households. The first household is poor (its per-adult-equivalent or 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 weighted[6] average of poverty statuses (or estimated poverty likelihoods) across households with participants. This is $\frac{1 \cdot 1 + 1 \cdot 0}{1 + 1} = \frac{1}{2} = 0.5 = 50$ percent. In the "$1 \cdot 1$" 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 "$1 \cdot 0$" term in the numerator, the "1" is the second household's weight,

[5] Even though consumption is sometimes measured in per-adult-equivalent units, the unit of analysis for estimates of poverty rates is always households or people.
[6] 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).

and the "0" is the second household's poverty status (non-poor). The "$1 + 1$" in the 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-weighted[7] 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 "$3 \cdot 1$" 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 "$4 \cdot 0$" 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 who directly participate in the program. For the example here, this means that some—but not all—household members are counted. The person-level rate is now the participant-weighted average[8] of the poverty statuses of households with

[7] Given simple random sampling at the household level, a household's person-level weight is the number of people in the household.

[8] Given simple random sampling at the household level, a household's participant-level weight is the number of participants in the household.

participants, or $\frac{1 \cdot 1 + 2 \cdot 0}{1 + 2} = \frac{1}{3} = 0.33 = 33$ percent. The first "1" in the "$1 \cdot 1$" in the numerator is the 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—whether household, household member, or participant—and explain why that unit is relevant.

Table 1 reports poverty lines and poverty rates for households and people in the 2011/12 HBS for Tanzania as a whole, for Tanzania's three poverty-line regions (Dar es Salaam, Other Urban, and Rural), for the construction/calibration sample, and for the validation sample.

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 and because sampling is almost always done at the level of households. This is also why the scorecard is constructed, calibrated, and validated with household weights. Person-level poverty rates are also included in Table 1 because these are the

rates reported by the government of Tanzania. Furthermore, popular discussions and policy discourse usually proceed in terms of person-level rates, and the goal of local pro-poor programs is to help people (not households) to improve their well-being.

2.3 Definition of *poverty*, and the national poverty line

A household's *poverty status* as poor or non-poor depends on whether its per-adult-equivalent or per-capita consumption is below a given poverty line. Thus, a definition of *poverty* is the combination of a poverty line along with a measure of consumption.

NBS (2014) documents the derivation of Tanzania's national poverty lines. Following the cost-of-basic-needs approach of Ravallion (1998), the food line is the cost of a basket of 153 food items, with each item's share derived from the observed consumption of people in the 10^{th} to 50^{th} percentiles of per-adult-equivalent total (food-plus-non-food) consumption in the 2011/12 HBS. The basket is scaled to provide 2,200 Calories. The prices of the food items are adjusted to units of average prices for Tanzania as a whole during the 2011/12 HBS fieldwork by adjusting for price changes across the four quarters of fieldwork and for price differences in urban and rural areas in the three poverty-line regions. For Tanzania overall in the 2011/12 HBS, the food line is TZS852 per adult equivalent per day, giving a household-level poverty rate of 7.2 percent and a person-level poverty rate of 9.7 percent (Table 1).

The national ("basic-needs") poverty line (usually called here "100% of the national line") is this food line, plus a non-food component that is defined as the average non-food consumption of households whose total consumption is between the food line and 120 percent of the food line (World Bank, 2015). Like the food component, the non-food component is adjusted for price changes during fieldwork and across the three poverty-line regions. This leads to an all-Tanzania national (food-plus-non-food) poverty line of TZS1,191 per adult equivalent per day (Table 1), with a household-level poverty rate of 21.5 percent and a person-level poverty rate of 28.2 percent.[9]

[9] These poverty rates match those in NBS (2014, pp. 96–97) and so provide some reassurance that this paper uses the same data as NBS and that the price adjustments in the poverty lines are correct.

2.4 Supported poverty lines

Because pro-poor organizations in Tanzania may want to use different or various poverty lines, this paper calibrates scores from its single scorecard to poverty likelihoods for 18 lines:

- Food
- 100% of national
- 150% of national
- 200% of national
- $1.25/day 2005 PPP
- $2.00/day 2005 PPP
- $2.50/day 2005 PPP
- $5.00/day 2005 PPP
- $1.90/day 2011 PPP
- $3.10/day 2011 PPP
- $3.80/day 2011 PPP
- $4.00/day 2011 PPP
- Line marking the poorest half of people below 100% of the national line
- First-quintile (20^{th}-percentile) line
- Second-quintile (40^{th}-percentile) line
- Median (50^{th}-percentile) line
- Third-quintile (60^{th}-percentile) line
- Fourth-quintile (80^{th}-percentile) line

The lines for 150% and 200% of the national line are multiples of 100% of the national line.

The international 2005 and 2011 PPP lines are derived from:

- PPP exchange rate for Tanzania for "individual consumption expenditure by households":
  - 2005 (World Bank, 2008): TZS482.451 per $1.00
  - 2011:[10] TZS585.520 per $1.00
- Average Consumer Price Index (CPI) for all of Tanzania:[11,12]
  - 2005 calendar-year: 120.8749
  - 2011 calendar-year: 201.4093
  - 2011/12 HBS fieldwork: 228.3548
- 100% of the national line in the 2011/12 HBS' three poverty-line regions
- Person-weighted average of 100% of the national line for all-Tanzania: TZS1,191.04 (Table 1)

Given this, the average $1.25/day 2005 PPP line in average prices in Tanzania during the 2011/12 HBS fieldwork is (Sillers, 2006):

$$\$1.25 \cdot \left(\frac{\text{2005 PPP factor}}{\$1.00}\right) \cdot \left(\frac{\text{CPI}_{\text{HBS}}}{\text{CPI}_{2005}}\right) = \$1.25 \cdot \left(\frac{\text{TZS482.451}}{\$1.00}\right) \cdot \left(\frac{228.3548}{120.8749}\right) = \text{TZS1,139.}$$

The other 2005 PPP lines are multiples of the $1.25/day line.

The 2005 PPP lines in the top row of Table 1 are all-Tanzania averages. In a given poverty-line region, the $1.25/day line is the all-Tanzania $1.25/day line,

[10] `iresearch.worldbank.org/PovcalNet`, retrieved 3 June 2016.
[11] Based on a spliced CPI series with base = 100 (January 2002) with data from: `nbs.go.tz/index.php?option=com_phocadownload&view=category&id=138:summary-cpi&download=411:cpi-summary-2002-2010-releases&Itemid=106` (retrieved 28 January 2011); `nbs.go.tz/nbs/takwimu/cpi/CPIApril2011.pdf` (retrieved 18 February 2016); `nbs.go.tz/nbs/takwimu/cpi/CPI_December2011_Eng.pdf` (retrieved 18 February 2016); `nbs.go.tz/nbs/takwimu/cpi/CPI_Jan2013_Eng.pdf` (retrieved 18 February 2016).
[12] Inflation implied by data in the 2007 and 2011/12 HBS is much higher than that from Tanzania's official CPI (World Bank, 2015; Arndt *et al.*, 2013 and 2015). Using HBS-based deflators to adjust the 2005 PPP factors would lead to higher lines and thus higher poverty rates. This issue matters less for the 2011 PPP factors.

multiplied by 100% of the national line for that region, and divided by 100% of the national line for Tanzania as a whole.

For example, the regional $1.25/day 2005 PPP line for Rural is the all-Tanzania $1.25/day line (TZS1,139), multiplied by 100% of the national line in Rural (TZS1,146, Table 1), and divided by 100% of the national line for Tanzania as a whole (TZS1,191). This is 1,139 x 1,146 ÷ 1,191 = TZS1,096.

The World Bank's PovcalNet reports a person-level poverty rate for $1.25/day 2005 PPP of 43.5 percent (versus the 40.7 percent here).[13] The $1.25/day estimates here are to be preferred (Schreiner, 2014b) because PovcalNet does not report:

- Its $1.25/day 2005 PPP line in TZS
- The time/place of its price units
- Whether/how it adjusts for regional differences in prices
- How it deflates 2005 PPP factors over time

The average $1.90/day 2011 PPP line in average prices in Tanzania during the 2011/12 HBS fieldwork is:

$$\$1.90 \cdot \left(\frac{\text{2011 PPP factor}}{\$1.00}\right) \cdot \left(\frac{\text{CPI}_{\text{HBS}}}{\text{CPI}_{2011}}\right) = \$1.90 \cdot \left(\frac{\text{TZS585.520}}{\$1.00}\right) \cdot \left(\frac{228.3548}{201.4093}\right) = \text{TZS1,261.}$$

The other 2011 PPP lines are multiples of the $1.90/day line.

[13] iresearch.worldbank.org/PovcalNetPPP2005/Detail.aspx?Format=Detail&C0=TZA_3&PPP0=482.45&PL0=1.25&Y0=2011.8&NumOfCountries=1, retrieved 6 June 2016.

As for the 2005 PPP lines, the 2011 PPP lines are adjusted for price differences across Tanzania's three poverty-line regions.

The World Bank's PovcalNet reports a $1.90/day 2005 PPP line for 2011/12 of TZS1,211 (versus TZS1,261 here) and person-level poverty rate of 46.6 percent (versus 49.0 percent here).[14] As noted above, this paper's derivation of the $1.90/day is to be preferred because it makes sense and is fully documented.

The line that marks the poorest half of people below 100% of the national line is defined as the median aggregate household per-capita consumption of people (not households) below 100% of the national line (U.S. Congress, 2004). Unlike all previous (non-relative) lines, this line (and the percentile-based lines below) is derived by putting all regional price adjustments in the measure of consumption rather than in the poverty line, deriving a single line for all of Tanzania, and then taking all price adjustments out of consumption and putting them back in the regional lines.[15]

[14] `iresearch.worldbank.org/PovcalNet/Detail.aspx?Format=Detail&C0=TZA_3&PPP0=585.52&PL0=1.90&Y0=2011.77&NumOfCountries=1`, retrieved 3 June 2016.

[15] This corrects how the scorecard derived this line prior to 2016 (in particular, in Schreiner 2013a). Formerly, price adjustments were left in the poverty line and compared with nominal consumption to find a line in each poverty-line region that marked the poorest half of people below 100% of the national line in that particular poverty-line region. Both approaches produce a person-level poverty rate that is half that of 100% of the national line, but the set of people who are identified as *poor* differs. Unlike the former approach, the current approach correctly identifies as *poor* the poorest half of all people in the country whose price-adjusted consumption is below the single, all-country national line. This implies that the correction in Schreiner (2014b) of the derivation used for this line by IRIS Center for its Poverty-Assessment Tool is itself

Microenterprise programs in Tanzania who use the scorecard to report the number of their participants who are "very poor" to USAID should use the $1.25/day 2005 PPP line. This is because USAID defines the "very poor" as those people in households whose daily per-capita consumption is below the highest of the following two poverty lines:

- The line that marks the poorest half of people below 100% of the national line (TZS693, with a person-level poverty rate of 14.1 percent, Table 1)
- $1.25/day 2005 PPP (TZS1,139, with a person-level poverty rate of 40.7 percent)

The scorecard also supports percentile-based poverty lines. This facilitates a number of types of analyses. For example, the second-quintile (40^{th}-percentile) line might be used to help track Tanzania's progress toward the World Bank's (2013) goal of "shared prosperity/inclusive economic growth", defined as income growth among the bottom 40 percent of the world's people.

The five quintile lines, analyzed together, could also be used to look at the relationship of consumption with health outcomes (or anything else related with the distribution of consumption). The scorecard thus offers an alternative for health-equity analyses that have typically used a "wealth index" such as that supplied with the data from the Demographic and Health Surveys (Rutstein and Johnson, 2004) to compare some aggregation of wealth with health outcomes.

---

wrong, and IRIS Center's approach (the one now used here) is correct (although IRIS Center still incorrectly derives this line based on households instead of people).

Of course, analysts could always do (and still can do) relative-wealth analyses with scores from the scorecard. But support for relative consumption lines now allows a more straightforward use of a single tool (the scorecard) to analyze any or all of:

- Relative wealth (via scores)
- Absolute consumption (via poverty likelihoods and absolute poverty lines)
- Relative consumption (via poverty likelihoods and percentile-based poverty lines)

Unlike the scorecard, wealth indexes only serve to analyze relative wealth. Furthermore, the scorecard—unlike wealth indexes based on Principal Component Analysis or similar approaches—is tied to a straightforward, well-understood standard whose definition is external to the scorecard itself (consumption related to a poverty line defined in monetary terms).

In contrast, a wealth index opaquely defines *poverty* in terms of its own indicators and points, without reference to an external standard. This means that two wealth indexes with different indicators or different points—even if derived from the same data for a given country—imply two different definitions of *poverty*. In the same set-up, two scorecards would both apply a single definition of *poverty*.

2.5 Non-comparability of estimates from the old 2007 scorecard and the new 2011/12 scorecard

Unfortunately, estimates from the new 2011/12 scorecard here are not comparable with estimates from the old 2007 scorecard in Schreiner (2013a). NBS (2014) does not compare poverty-rate estimates across the 2007 and 2011/12 HBS surveys, and it repeatedly warns against making such comparisons. World Bank (2015) details the differences between the two HBS rounds.

Thus, it is not possible for existing "legacy" users of the old 2007 scorecard to switch to the new 2011/12 scorecard and still estimate changes in poverty rates over time with a baseline from the old 2007 scorecard and a follow-up from the new 2011/12 scorecard. Estimates are non-comparable even for poverty lines (such as \$1.25/day and \$2.50/day 2005 PPP) that are supported for both scorecards because the each of the two scorecards support a different definition of these lines. That is, the \$1.25/day 2005 PPP line supported for the old 2007 scorecard in Schreiner (2013a) is defined differently than is the \$1.25/day 2005 PPP line supported for the new 2011/12 scorecard here.

The non-comparability arises because between 2007 and 2011/12, the NBS:

- Made major improvements in how the HBS measures consumption
- Changed some details in the derivation of the national poverty lines
- Overhauled the non-consumption sections of the questionnaire, changing the wording or response options of many indicators and also how many indicators were collected

These changes mean that it is not possible in Tanzania to use the techniques used in elsewhere to enable comparisons of estimates from different scorecards when the definition of *poverty* has changed. In particular, it is not possible to apply the NBS' old 2007 method for measuring consumption with the 2011/12 HBS. Likewise, it is not possible to apply the NBS' new 2011/12 method for measuring consumption with the 2007 NBS. And even if it were possible, the changes in the non-consumption sections of the questionnaire render invalid the required "identical indicators" assumption (Schreiner, 2013d).

This is not a desired situation, but there is nothing that can be done about it.

3. Scorecard construction

For Tanzania, about 90 candidate indicators are initially prepared in the areas of:

- Household composition (such as the number of members)
- Education (such as the literacy of the female head/spouse)
- Housing (such as the main material of the roof)
- Ownership of durable assets (such as televisions or tables)
- Employment (such as the number of household members who work)
- Agriculture (such as the ownership of cattle)

Table 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.[16]

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 radio is probably more likely to change in response to changes in poverty than is the literacy of the female head/spouse.

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

[16] The uncertainty coefficient is not used to help select scorecard indicators; it is just a way to order the candidate indicators listed in Table 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 statistical accuracy with the non-statistical criteria. These steps are repeated until the scorecard has 10 indicators that work well together.[17]

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).

[17] For Tanzania, indicator selection was also informed by feedback from user review of a set of draft indicators.

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 statistical[18] and non-statistical criteria. The use of non-statistical criteria can improve robustness through time and helps ensure that indicators are simple, common-sense, and acceptable to users.

The single scorecard here applies to all of Tanzania. 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. 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).

[18] 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.

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, Tanzania'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)

The scorecard (and its "Back-page Worksheet") is ready to be photocopied. A field worker using the scorecard in Tanzania would:

- Record the interview identifier, interview date, country code ("TZA"), scorecard code ("002") and the sampling weight assigned by the organization's 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), the field agent, and the relevant organizational service point
- Complete the "Back-page Worksheet" with each household member's first name, age, and school-attendance status
- Based on what has already been recorded on the "Back-page Worksheet", record household size (the number of household members, regardless of age) in the scorecard header next to "Number of household members:"
- Based on what has already been recorded on the "Back-page Worksheet", mark the response to the first scorecard indicator based on the number of household members who are 18-years-old or younger
- Based on what has already been recorded on the "Back-page Worksheet", mark the response to the second scorecard indicator based on whether all household members ages 6 to 18 are currently in school
- Read the rest of the scorecard indicators to the respondent one-by-one, 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).[19] IRIS Center

[19] 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

(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 Scorecard Indicators" found at the end of this paper, as these "Guidelines"—along with the "Back-page Worksheet"—are integral parts of the Simple Poverty Scorecard tool.[20]

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 a 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 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

---

office. Even if points are hidden, however, 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.

[20] 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 Tanzania's NBS did in the HBS.

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 organizations who use scoring for targeting in Tanzania.

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

- Who will do the interviews
- How responses and scores will be recorded
- What participants will be interviewed
- How many participants will be interviewed
- How frequently participants will be interviewed
- 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 and/or in a representative sample of relevant field agents
- A representative sample of relevant participants in a representative sample of relevant field offices and/or in a representative sample of relevant field agents

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. To have a chance to meaningfully inform questions that matter to the organization, however, the focus 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 that is relevant for an issue that matters to the program.

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 a Simple Poverty Scorecard tool for Bangaldesh (Schreiner, 2013b) 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 a household's poverty likelihood

The sum of scorecard points for a household is called the *score*. For Tanzania, 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 are converted to *poverty likelihoods*, that is, probabilities of being below a poverty line. This is done via simple look-up tables. For the example of 100% of the national line, scores of 25–29 have a poverty likelihood of 40.3 percent, and scores of 30–34 have a poverty likelihood of 32.9 percent (Table 3).

The poverty likelihood associated with a score varies by poverty line. For example, scores of 25–29 are associated with a poverty likelihood of 40.3 percent for 100% of the national line but 70.6 percent for the $1.90/day 2011 PPP line.[21]

[21] From Table 3 on, many tables have 18 versions, one for each of the 18 poverty lines. To keep them straight, they are grouped by line. Single tables pertaining to all lines appear with the first group of tables for 100% of the 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-adult-equivalent or per-capita consumption below a given poverty line.

For the example of 100% of the national line (Table 4), there are 10,104 (normalized) households in the calibration sub-sample with a score of 25–29. Of these, 4,076 (normalized) are below the poverty line. The estimated poverty likelihood associated with a score of 25–29 is then 40.3 percent, because 4,076 ÷ 10,104 = 40.3 percent.

To illustrate with 100% of the national line and a score of 30–34, there are 11,924 (normalized) households in the calibration sub-sample, of whom 3,929 (normalized) are below the line (Table 4). The poverty likelihood for this score range is then 3,929 ÷ 11,924 = 32.9 percent.

The same method is used to calibrate scores with estimated poverty likelihoods for all 18 poverty lines.[22]

[22] 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.

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 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 their objectivity depends on using data in score calibration, not on using data (and nothing else) in scorecard construction.

Although the points in the Tanzania 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.[23]

Of course, the relationships between indicators and poverty do change to some unknown extent over time and also across sub-national groups in Tanzania's population. Thus, the scorecard will generally be biased when applied after October 2012 (the last month of fieldwork for the 2011/12 HBS) or when applied with sub-groups that are not nationally representative.

[23] This is 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 Tanzania as a whole? To find out, the scorecard is applied to 1,000 bootstrap samples of size $n = 16{,}384$ with the validation sample. Bootstrapping means to:

- Score each household in the 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 (Table 3) 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 $n = 16{,}384$, Table 5 shows the average difference between estimated and true poverty likelihoods. It also shows confidence intervals for the differences.

For the 100% of the national line, the average poverty likelihood across bootstrap samples for scores of 25–29 in the validation sample is too low by 4.4 percentage points. For scores of 30–34, the estimate is too high by 5.9 percentage points.[24]

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

The 90-percent confidence interval for the differences for scores of 25–29 is ±3.3 percentage points (Table 5). This means that in 900 of 1,000 bootstraps, the average difference between the estimate and the true value for households in this score range is between –7.7 and –1.1 percentage points (because –4.4 – 3.3 = –7.7, and –4.4 + 3.3 = –1.1). In 950 of 1,000 bootstraps (95 percent), the difference is –4.4 ± 3.5 percentage points, and in 990 of 1,000 bootstraps (99 percent), the difference is –4.4 ± 4.1 percentage points.

A few of the absolute differences between estimated poverty likelihoods and true values in Table 5 for 100% of the national line 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 Tanzania's population (and also because there are very few households in some score ranges). For targeting, however, what matters is less the difference in all score ranges and more the difference 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). Section 8 below looks at targeting accuracy in detail.

In addition, if estimates of groups' poverty rates are to be usefully accurate, then errors for individual households' poverty likelihoods must largely balance out. As discussed in the next section, this is generally the case for nationally representative samples, although it holds less well for samples from sub-national populations.

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 HBS fieldwork in October 2012. That is, the scorecard may fit the data from 2011/12 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 2011/12 HBS data but not in the overall population of Tanzania. 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 poverty rates for nationally representative samples (see the next two sections). Furthermore, at least some of the differences in change-over-time estimates 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 price 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 poverty rate at a point in time

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

To illustrate, suppose a program samples three households on 1 January 2016 and that they have scores of 20, 30, and 40, corresponding to poverty likelihoods of 51.2, 32.9, and 14.0 percent (100% of the national line, Table 3). The group's estimated poverty rate is the households' average poverty likelihood of $(51.2 + 32.9 + 14.0) \div 3 =$ 32.7 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 32.9 percent. This differs from the 32.7 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. There are a few contexts in which the analysis of scores is appropriate, but, in general, the safest rule to follow is: If you are not completely sure what to do, then use poverty likelihoods, not scores.

Scores from the scorecard for Tanzania are calibrated with data from the 2011/12 HBS for all 18 poverty lines. The process of calibrating scores to poverty likelihoods and the approach to estimating poverty rates is exactly the same for all poverty lines. For users, the only difference in what is they do with one poverty line versus with another is the specific look-up table used to convert scores to poverty likelihoods.

6.1 Accuracy of estimated poverty rates at a point in time

For the Tanzania scorecard applied to 1,000 bootstraps of $n = 16{,}384$ from the validation sample and 100% of the national poverty line, the average error (difference between the estimate and the true value) for a poverty rate at a point in time is 0.0 percentage points (Table 7, summarizing Table 6 across all poverty lines). Across all 18 poverty lines in the validation sample, the maximum average absolute error is 1.8 percentage points, and the average absolute error is about 0.5 percentage points. At least part of these differences is due to sampling variation in the division of the 2011/12 HBS into sub-samples.

When estimating poverty rates at a point in time for a given poverty line, the average error reported in Table 7 should be subtracted from the average poverty likelihood to give a corrected estimate. For the example of the scorecard here and 100% of the national line in the validation sample, the error happens to be 0.0 percentage points, so the corrected estimate in the three-household example above is the same as

the uncorrected estimate, 32.7 – (0.0) = 32.7 percent. In general—and in 15 of the 18 poverty lines in Table 7—the error is not zero and so the corrected and uncorrected estimates are not the same (unlike in the example above).

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 ±0.7 percentage points or better for all poverty lines (Table 7). This means that in 900 of 1,000 bootstraps of this size, the estimate (after correcting for the known average error) is within 0.7 percentage points of the true value.

For example, suppose that the (uncorrected) average poverty likelihood in a sample of $n = 16{,}384$ with the 2011/12 scorecard and 100% of the national line is 32.7 percent. Then estimates in 90 percent of such samples would be expected to fall in the range of 32.7 – (0.0) – 0.7 = 32.0 percent to 32.7 – (0.0) + 0.7 = 33.4 percent, with the most likely true value being the corrected estimate in the middle of this range, that is, 32.7 – (0.0) = 32.7 percent. This is because the original (uncorrected) estimate is 32.7 percent, the average error is 0.0 percentage points, and the 90-percent confidence interval for 100% of the national line in the validation sample with this sample size is ±0.7 percentage points (Table 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 error (average difference vis-à-vis true values), together with their standard error (precision).

Schreiner (2008) proposes an approach to deriving a formula for the standard errors of estimated poverty rates at a point in time from indirect measurement via scorecards. 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 the 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, Tanzania's 2011/12 HBS gives a direct-measurement estimate of the household-level poverty rate for 100% of the national line in the validation sample of $\hat{p} = 21.5$ percent (Table 1). If this estimate came from a sample of $n = 16{,}384$ households from a population N of 8,444,035 (the number of households in Tanzania in 2011/12 according to the HBS sampling weights), then the finite population correction ϕ is $\sqrt{\frac{8{,}444{,}035 - 16{,}384}{8{,}444{,}035 - 1}} = 0.9990$, which very close to $\phi = 1$. If the desired confidence level is 90-percent ($z = 1.64$), then the confidence interval $\pm c$ is

$$\pm z \cdot \sqrt{\frac{\hat{p} \cdot (1 - \hat{p})}{n}} \cdot \sqrt{\frac{N - n}{N - 1}} = \pm 1.64 \cdot \sqrt{\frac{0.215 \cdot (1 - 0.215)}{16{,}384}} \cdot \sqrt{\frac{8{,}444{,}035 - 16{,}384}{8{,}444{,}035 - 1}} = \pm 0.526$$

percentage points. (If ϕ were taken as 1, then the interval is still ± 0.526 percentage points.)

Unlike the 2011/12 HBS, however, the scorecard does not measure poverty directly, so this formula is not applicable. To derive a formula for the Tanzania scorecard, consider Table 6, which reports empirical confidence intervals $\pm c$ for the errors for the scorecard applied to 1,000 bootstraps of various sizes from the validation sample. For example, with $n = 16{,}384$ and 100% of the national line in the validation sample, the 90-percent confidence interval is ± 0.654 percentage points.[25]

[25] Due to rounding, Table 6 displays 0.7, not 0.654.

Thus, the 90-percent confidence interval with $n = 16{,}384$ is ± 0.654 percentage points for the 2011/12 scorecard and ± 0.526 percentage points for direct measurement. The ratio of the two intervals is $0.654 \div 0.526 = 1.24$.

Now consider the same exercise, but with $n = 8{,}192$. The confidence interval under direct measurement and 100% of the national line in the validation sample is $\pm 1.64 \cdot \sqrt{\frac{0.215 \cdot (1 - 0.215)}{8{,}192}} \cdot \sqrt{\frac{8{,}444{,}035 - 8{,}192}{8{,}444{,}035 - 1}} = \pm 0.744$ percentage points. The empirical confidence interval with Tanzania's scorecard (Table 6) is ± 0.868 percentage points. Thus for $n = 8{,}192$, the ratio of the two intervals is $0.868 \div 0.744 = 1.17$.

This ratio of 1.17 for $n = 8{,}192$ is not far from the ratio of 1.24 for $n = 16{,}384$. Across all sample sizes of 256 or more in Table 6, these ratios are generally close to each other, and the average of these ratios in the validation sample turns out to be 1.20, implying that confidence intervals for indirect estimates of poverty rates via Tanzania's scorecard and 100% of the national line are—for a given sample size—about 20-percent wider than confidence intervals for direct estimates via the 2011/12 HBS. This 1.20 appears in Table 7 as the "α factor for precision" because if $\alpha = 1.20$, then the formula for confidence intervals c for the 2011/12 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 greater than 1.00, it means that the scorecard is less precise than direct measurement. It turns out that α is more than 1.00 for nine of the 18 poverty lines in Table 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 $\widetilde{p}$ 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 $n = N \cdot \left(\frac{z^2 \cdot \alpha^2 \cdot \widetilde{p} \cdot (1 - \widetilde{p})}{z^2 \cdot \alpha^2 \cdot \widetilde{p} \cdot (1 - \widetilde{p}) + c^2 \cdot (N - 1)} \right)$. 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 $n = \left(\frac{\alpha \cdot z}{c} \right)^2 \cdot \widetilde{p} \cdot (1 - \widetilde{p})$.

To illustrate how to use this, suppose the population N is 8,444,035 (the number of households in Tanzania in 2011/12), suppose $c = 0.05235$, $z = 1.64$ (90-percent confidence), and the relevant poverty line is 100% of the national line so that the most sensible expected poverty rate $\widetilde{p}$ is Tanzania's overall poverty rate for that line in 2011/12 (21.5 percent at the household level, Table 1). The α factor is 1.20 (Table 7). Then the sample-size formula gives $n = 8{,}444{,}035 \cdot \left(\frac{1.64^2 \cdot 1.20^2 \cdot 0.215 \cdot (1 - 0.215)}{1.64^2 \cdot 1.20^2 \cdot 0.215 \cdot (1 - 0.215) + 0.05235^2 \cdot (8{,}444{,}035 - 1)} \right) = 239$, which is not far from the sample size of 256 observed for these parameters in Table 6 for 100%

of the national line. Taking the finite population correction factor ϕ as one (1) gives the same result, as $n = \left(\frac{1.20 \cdot 1.64}{0.05235}\right)^2 \cdot 0.215 \cdot (1 - 0.215) = 239$.[26]

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

[26] 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's microenterprise partners in Tanzania should report using the $1.25/day 2005 PPP line. Given the α factor of 1.10 for this line (Table 7), an expected before-measurement household-level poverty rate of 32.2 percent (the all-Tanzania rate for this line in 2011/12, Table 1), and a confidence level of 90 percent ($z = 1.64$), then $n = 300$ implies a confidence interval of $\pm 1.64 \cdot 1.10 \cdot \sqrt{\frac{0.322 \cdot (1 - 0.322)}{300}} = \pm 4.9$ percentage points.

In practice after the end of fieldwork for the HBS in October 2012, a program would select a poverty line (say, 100% of the 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 $\widetilde{p}$ (perhaps based on a previous measurement such as the household-level poverty rate for 100% of the national line for Tanzania of 21.5 percent in the 2011/12 HBS in Table 1), look up α (here, 1.20 in Table 7), assume that the scorecard will still work in the future and for sub-groups that are not nationally representative,[27] and then compute the required sample size. In this illustration, $n = 10{,}000 \cdot \left(\frac{1.64^2 \cdot 1.20^2 \cdot 0.215 \cdot (1 - 0.215)}{1.64^2 \cdot 1.20^2 \cdot 0.215 \cdot (1 - 0.215) + 0.02^2 \cdot (10{,}000 - 1)} \right) = 1{,}405.$

[27] This paper reports accuracy for the scorecard applied to its validation sample, but it cannot test accuracy for later years or for sub-populations that are not nationally representative. Performance after October 2012 will resemble that in the 2011/12 HBS 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 population's poverty rate between two points in time is estimated as the change in the average poverty likelihood of a sample of households from the population.

Because the definition of *poverty* in the data from the 2011/12 HBS used here to construct the new 2011/12 scorecard differs from the definition of *poverty* in the data used by Schreiner (2013a) to construct the old 2007 scorecard, and because many of the indicators in the new 2011/12 scorecard were asked differently (or not at all) in the data used by Schreiner (2013a) for the old 2007 scorecard, this paper cannot test estimates of change over time for Tanzania, and it can only suggest approximate formulas for standard errors. Nonetheless, the relevant concepts are presented here because, in practice, pro-poor organizations in Tanzania can apply the scorecard to collect their own data and to estimate change through time.

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 merely estimates change, and it does not, in and of itself, indicate the reasons 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 51.2, 32.9, and 14.0 percent (100% of the national line, Table 3). Correcting for the known average error in the validation sample of 0.0 percentage points (Table 7), the corrected baseline estimated poverty rate is the households' average poverty likelihood of [(51.2 + 32.9 + 14.0) ÷ 3] – (0.0) = 32.7 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 40.3, 20.2, and 10.9 percent, 100% of the national line, Table 3). Adjusting for the known average error, the average poverty likelihood at follow-up is [(40.3 + 20.2 + 10.9) ÷ 3] – (0.0) = 23.8 percent, an improvement of 32.7 – 23.8 = 8.9

percentage points.[28] Supposing that exactly two years passed between the average baseline interview and the average follow-up interview, the estimated annual rate of decrease in poverty is 8.9 ÷ 2 = 4.5 percentage points per year. About one in 11 participants in this hypothetical example cross the poverty line in 2016/8.[29] Among those who start below the line, about one in four (8.9 ÷ 32.7 = 27.2 percent) on net end up above the line.[30]

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 40.3, 20.2, and 10.9 percent. The average across households of the difference in each given household's baseline poverty likelihood and its follow-up poverty likelihood is [(51.2 – 40.3) + (32.9 – 20.2) + (14.0 – 10.9)] ÷ 3 = 8.9 percentage points.[31] Assuming in this example that there are exactly two years between each household's interviews, the estimated annual decrease in poverty is (again) 8.9 ÷ 2 = 4.5 percentage points per year.

[28] 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.
[29] This is a net figure; some start above the line and end below it, and vice versa.
[30] The scorecard does not reveal the reasons for this change.
[31] In this case, the error for this line in Table 7 should *not* be subtracted off.

Given the assumptions of the scorecard, 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 (Schreiner, 2014a).

7.3 Precision for estimates of change in two samples

For two equal-sized independent samples, the same logic as in the previous section can be used to derive a formula relating the confidence interval $\pm c$ 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,[32] 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.

[32] 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 interviews (not twice as many) as does estimating a poverty rate at a point in time.

As before, the formula for standard errors can be rearranged to give a formula for sample sizes before indirect measurement via a scorecard, where $\widetilde{p}$ is based on previous measurements and is assumed equal at both baseline and follow-up: $n = 2 \cdot N \cdot \left(\frac{z^2 \cdot \alpha^2 \cdot \widetilde{p} \cdot (1 - \widetilde{p})}{z^2 \cdot \alpha^2 \cdot \widetilde{p} \cdot (1 - \widetilde{p}) + c^2 \cdot (N - 1)} \right)$. If ϕ can be taken as one, then the formula becomes $n = 2 \cdot \left(\frac{\alpha \cdot z}{c} \right)^2 \cdot \widetilde{p} \cdot (1 - \widetilde{p})$.

This α has been measured for 14 countries (Schreiner, 2016, 2015a, 2015b, 2015c, 2015d, 2013b, 2013c, 2012c, 2010, 2009a, 2009b, 2009c; Schreiner and Woller (2010); and Chen and Schreiner, 2009). The simple average of α across countries—after averaging α across poverty lines and survey years within each country—is 1.08. This rough figure is as reasonable as any to use for Tanzania.

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 national line, $\alpha = 1.08$, $\hat{p} = 0.215$ (the household-level poverty rate in 2011/12 for 100% of the national line in Table 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 (1). Then the baseline sample size is $n = 2 \cdot \left(\frac{1.08 \cdot 1.64}{0.02} \right)^2 \cdot 0.215 \cdot (1 - 0.215) \cdot 1 = 2{,}648$, and the follow-up sample size is also 2,648.

7.4 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:[33]

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

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. With the available data for Tanzania, 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:

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

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

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

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

Given this, a sample-size formula for a group of households to whom the 2011/12 scorecard is applied twice (once after October 2012 and then again later) is

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

In Peru (the only source of a data-based estimate, Schreiner, 2009d), 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 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 (1). The pre-baseline poverty rate p_{2016} is taken as 21.5 percent (Table 1), and α is assumed to be 1.30. Then the baseline sample size is

$n = 2 \cdot \left(\frac{1.30 \cdot 1.64}{0.02}\right)^2 \cdot \left\{-0.02 + 0.016 \cdot 3 + 0.47 \cdot \left[0.215 \cdot (1 - 0.215)\right]\right\} \cdot 1 = 2{,}440$. The same group of 2,440 households is scored at follow-up as well.

8. Targeting

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

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

Households who score at or below a given cut-off should be labeled as *targeted*,[34] 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 scoring, the terms *poor* and *non-poor* have specific definitions. Using these same terms for targeting status is incorrect and misleading.

[34] Others labels are acceptable as long as they describe the segment and do 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*).

Table 8 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).

Table 9 shows the distribution of households by targeting outcome for Tanzania. For an example cut-off of 29 or less, outcomes for 100% of the national line in the validation sample are:

- Inclusion: 11.5 percent are below the line and correctly targeted
- Undercoverage: 10.0 percent are below the line and mistakenly not targeted
- Leakage: 11.7 percent are above the line and mistakenly targeted
- Exclusion: 66.8 percent are above the line and correctly not targeted

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

- Inclusion: 15.0 percent are below the line and correctly targeted
- Undercoverage: 6.5 percent are below the line and mistakenly not targeted
- Leakage: 20.2 percent are above the line and mistakenly targeted
- Exclusion: 58.3 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 Table 9 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.	

Table 9 shows the hit rate for all cut-offs for the Tanzania scorecard. For 100% of the national line in the validation sample, total net benefit—under the hit rate—is greatest (80.6) for a cut-off of 19 or less, with about four in five households in Tanzania 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 x Households correctly included) + (1 x Households correctly excluded).[35]

[35] Figure 9 also reports BPAC, the Balanced Poverty Accuracy Criteria adopted by USAID for certifying poverty-assessment tools. IRIS Center (2005) made BPAC to consider accuracy in terms of the bias of estimated poverty rates and in terms of targeting inclusion. BPAC = (Inclusion – |Undercoverage – Leakage|) x [100 ÷ (Inclusion + Undercoverage)]. Schreiner (2014b) explains why BPAC does not add any useful information beyond that provided by the more-standard measures used here.

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 Table 10 ("% targeted HHs who are poor") shows, for the Tanzania scorecard applied to the validation sample, the expected poverty rate among households who score at or below a given cut-off. For the example of 100% of the national line, targeting households in the validation sample who score 29 or less would target 23.3 percent of all households (second column) and would be associated with a poverty rate among those targeted of 49.5 percent (third column).

Table 10 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 national line with the validation sample and a cut-off of 29 or less, 53.6 percent of all poor households are covered.

The final targeting measure in Table 10 is the number of successfully targeted poor households for each non-poor household mistakenly targeted (right-most column). For 100% of the national line with the validation sample and a cut-off of 29 or less, covering 1.0 poor households means leaking to 1 non-poor household.

9. Context of poverty-measurement tools in Tanzania

This section discusses five existing poverty-measurement tool for Tanzania in terms of their goals, methods, definitions of *poverty*, data, indicators, bias, precision, and cost. In general, the advantages of the scorecard are its:

- Using data from the most-recent available nationally representative consumption survey
- Having fewer and lower-cost indicators
- Using a consumption-based definition of *poverty* that is widely understood and that is used by the government of Tanzania
- Reporting errors and precision for estimates of poverty rates at a point in time from out-of-sample tests, including formulas for standard errors
- Reporting targeting accuracy, and having targeting accuracy that is likely similar to that of alternative approaches
- Being feasible for pro-poor programs in Tanzania, due to its low cost and transparency

9.1 Gwatkin *et al.*

Gwatkin *et al.* (2007) construct a poverty-measurement tool for Tanzania 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 low-cost indicators available for the 9,735 households in Tanzania's 2004 DHS.[36] 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.[37] 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).

[36] All DHS datasets for Tanzania since 1996 include each household's asset-index score (`dhsprogram.com/topics/wealth-index/Wealth-Index-Construction.cfm`, retrieved 3 June 2016).

[37] Nevertheless, the indicators are similar and the "flat maximum" is important, so carefully built PCA indexes and consumption-based scorecards 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 of households by PCA indexes, directly-measured consumption, and consumption-based scorecards 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).

Most of the 20 indicators in Gwatkin *et al.* are similar to those in the scorecard in terms of their low cost and verifiability:

- Characteristics of the residence:
  - Type of floor
  - Type of walls
  - Type of roof
  - Type of fuel for cooking
  - Source of energy for lighting
  - Source of drinking water
  - Type of toilet arrangement
  - Number of people per sleeping room
- Ownership of consumer durables:
  - Radio
  - Television
  - Refrigerator
  - Bicycle
  - Motorcycle or scooter
  - Car or truck
  - Telephone
  - Iron
- Ownership of a bank account
- Presence of a domestic worker not related to the head of the household
- Land ownership:
  - Hectares for farming
  - Hectares for grazing

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

- Segmenting households by the quintile of their index 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 segmentation, and the last two goals deal with performance monitoring, so the asset index would be used much like the scorecard here. In particular, the scorecard's support for relative (percentile-based) poverty lines allows

the segmentation of households by quintile to see how health (or other things) vary with consumption. Of course, it is also possible to segment households by quintiles based on scores from the scorecard.

Still, the Gwatkin *et al.* index is more costly and difficult-to-use than the scorecard. The index has 20 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 58 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 constructed from data from 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. Like an asset index, the scorecard can be applied to data from a "light" survey that does not collect consumption as long as the "light" survey collects indicators that match those in the scorecard (Schreiner, 2011a).

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 are not comparable across different asset indexes because the definition of *poverty* varies with a given index's indicators and points.

In general, 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.

9.2 Sahn and Stifel (2000)

Sahn and Stifel (2000) use factor analysis—a sister to PCA—to construct an asset index that measures poverty in Tanzania in terms of long-term wealth. Their purpose is to inform governments and donors on the broad progress of poverty-reduction efforts in Africa.

Like Gwatkin *et al.*, Sahn and Stifel (2000) share many of the strengths of the poverty-scoring approach in that it can be used for targeting and in that it is flexible, low-cost, and adaptable to diverse contexts. In particular, an asset index does not require price adjustments over time and nor consumption data.

Sahn and Stifel (2000) construct their index by pooling data from Tanzania's 1991 and 1996 DHS. The nine indicators are simple, low-cost, and verifiable:

- Characteristics of the residence:
  - — Type of floor
  - — Source of drinking water
  - — Type of toilet arrangement
- Education of the household head
- Ownership of consumer durables:
  - — Radio
  - — Television
  - — Refrigerator
  - — Bicycle
  - — Motorized transport

After defining *poverty* according to lines set at the 25^{th} and 40^{th} percentiles of scores from the index, they compare the distribution of the scores and index-based poverty rates over time within Tanzania.

Sahn and Stifel (2000) also construct a single index from pooled DHS data from 11 sub-Saharan countries—including Tanzania—with two DHS rounds, along with five other sub-Saharan countries with only a single DHS round. This elegantly allows them to compare asset-based poverty across time (within a country) and across countries[38] based on a single index with a definition of *poverty* that—unlike a consumption-based definition—is measured consistently across time and countries.

Booysen *et al.* (2008) is like Sahn and Stifel (2000), except that Booysen *et al.* look at both poverty rates and inequality measures, they use more recent data, and they use three rounds of DHS data (1991/2, 1996, and 1999 for Tanzania) rather than two. In addition, Booysen *et al.* use Multiple Correspondence Analysis instead of factor analysis. MCA is PCA, *sans* the assumption that indicators have Normal distributions. In principle, this makes MCA better suited for categorical indicators, although Booysen *et al.* do not show that this matters in practice. Using DHS data and setting poverty lines at the 40th and 60th percentile of scores from their index, Booysen *et al.* find that Tanzania was the second-poorest of their seven countries and that asset-based poverty in Tanzania decreased from 1991/2 to 1999.

[38] Their ranking ignores that the time between the two DHS surveys varies by country.

9.3 Minot

Minot (2007) applies a "poverty mapping" approach (Hentschel *et al.*, 2000; Elbers, Lanjouw, and Lanjouw, 2003) to estimate changes in poverty rates in Tanzania between 1991/2 and 2003 and to estimate regional point-in-time poverty rates. The goal is to inform policy and to help target assistance by improving data both on poverty trends and on the spatial distribution of poverty.

To do this, Minot first constructs a poverty-assessment tool using only indicators matched across the 1991/2 HBS and Tanzania's DHS rounds of 1991/2, 1996, 1999, and 2003. Four regional tools (Dar es Salaam, Large Towns, Small Towns, and Rural) are built with stepwise ordinary least-squares regression of the matched indicators on the logarithm of per-capita consumption. The four tools include an average of 17 indicators from among the following 22, most of which are similar to those in Gwatkin *et al.* (2007) and Sahn and Stifel (2000):

- Demographics:
  - Number of household members (and its square)
  - Share of males in the household who are:
    - 4-years-old or younger
    - 5- to 15-years-old
    - 16- to 30-years-old
    - 56-years-old or older
  - Share of females in the household who are:
    - 4-years-old or younger
    - 5- to 15-years-old
    - 16- to 30-years-old
    - 31- to 55-years-old
    - 56-years-old or older
  - Age of the head

- Highest educational attainment of:
  - Head
  - Spouse of head
- Characteristics of the residence:
  - Presence of electricity
  - Type of floor
  - Source of drinking water
  - Type of toilet arrangement
- Ownership of consumer durables:
  - Radio
  - Television
  - Refrigerator
  - Motorbike
  - Car

The four tools are then used to estimate per-capita consumption for each surveyed household in each DHS round. Following Hentschel *et al.* (2000), each household's estimated consumption is converted to an estimated poverty likelihood, and finally the poverty likelihoods are averaged to estimate poverty rates.

The approach estimates a decrease in the person-level poverty rate from 1991/2 to 2003 of about 9 percentage points, contrasting with the 2.9-percentage-point decrease found between the 1991/2 and 2000/1 HBS. Unexpectedly, the estimates also suggest that the decrease in poverty is not smaller for more-remote households.

In terms of error, Minot reports that the poverty-assessment tool, when applied to the 1991/2 DHS, gives a person-level poverty rate for Tanzania as a whole that is 0.5 percentage points lower than the direct-measure estimate from the 1991/2 HBS. Because Minot reports sample sizes, poverty-rate estimates, and standard errors for 20 regions, it is possible to derive an average α factor for precision (2.47).

Minot *et al.* (2006) say that this is the first application of a poverty-measurement tool constructed from data from a household consumption survey with data from a survey that does not include consumption. Other examples include Mathiassen (2009), Stifel and Christiaensen (2007), Azzarri *et al.* (2005), Simler, Harrower, and Massingarela (2003), and Schreiner (2012d, 2011a, and 2011b).

9.4 Ward *et al.*

Ward, Owens, and Kahyrara (2002) and Ward and Abassy (2003) ("Ward *et al.*") develop a "core poverty module" with indicators that both matter directly for well-being and development (such as school attendance) and predict consumption-based poverty. They propose that the indicators be included in all national household surveys in Tanzania to facilitate "poverty-mapping" approaches (as in Minot, 2007) that allow estimating consumption for households in surveys that do not collect consumption. Like the scorecard, Ward *et al.* encourage adoption of their core indicators by providing a formatted questionnaire as well as guidelines for how to ask questions and interpret answers based on the HBS *Enumerator Manual*.

Ward *et al.* build poverty-assessment tools via ordinary least-squares stepwise regressions that predict the logarithm of per-adult-equivalent consumption with indicators from the 2000/1 HBS and also with a reduced set of indicators that are matched between the 2000/1 HBS and the 2002 Population Census. They also use

judgment to select indicators that matter directly for well-being and development, that respond to feedback from users, and that address region-specific characteristics.

Ward *et al.* repeatedly run into the "flat maximum", finding sharply decreasing returns to additional indicators across many aspects of out-of-sample performance (R^2, targeting accuracy, and concordance of estimated with actual tercile ranks).[39] They also find little improvement when replacing the all-Tanzania tool with three regional-specific tools, and they find little degradation when they limit themselves to census indicators.

There are 17 indicators in the core set for Tanzania as a whole (Ward and Abassy, 2003, pp. 14–15):

- Number of household members (and its square)
- Education of the household head
- Type of employment of the household head
- Characteristics of the residence:
  - Type of roof
  - Source of drinking water
  - Type of toilet arrangement
  - Administrative region
  - Urban/rural
- Ownership of consumer durables:
  - Radio
  - Iron
  - Telephone
  - Bicycle
- Hectares of land owned/used for agriculture/herding (and its square)
- Food consumption:
  - Number of meals per day
  - Number of days in the past week in which the household ate meat
  - How often the household has not had enough to eat
- Percentage of households in the Primary Sampling Unit with a bank account

[39] Ward, Owens, and Kahyrara (2002) say that there are "limited gains from including more than 10 additional variables."

Ward *et al.*'s core set includes some indicators that the scorecard avoids. In particular, the three food-security questions refer to the past and thus are not verifiable. The question about getting enough to eat is subjective and will embarrass some respondents, as will the question about formal bank accounts. Finally, scorecard users—unlike the NBS, the chief user envisioned by Ward *et al.*—would not have access to data on bank-account ownership at the level of Primary Sampling Units (even if an interviewed household lived in a PSU covered by a NBS survey, and even if the identity of the PSU were known).

All in all, Ward *et al.*'s approach is close to that of the scorecard. It is also close to that of IRIS Center's Poverty-Assessment Tool (PAT), a first cousin of the scorecard (Schreiner, 2014b).

9.5 Tatem *et al.*

Tatem *et al.* (2014) make a poverty map[40] of Tanzania with consumption data from the 3,924 households in the 2010/11 National Panel Survey (Wave 2) with $1.25 and $2.00/day 2005 PPP poverty lines.[41] Their goal is to help governments to target resources more finely.

Tatem *et al.* make a single all-Tanzania poverty-assessment tool with a Bayesian regression of the logarithm of average per-capita aggregate consumption for geo-coded National Panel Survey households against satellite-sensed environmental indicators. These indicators are known for small areas of Tanzania, so they can be fed into the all-Tanzania poverty-assessment tool to estimate average consumption—and poverty rates—in all areas, including those not covered by the National Panel Survey.

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

- Accessability to cities with more than 50,000 people via all transport methods
- Accessability 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`

[40] Called a *high-resolution gridded poverty surface* with an estimate for every km^2.
[41] Tatem *et al.* do not report poverty lines in TZS nor poverty rates. They derive a single all-Tanzania 2005 PPP line based on price indexes from PovcalNet.

The poverty map's accuracy is tested out-of-sample via ten-fold cross-validation. Tatem *et al.* say that their estimates "for Tanzania are essentially unbiased (mean square error = –0.05 [$1.25/day]), . . . indicating no overall tendency to over- or under-predict the poverty-headcount ratio. The mean absolute errors are between 0.13 and 0.19." (p. 37). They also say, "The correlations between actual and predicted values are around 0.48, indicating a satisfactory degree of linear association." Of course, the adequacy of the association depends on the purpose and context of the estimates. And there are some errors in the assessment of bias. In particular, mean squared errors cannot be negative, and mean squared errors need not indicate much about bias. Furthermore, mean absolute bias in the range of 13 to 19 percentage points—and a correlation of estimated-to-true values of 0.48—are not encouraging in terms of precision, even if they do not necessarily imply bias.

10. Conclusion

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

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

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

The scorecard is constructed with half of the data from Tanzania's 2011/12 HBS. Its scores are then calibrated with that same data to poverty likelihoods for 18 poverty lines. The accuracy of the scorecard is tested out-of-sample on data that is not used in scorecard construction. Errors 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.

When the scorecard is applied to the 18 poverty lines in the validation sample, the maximum absolute error for point-in-time estimates of poverty rates is 1.8 percentage points, and the average absolute error is about 0.5 percentage points. Corrected estimates may be had by subtracting the known error for a given poverty line from original, uncorrected estimates.

For $n = 16{,}384$ and 90-percent confidence, the precision of point-in-time estimates of poverty rates is ± 0.7 percentage points or better. With $n = 1{,}024$, the 90-percent confidence intervals are ± 2.6 percentage points or better.

If an organization wants to use the scorecard for segmenting clients for differentiated services, then the results here provide useful information for selecting a targeting 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 10 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 to 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 Tanzania to estimate consumption-based poverty rates, track changes in poverty rates over time, and segment participants for differentiated 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; Demery, Lionel; McKay, Andy; and Finn Tarp. (2015) "Growth and Poverty Reduction in Tanzania", WIDER Working Paper 2015/051, `wider.unu.edu/sites/default/files/wp2015-051.pdf`, retrieved 3 June 2016.

_____; Hussain, M. Azhar; Leyaro, Vincent; Jones, E. Samuel; and Finn Tarp. (2013) "Poverty and Growth in Tanzania", `www1.wider.unu.edu/inclusivegrowth/sites/default/files/IGA/Leyaro.pdf`, retrieved 3 June 2016.

Azzarri, Carlo; Carletto, Gero; Davis, Benjamin; and Alberto Zezza. (2005) "Monitoring Poverty without Consumption Data: An Application Using the Albania Panel Survey", *Eastern European Economics*, Vol. 44, No. 1, pp. 59–82.

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

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

Booysen, 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.

Caire, Dean. (2004) "Building Credit Scorecards for Small-Business Lending in Developing Markets", `microfinance.com/English/Papers/Scoring_SMEs_Hybrid.pdf`, retrieved 3 June 2016.

_____; 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 3 June 2016.

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.

Chen, Shiyuan; and Mark Schreiner. (2009) "Simple Poverty Scorecard Poverty-Assessment Tool: Vietnam", `SimplePovertyScorecard.com/VNM_2006_ENG.pdf`, retrieved 3 June 2016.

Coady, David; Grosh, Margaret; and John Hoddinott. (2004) *Targeting of Transfers in Developing Countries*, `hdl.handle.net/10986/14902`, retrieved 3 June 2016.

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

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

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.

Elbers, Chris; 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.

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 3 June 2016.

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/W9OWN57PD0, retrieved 3 June 2016.

Gwatkin, Davidson R.; Rutstein, Shea; Johnson, Kiersten; Suliman, Eldaw; Wagstaff, Adam; and Agbessi Amouzou. (2007) "Socio-Economic Differences in Health, Nutrition, and Population: Tanzania", World Bank Country Reports on HNP and Poverty, go.worldbank.org/T6LCN5A340, retrieved 3 June 2016.

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 3 June 2016.

Hentschel, Jesko; Lanjouw, Jean Olson; Lanjouw, Peter; and Javier Poggi. (2000) "Combining Census and Survey Data to Trace the Spatial Dimensions of Poverty: A Case Study of Ecuador", *World Bank Economic Review*, Vol. 14, No. 1, pp. 147–165.

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. (2007a) "Manual for the Implementation of USAID Poverty Assessment Tools", povertytools.org/training_documents/Manuals/USAID_PAT_Manual_Eng.pdf, retrieved 3 June 2016.

_____. (2007b) "Introduction to Sampling for the Implementation of PATs", povertytools.org/training_documents/Sampling/Introduction_Sampling.ppt, retrieved 3 June 2016.

_____. (2005) "Notes on Assessment and Improvement of Tool Accuracy", povertytools.org/other_documents/AssessingImproving_Accuracy.pdf, retrieved 3 June 2016.

Johnson, Glenn. (2007) "Lesson 3: Two-Way Tables—Dependent Samples", onlinecourses.science.psu.edu/stat504/node/96, retrieved 3 June 2016.

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

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

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

Martinelli, César; and Susan W. Parker. (2007) "Deception and Misreporting in a Social Program", *Journal of the European Economic Association*, Vol. 4, No. 6, pp. 886–908.

Mathiassen, Astrid. (2009) "A Model-Based Approach for Predicting Annual Poverty Rates without Expenditure Data", *Journal of Economic Inequality*, 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 3 June 2016.

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

Minot, Nicholas. (2007) "Are Poor, Remote Areas Left behind in Agricultural Development? The Case of Tanzania", *Journal of African Economies*, Vol. 17, No. 2, pp. 239–276.

_____; Simler, Ken; Benson, Todd; Kilama, Blandina; Luvanda, Eliab; and Ahmed Makbel. (2006) "Poverty and Malnutrition in Tanzania: New Approaches for Examining Trends and Spatial Patterns", repoa.or.tz/documents_storage/Research%20and%20Analysis/IFPRI%20Report.pdf, retrieved 3 June 2016.

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

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

Narayan, Ambar; and Nobuo Yoshida. (2005) "Proxy-Means Tests for Targeting Welfare Benefits in Sri Lanka", World Bank Report No. SASPR–7, documents.worldbank.org/curated/en/2005/07/6209268/proxy-means-test-targeting-welfare-benefits-sri-lanka, retrieved 3 June 2016.

National Bureau of Statistics. (2014) *Household Budget Survey Main Report 2011/12 (Mainland)*, nbs.go.tz/nbs/takwimu/hbs/The_2011-12_HBSFinalReport.zip, retrieved 3 June 2016.

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.

Ravallion, Martin. (1998) "Poverty Lines in Theory and Practice", World Bank LSMS Working Paper No. 133, go.worldbank.org/8P3IBJPQS1, retrieved 3 June 2016.

Rutstein, Shea Oscar; and Kiersten Johnson. (2004) "The DHS Wealth Index", DHS Comparative Reports No. 6, measuredhs.com/pubs/pdf/CR6/CR6.pdf, retrieved 3 June 2016.

Sahn, David E.; and David C. 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", *SAS/STAT User's Guide, Version 9*, support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_logistic_sect035.htm, retrieved 3 June 2016.

Schreiner, Mark. (2016) "Simple Poverty Scorecard Poverty-Assessment Tool: India", SimplePovertyScorecard.com/IND_2011_ENG.pdf, retrieved 3 June 2016.

_____. (2015a) "Simple Poverty Scorecard Poverty-Assessment Tool: Ghana", SimplePovertyScorecard.com/GHA_2012_ENG.pdf, retrieved 3 June 2016.

_____. (2015b) "Simple Poverty Scorecard Poverty-Assessment Tool: Bolivia", SimplePovertyScorecard.com/BOL_2013_ENG.pdf, retrieved 3 June 2016.

_____. (2015c) "Simple Poverty Scorecard Poverty-Assessment Tool: Malawi", SimplePovertyScorecard.com/MWI_2010_ENG.pdf, retrieved 3 June 2016.

_____. (2015d) "Simple Poverty Scorecard Poverty-Assessment Tool: Cambodia", SimplePovertyScorecard.com/KHM_2011_ENG.pdf, retrieved 3 June 2016.

_____. (2014a) "The Process of Poverty-Scoring Analysis", SimplePovertyScorecard.com/Process_Poverty_Scoring_Analysis.pdf, retrieved 3 June 2016.

_____. (2014b) "How Do the Poverty Scorecard and the PAT Differ?", microfinance.com/English/Papers/Scorecard_versus_PAT.pdf, retrieved 3 June 2016.

_____. (2013a) "Simple Poverty Scorecard Poverty-Assessment Tool: Tanzania", SimplePovertyScorecard.com/TZA_2007_ENG.pdf, retrieved 3 June 2016.

_____. (2013b) "Simple Poverty Scorecard Poverty-Assessment Tool: Bangladesh", SimplePovertyScorecard.com/BGD_2010_ENG.pdf, retrieved 3 June 2016.

_____. (2013c) "Simple Poverty Scorecard Poverty-Assessment Tool: Nicaragua", SimplePovertyScorecard.com/NIC_2009_ENG.pdf, retrieved 3 June 2016.

_____. (2013d) "Simple Poverty Scorecard Poverty-Assessment Tool: Philippines", SimplePovertyScorecard.com/PHL_2009_ENG.pdf, retrieved 3 June 2016.

_____. (2012a) "An Expert-Based Poverty Scorecard for Rural China", microfinance.com/English/Papers/Scoring_Poverty_China_EN.pdf, retrieved 3 June 2016.

_____. (2012b) "Simple Poverty Scorecard Poverty-Assessment Tool: Colombia", SimplePovertyScorecard.com/COL_2009_ENG.pdf, retrieved 3 June 2016.

_____. (2012c) "Simple Poverty Scorecard Poverty-Assessment Tool: Peru", SimplePovertyScorecard.com/PER_2010_ENG.pdf, retrieved 3 June 2016.

_____. (2012d) "Estimating Consumption-Based Poverty in the Ethiopia Demographic and Health Survey", *Ethiopian Journal of Economics*, Vol. 21, No. 1, pp. 73–106.

_____. (2011a) "Estimating Expenditure-Based Poverty in Demographic and Health Surveys".

_____. (2011b) "Estimating Expenditure-Based Poverty in the Bangladesh Demographic and Health Survey", *Bangladesh Journal of Development Studies*, Vol. 34, No. 4, pp. 65–94.

_____. (2010) "Simple Poverty Scorecard Poverty-Assessment Tool: Honduras", `SimplePovertyScorecard.com/HND_2007_ENG.pdf`, retrieved 3 June 2016.

_____. (2009a) "Simple Poverty Scorecard Poverty-Assessment Tool: Philippines", `SimplePovertyScorecard.com/PHL_2004_ENG.pdf`, retrieved 3 June 2016.

_____. (2009b) "Simple Poverty Scorecard Poverty-Assessment Tool: Pakistan", `SimplePovertyScorecard.com/PAK_2005_ENG.pdf`, retrieved 3 June 2016.

_____. (2009c) "Simple Poverty Scorecard Poverty-Assessment Tool: Mexico", `SimplePovertyScorecard.com/MEX_2008_ENG.pdf`, retrieved 3 June 2016.

_____. (2009d) "Simple Poverty Scorecard Poverty-Assessment Tool: Peru", `SimplePovertyScorecard.com/PER_2007_ENG.pdf`, retrieved 3 June 2016.

_____. (2008) "Simple Poverty Scorecard Poverty-Assessment Tool: Peru", `SimplePovertyScorecard.com/PER_2003_ENG.pdf`, retrieved 3 June 2016.

_____. (2006) "Is One Simple Poverty Scorecard Poverty-Assessment Tool Enough for India?", `microfinance.com/English/Papers/Scoring_Poverty_India_Segments.pdf`, retrieved 3 June 2016.

_____. (2005a) "La Herramienta del Índice de Calificación de la Pobreza™: México", `SimplePovertyScorecard.com/MEX_2002_SPA.pdf`, retrieved 3 June 2016.

_____. (2005b) "IRIS Questions on the Simple Poverty Scorecard Poverty-Assessment Tool", `microfinance.com/English/Papers/Scoring_Poverty_Response_to_IRIS.pdf`, retrieved 3 June 2016.

_____. (2002) *Scoring: The Next Breakthrough in Microfinance?* CGAP Occasional Paper No. 7, `microfinance.com/English/Papers/Scoring_Breakthrough_CGAP.pdf`, retrieved 3 June 2016.

_____; 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.*

_____; and Gary Woller. (2010) "Simple Poverty Scorecard Poverty-Assessment Tool: Guatemala", `SimplePovertyScorecard.com/GTM_2006_ENG.pdf`, retrieved 3 June 2016.

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`, retrieved 3 June 2016.

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 3 June 2016.

Simler, Kenneth R.; Harrower, Sarah; and Claudio Massingarela. (2003) "Estimating Poverty Indices from Simple Indicator Surveys", `citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.583.76&rep=rep1&type=pdf`, retrieved 3 June 2016.

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

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

Tarozzi, Alessandro; and Angus Deaton. (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", worldpop.org.uk/resources/docs/Poverty-mapping-report.pdf, retrieved 3 June 2016.

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 3 June 2016.

USAID. (2014) *Microenterprise Results Reporting: Annual Report to Congress, Fiscal Year 2013*, usaid.gov/sites/default/files/documents/1865/MRR.FY13.pdf, retrieved 3 June 2016.

United States Congress. (2004) "Microenterprise Results and Accountability Act of 2004 (HR 3818 RDS)", November 20, smith4nj.com/laws/108-484.pdf, retrieved 3 June 2016.

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.

Ward, Patrick; and Mlemba Abassy. (2003) "Report on the Development of a Core Poverty Module".

_____; Owens, Trudy; and Godius Kahyrara. (2002) "Developing Proxy Predictors for Household Expenditure and Income Poverty".

World Bank. (2015) *Tanzania Mainland Poverty Assessment*, hdl.handle.net/10986/22021, retrieved 3 June 2016.

_____. (2013) "Shared Prosperity: A New Goal for a Changing World", May 8, worldbank.org/en/news/feature/2013/05/08/shared-prosperity-goal-for-changing-world, retrieved 3 June 2016.

_____. (2012) *Targeting Poor and Vulnerable Households in Indonesia*, documents.worldbank.org/curated/en/2012/01/15879773/targeting-poor-vulnerable-households-indonesia, retrieved 3 June 2016.

_____. (2008) "International Comparison Project: Tables of Results", siteresources.worldbank.org/ICPINT/Resources/icp-final-tables.pdf, retrieved 3 June 2016.

Zeller, Manfred. (2004) "Review of Poverty Assessment Tools", pdf.usaid.gov/pdf_docs/PNADH120.pdf, retrieved 3 June 2016.

_____; 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.

Guidelines for the Interpretation of Scorecard Indicators

The excerpts quoted below come from:

National Bureau of Statistics. (2011) "Instruction Manual: Household Budget Survey, 2011/12", [the *Manual*], `catalog.ihsn.org/index.php/catalog/4846/download/60555`, retrieved 3 June 2016.

and

National Bureau of Statistics. (2011) "Questionnaire: Household Budget Survey, 2011/12", [the *Questionnaire*], `nbs.go.tz/nbs/takwimu/hbs/201-12_HBS_Technical-Report.zip`, retrieved 3 June 2016.

The *Manual* for Tanzania is unusually uninformative. When, as will happen frequently, an issue comes up that is not addressed here, its resolution should be left to the unaided judgment of the enumerator, as that seems to have been what Tanzania's National Bureau of Statistics (NBS) did in the 2011/12 HBS. That is, an organization using the scorecard should not promulgate any definitions nor 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 each individual enumerator.

Admittedly, this is a not a satisfactory way to resolve such issues, but in the absence of knowledge (from some source other than the *Manual*) of what the NBS taught enumerators to do in the 2011/12 HBS, there is no better alternative.

General guidelines for asking scorecard questions

Fill out the scorecard header and the "Back-page Worksheet" first, following the directions on the "Back-page Worksheet".

Do not ask the first scorecard indicator directly ("How many household members are 18-years-old or younger?"). Instead, use the information recorded on the "Back-page Worksheet" to determine the response to mark. You must also record the number of household members (regardless of age) in the scorecard header next to "Number of household members:".

In the same way, do not ask the second indicator directly ("Are all household members ages 6 to 18 currently in school?"). Rather, use the information recorded on the "Back-page Worksheet" to determine the response to mark.

In general, do not read the response options to the respondent. Just 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 help 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 happen to see something yourself—such as a consumer durable that the respondent avers not to possess, or a child eating in the room who has not been counted as a member of the household—that suggests that a response may not be accurate.

In general, your application of the scorecard should mimic as closely as possible NBS's application of the 2011/12 HBS. For example, poverty-scoring interviews should take place in respondents' homesteads because the 2011/12 HBS took place in respondents' homesteads.

Questionnaire translation:
It seems that the 2011/12 HBS left translation of the survey instrument to languages other than English to each individual enumerator (perhaps with the help of local translators). When such translation was needed, it was done on the fly.

While the application of the scorecard should, in general, mimic the application of the 2011/12 HBS, it nevertheless makes sense to have a standard, well-done, checked translation to languages that are common in Tanzania (such as Kiswahili, among others).

Without a standard translation, the variation in translations and interpretations across enumerators could greatly harm data quality. Of course, any translation should reflect the meaning in the original English HBS survey instrument as closely as possible. Ideally, all organizations using the scorecard in a given dialect or language in Tanzania would coordinate to produce and use a single translation.

Confidentiality:
According to the title page of the *Questionnaire*, all information collected "is strictly confidential and is to be used for statistical purposes only."

According to p. 2 of the *Manual*, "Keep the information collected strictly confidential. Do not disclose it to anyone except those who are involved in the survey. Tell respondents that their information will be combined and processed to make a statistical report. Assure respondents that the information will be kept confidential and will be used only for statistical purposes."

Who should be the respondent?
According to p. 4 of the *Manual*, "If the head of the household is not available, then any other member of household can be interviewed." This implies that the preferred respondent is the head of the household.

According to p. 14 of the *Manual*, the *head of the household* is "the person recognized as the head by the rest of the members of the household. Often he/she is responsible for the financial support and welfare of the household members."

What to take with you:
According to p. 5 of the *Manual*, you should carry the following to the interview:

- Your identity card
- Letter of introduction from [your organization]
- [These "Guidelines"]
- List of households to interview
- Copies of [the scorecard]

Art of interviewing:

Meeting the household for the first time
According to p. 7 of the *Manual*, "Your behaviour and attitude is vital. As an interviewer, your first responsibility is to establish a good rapport. At the beginning of an interview, you and the respondent do not know each other. The respondent's first impression of you will influence his/her willingness to cooperate. Be friendly as you introduce yourself.

"Carry with you a letter and an identification card that indicates that you work with [your organization].

"Remember to wear respectable clothes. Women are to have *khanga* or *kitenge* for protecting them, as not all households have chairs."

Your role as enumerator
According to pp. 7–8 of the *Manual*, "You play the leading role in the survey. As an enumerator, your duty is to collect the required data. The quality of your work is determined by the quality of the data that you collect. The fate of the survey depends on you.

"Carefully follow [these instructions].

"Create a friendly atmosphere with the respondent so that you can get accurate data. Building an atmosphere of trust helps, and it also helps to familiarize yourself with the culture and habits of the community you are working with.

"Your appearance and your first words and actions are vital for winning the respondent's cooperation.

"Convince respondents that you are knowledgeable and that their participation in the survey exercise will benefit [the participants of your organization].

"Do not mention—nor promise—any immediate benefit to the responding household (nor to anyone else) as a result of the survey, as this may bias responses.

"Highlight the confidentiality of responses. Do the interview with only yourself and the respondent present.

"Make interviews short, but do not hurry the respondent.

"Unlike a normal conversation, in the interview you are asking all the questions, and the respondent is giving all the answers. Never express your own opinion, and do not react in any way to what the respondent says. Do not show any approval nor disapproval; always remain neutral.

"If the respondent is reluctant or unwilling to answer a question, then try to overcome that. Explain that all information is confidential. But do not push too hard for an answer; a respondent cannot be forced to give answers.

"Strictly follow the sequence of questions as written." Read them word-for-word as written. In particular, do not mark a response unless you have asked the question of the respondent and the respondent has given an answer.

"Be in control of the situation. Maintain the interest of the respondent at all times. If the respondent digresses or rambles on, then do not stop him/her abruptly or rudely; listen to what he/she has to say, then gently try to steer him/her back to the question."

<u>Persuading reluctant households</u>

According to pp. 8–9 of the *Manual*, "Results from surveys where stronger attempts are made to convince sampled households to participate are more representative than surveys where only willing respondents are interviewed. For example, surveys using replacement households often over-represent the elderly and people who stay at home. The households on your sample list have been randomly chosen. They cannot be replaced with other households.

"Try your level best to get accurate data. Try to build a good relationship with the household. Try to be faithful to them always and to show honest trustfulness. If you have an appointment with the household, try to be there at the agreed time.

"You will have your own style for persuading unwilling household to cooperate. The main points to consider are:

- Make the household feel that it the most important household in the survey. Tailor your response to their particular circumstances. If they are busy, ill, or harassed, then become involved in their problem
- Be prepared to be flexible. If the respondent wants change an appointment or wants you to call at a different time, then make it clear that you are happy to accommodate them
- Do as much as possible to convince the household to cooperate
- Be sure to tell the household that this survey is for the benefit of [the participants of your organization]
- If the household still refuses, seek assistance from a local leader in the area or from your supervisor

"Here are a some of effective approaches for countering refusals:

- Remind the respondent that the survey need not be completed in a single visit and that you can return as many times as needed
- If the household fears that questions will be too personal or intrusive, remind the household of its right to refuse to answer individual questions, saying, 'The interview is completely voluntary, and we encourage you to refuse questions that you find too personal, sensitive, or intrusive. The questions you do answer will still be important to us, no matter how few'
- If the household protests that it is too busy, then stress that it is precisely this group of busy households that is the most important group in the survey. Say that because of the large numbers of unemployed and retired people, you need all the information that only full-time workers can supply. Always make it clear that you are happy to accommodate their requirements
- If the household is worried about confidentiality, inform them that the files are protected by secret code words which are changed frequently. In addition, you can guarantee that no information will ever be released that can be linked to their name and address. Stressing the 'number crunching' aspect of statistics often helps
- If the household worries that its data might, for example, fall into the hands of government departments, then note that the statistics reveal only a percentage figure of the population rather than singling out any particular household"

Guidelines for specific scorecard indicators

1. How many household members are 18-years-old or younger?
    A. Six or more
    B. Five
    C. Four
    D. Three
    E. Two
    F. One
    G. None

Page 14 of the *Manual* defines *household* with the following list of statements:

- "Only normal members of the household are relevant
- A household may have one person or more than one person
- A *one-person household* is a person who lives alone in all or in part of a housing unit and who has independent consumption
- A *multi-person household* is a group of two or more persons who occupy all or part of a housing unit and who share consumption. Often households of this type consist of a husband, wife, and children. Other relatives, boarders, and visitors are included as members of the household if they pool resources, share consumption, and have been living with the household for at least two weeks
- Household servants count as members of the household if and only if they take their meals with the household and recognize the same head
- Children who are at boarding school are counted as household members
- Lodgers who share the residence and meals by paying are counted as household members. If a lodger does not contribute or does not share meals with the members of the household, then the lodger is not a member of the household
- A husband with more than one wife who spends his time in more than one household is counted as household member only in that household in which he spends the largest share of his time

"A household is made up of those who share consumption and who pool resources."

According to p. 2 of the *Questionnaire*, [when you list household members on the "Back-page Worksheet"] you should "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.

"To make a comprehensive list of household members:

- Ask for the names and ages of all the members of the immediate (nuclear) family who normally live and eat their meals together here
- Ask for the names and ages of any other persons related to household members who normally live and eat their meals together here
- Ask for the names and ages of other persons who are not here now but who normally live and eat their meals here (for example, household members studying elsewhere or traveling)
- Ask for the names and ages of any other persons not related to household members who normally live and eat their meals together here, such as live-in servants"

2. Are all household members ages 6 to 18 currently in school?
    A. No
    B. Yes
    C. No members ages 6 to 18

The *Manual* provides no information related to this indicator.

3. What is the main building material used for the walls of the main building?
    A. Baked bricks
    B. Poles and mud, grass, sun-dried bricks, or other
    C. Stones, cement bricks, or timber

According to p. 26 of the *Manual*, "If the dwelling unit is constructed with more than one building material, then consider the main building material that was used during the construction process."

According to p. 15 of the *Manual*, a household's *residence* or *dwelling unit* is "all the living space occupied by the household, regardless of the physical arrangement of facilities. For example, it may be one room occupied by lodgers or it may be one, two, or more housing units occupied by an extended family or household."

4. What is the main building material used for the roof of the main building?
    A. Grass/leaves, mud and leaves, or other
    B. Iron sheets, tiles, concrete, or asbestos

According to p. 26 of the *Manual*, "If the dwelling unit is constructed with more than one building material, then consider the main building material that was used during the construction process."

According to p. 15 of the *Manual*, the household's *residence* or *dwelling unit* is "all the living space occupied by the household, regardless of the physical arrangement of facilities. For example, it may be one room occupied by lodgers or it may be one, two, or more housing units occupied by an extended family or household."

5. What is the main fuel used for cooking?
    A. Firewood, coal, solar, gas (biogas), wood/farm residuals, or animal residuals
    B. Charcoal, paraffin, gas (industrial), electricity, generator/private source, or other

The *Manual* provides no information related to this indicator.

6. Does your household have any televisions?
    A. No
    B. Yes

The *Manual* provides no information related to this indicator.

7. Does your household have any radios, cassette/tape recorders, or hi-fi systems?
    A. No
    B. Yes

The *Manual* provides no information related to this indicator.

8. Does your household have any lanterns?
    A. No
    B. Yes

The *Manual* provides no information related to this indicator.

9. Does your household have any tables?
    A. No
    B. Yes

The *Manual* provides no information related to this indicator.

10. If the household cultivated any crops in the last 12 months, does it currently own any bulls, cows, steers, heifers, male calves, female calves, or oxen?
    A. No crops, and no cattle
    B. No crops, and cattle
    C. Crops, but no cattle
    D. Crops, and cattle

According to p. 16 of the *Manual*, the *last 12 months* is "the previous 12 months ending the day before the interview."

According to p. 60 of the *Manual*, the distinctions among types of cattle are:

- *"Bull*: Mature, uncastrated male cattle used for breeding
- *Cow*: Mature female cattle that has given birth at least once
- *Steer*: Castrated male cattle over 1 year of age
- *Heifer*: Female cattle of 1 year of age, up to the first calving
- *Oxen*: Castrated male cattle over 1 year of age used as a draft animal (for example, for cultivation)
- *Calves*: Young cattle under 1 year of age"

Mark the appropriate response option based on the combination of the cultivation of crops and the ownership of cattle:

Crops	**Cattle**	**Response**
No	No	A
No	Yes	B
Yes	No	C
Yes	Yes	D

Table 1: National poverty lines, poverty rates, and sample size for Tanzania, Dar es Salaam, Other Urban, and Rural, for the construction and validation samples, by households and people, for 2011/12

Region	Line or Rate	HHs or People	HHs Surveyed	Poverty lines (TZS/adult equivalent/day) and rates (%)			
					National poverty lines		
				Food	100%	150%	200%
All Tanzania							
	Line	People		852	1,191	1,787	2,382
	Rate	HHs	10,186	7.2	21.5	48.2	66.5
	Rate	People		9.7	28.2	57.6	75.7
Dar es Salaam							
	Line	People		1,019	1,426	2,139	2,852
	Rate	HHs	3,016	0.7	2.6	12.4	28.2
	Rate	People		1.0	4.1	18.1	38.2
Other Urban							
	Line	People		883	2,235	3,352	4,470
	Rate	HHs	3,040	6.0	16.0	37.5	57.0
	Rate	People		8.7	21.7	46.4	66.0
Rural							
	Line	People		820	1,146	1,719	2,292
	Rate	HHs	4,130	8.8	26.7	58.1	76.6
	Rate	People		11.3	33.3	66.1	83.5
Construction and calibration (Selecting indicators and points, and associating scores with poverty likelihoods)							
	Rate	HHs	5,109	7.0	21.5	48.1	66.6
Validation (Measuring accuracy)							
	Rate	HHs	5,077	7.2	21.5	48.2	66.5

Source: 2011/12 HBS.

Poverty lines in units of daily per-adult equivalent TZS in average prices for all of Tanzania during 2011/12 HBS fieldwork.

Table 1: International 2005 and 2011 PPP poverty lines, poverty rates, and sample size for Tanzania, Dar es Salaam, Other Urban, and Rural, for the construction and validation samples, by households and people, for 2011/12

	Line or	HHs or		Poverty lines (TZS per person per day) and poverty rates (%)							
				2005 PPP poverty lines				2011 PPP poverty lines			
Region	Rate	People	HHs Surveyed	$1.25	$2.00	$2.50	$5.00	$1.90	$3.10	$3.80	$4.00
All Tanzania											
	Line	People		1,139	1,823	2,279	4,557	1,261	2,058	2,523	2,655
	Rate	HHs	10,186	32.2	63.0	74.9	93.7	39.4	69.5	79.0	80.7
	Rate	People		40.7	72.8	83.2	96.7	49.0	78.6	86.3	87.8
Dar es Salaam											
	Line	People		1,364	2,182	2,728	5,455	1,510	2,464	3,020	3,179
	Rate	HHs	3,016	4.5	23.1	37.2	78.4	7.5	30.4	43.7	46.9
	Rate	People		6.9	31.9	48.1	86.0	11.3	40.2	54.9	58.0
Other Urban											
	Line	People		1,181	1,890	2,362	4,725	1,308	2,134	2,616	2,753
	Rate	HHs	3,040	22.5	52.1	65.9	90.0	28.6	59.0	70.5	72.8
	Rate	People		29.7	61.9	74.9	94.2	36.9	68.5	79.1	81.4
Rural											
	Line	People		1,096	1,754	2,193	4,386	1,214	1,981	2,428	2,556
	Rate	HHs	4,130	40.4	73.8	84.7	97.6	48.6	80.0	88.2	89.5
	Rate	People		48.4	81.5	90.3	98.8	57.5	86.6	92.7	93.7
Construction and calibration (Selecting indicators and points, and associating scores with poverty likelihoods)											
	Rate	HHs	5,109	32.2	62.7	74.9	93.9	39.4	69.3	79.0	80.9
Validation (Measuring accuracy)											
	Rate	HHs	5,077	32.2	63.0	74.9	93.7	39.4	69.5	79.0	80.7

Source: 2011/12 HBS.

Poverty lines in units of daily per-person TZS in average prices for all of Tanzania during 2011/12 HBS fieldwork.

Table 1: Relative and percentile-based poverty lines, poverty rates, and sample size for Tanzania, Dar es Salaam, Other Urban, and Rural, for the construction and validation samples, by households and people, for 2011/12

Region	Line or Rate	HHs or People	HHs Surveyed	Poverty lines (TZS per person per day) and poverty rates (%)					
				Poorest half	Percentile-based lines				
				<100% Natl.	20th	40th	50th	60th	80th
All Tanzania									
	Line	People		693	766	1,071	1,225	1,432	2,179
	Rate	HHs	10,186	10.5	14.6	31.4	40.3	49.8	71.6
	Rate	People		14.1	20.0	40.0	50.0	60.0	80.0
Dar es Salaam									
	Line	People		829	917	1,281	1,466	1,714	2,608
	Rate	HHs	3,016	0.2	0.3	1.0	1.7	3.2	13.8
	Rate	People		0.3	0.6	1.7	2.8	4.9	19.6
Other Urban									
	Line	People		718	795	1,110	1,270	1,484	2,259
	Rate	HHs	3,040	4.7	6.4	16.7	22.9	32.5	59.3
	Rate	People		6.7	9.5	22.8	30.1	41.1	69.1
Rural									
	Line	People		667	738	1,030	1,179	1,378	2,097
	Rate	HHs	4,130	14.1	19.8	41.6	52.8	63.8	86.2
	Rate	People		18.0	25.5	50.0	61.9	72.8	91.4
Construction and calibration (Selecting indicators and points, and associating scores with poverty likelihoods)									
	Rate	HHs	5,109	10.5	14.8	31.6	40.1	49.8	71.4
Validation (Measuring accuracy)									
	Rate	HHs	5,077	10.5	14.6	31.4	40.3	49.8	71.6

Source: 2011/12 HBS.

Poverty lines in units of daily per-person TZS in average prices for all of Tanzania during 2011/12 HBS fieldwork.

Table 2: Poverty indicators

Uncertainty coefficient	Indicator (Responses ordered starting with those linked with higher poverty likelihoods)
90	How many household members are there? (Six or more; Five; Four; Three; Two; One; None)
88	How many household members are 18-years-old or younger? (Six or more; Five; Four; Three; Two; One; None)
86	How many household members are 15-years-old or younger? (Five or more; Four; Three; Two; One; None)
85	How many household members are 16-years-old or younger? (Five or more; Four; Three; Two; One; None)
84	How many household members are 17-years-old or younger? (Six or more; Five; Four; Three; Two; One; None)
79	How many household members are 14-years-old or younger? (Five or more; Four; Three; Two; One; None)
73	How many household members are 13-years-old or younger? (Four or more; Three; Two; One; None)
70	Are all household members ages 6 to 16 currently in school? (No; Yes; No members ages 6 to 16)
69	How many household members are 12-years-old or younger? (Four or more; Three; Two; One; None)
69	What is the main fuel used for cooking? (Firewood, coal, solar, gas (biogas), wood/farm residuals, or animal residuals; Charcoal, paraffin, gas (industrial), electricity, generator/private source, or other)
69	Are all household members ages 6 to 15 currently in school? (No; Yes; No members ages 6 to 15)
65	How many household members are 11-years-old or younger? (Four or more; Three; Two; One; None)
64	Are all household members ages 6 to 14 currently in school? (No; Yes; No members ages 6 to 14)
64	Are all household members ages 6 to 18 currently in school? (No; Yes; No members ages 6 to 18)
63	Are all household members ages 6 to 17 currently in school? (No; Yes; No members ages 6 to 17)
62	What is the main fuel used for lighting? (Firewood, or other; Paraffin, candles, or gas (biogas); Electricity, or solar)
62	Does your household have any stoves (firewood, coal, charcoal, or electric or gas (including microwaves))? (No; Only firewood or coal; Charcoal, but no electric or gas (including microwaves), regardless of firewood or coal; Electric or gas (including microwaves), regardless of any others)
61	Are all household members ages 6 to 13 currently in school? (No; Yes; No members ages 6 to 13)

Table 2 (cont.): Poverty indicators

Uncertainty coefficient	Indicator (Responses ordered starting with those linked with higher poverty likelihoods)
61	What is the main building material used for the walls of the main building? (Baked bricks; Poles and mud, grass, sun-dried bricks, or other; Stones, cement bricks, or timber)
60	Are all household members ages 6 to 11 currently in school? (No; Yes; No members ages 6 to 11)
58	What is the main building material of the floor? (Earth/sand, dung, palm/bamboo, or other; Cement, ceramic tiles, wood planks, parquet or polished wood, or vinyl or asphalt strips)
57	Looking at this card (show card), can you tell me what is the main toilet facility used by this household? (No toilet/bush/field, composting toilet/ecosan latrine, or other; Open pit without slab; Pit latrine with slab (not washable); Ventilated improved pit latrine; Pit latrine with slab (washable); Pour-flush toilet, or flush toilet with cistern)
57	Are all household members ages 6 to 12 currently in school? (No; Yes; No members ages 6 to 12)
56	Does your household have any charcoal stoves? (No; Yes)
56	Is the main building connected to TANESCO? (No; Yes)
52	Does your household have any televisions? (No; Yes)
51	Does your household have any televisions, video recorders/DVD players, or satellite dishes/decoders? (No; Only television; Video recorders/DVD players, or satellite dishes/decoders (regardless of televisions))
48	In the past 12 months, how many household members had as their main status employee or self-employed (with or without employees)? (None; One; Two; Three; Four or more)
48	What is the highest grade completed by the male head/spouse? (None; Pre-primary or nursury, adult education, or Primary 1 to 3; Primary 4 to 6; Primary 7 or 8, training after primary, or secondary form I, II, or III; No male head/spouse; Secondary form IV, training after secondary, form V or VI, training after form VI, diploma, other course, university year 1 or more, Master's degree, or PhD)
45	Did you or anyone in this household own or cultivate any plots in the last 12 months? (Yes; No)
45	If the household cultivated any crops in the last 12 months, does it currently own any billy goats, she-goats, male kids, or female kids? (Crops, but no goats; Crops, and goats; No crops (regardless of goats))

Table 2 (cont.): Poverty indicators

Uncertainty coefficient	Indicator (Responses ordered starting with those linked with higher poverty likelihoods)
44	If the household cultivated any crops in the last 12 months, does it currently own any bulls, cows, steers, heifers, male calves, female calves, oxen, billy goats, she-goats, male kids, female kids, rams (including castrated), ewes, male lambs, female lambs, boars, sows, male gilts or piglets, or female gilts or piglets? (Crops, but no large livestock; Crops, and large livestock; No crops (regardless of large livestock))
42	If the household cultivated any crops in the last 12 months, does it currently own any ams (including castrated), ewes, male lambs, or female lambs? (Crops, but no sheep; Crops, and sheep; No crops (regardless of sheep))
42	If the household cultivated any crops in the last 12 months, does it currently own any boars, sows, male gilts or piglets, or female gilts or piglets? (Crops, but no pigs; Crops, and pigs; No crops (regardless of pigs))
42	If the household cultivated any crops in the last 12 months, does it currently own any bulls, cows, steers, heifers, male calves, female calves, or oxen? (No crops, and no cattle; No crops, and cattle; Crops, but no cattle; Crops, and cattle)
42	Did the household cultivate any crops in the last 12 months? (Yes; No)
42	Does your household have any mobile phones or landline telephones? (No; Yes)
40	Does your household have any video recorders/DVD players or satellite dishes/decoders? (No; Yes)
39	Does the household live in the region of Dar es Salaam? (No; Yes)
39	What is the main building material used for the roof of the main building? (Grass/leaves, mud and leaves, or other; Iron sheets, tiles, concrete, or asbestos)
38	Can the male head/spouse read and write a short sentence in Kiswahali, English, Kiswahali and English, or any other language? (None; No male head/spouse; Only Kiswahili; Only English, Kiswahili and English, or other language)

Table 2 (cont.): Poverty indicators

Uncertainty coefficient	Indicator (Responses ordered starting with those linked with higher poverty likelihoods)
38	Can the male head/spouse read a short sentence from a card in Kiswahali or English? (Cannot read at all, or blind or visually impaired; Able to read only parts of sentence in Kiswahili; Able to read whole sentence in Kiswahili; No male head/spouse; Able to read whole sentence or only parts of sentence in English)
37	Does your household have any cupboards? (No; Yes)
34	What was the main status of the female head/spouse in the past 12 months? (Self-employed without employees, or unpaid helper in family business; Student, long-term unemployed, retired, too young to work, too old to work, disabled, or other; Housewife/househusband; Self-employed with employees; No female head/spouse; Employee)
34	How many household members are 6-years-old or younger? (Three or more; Two; One; None)
34	Does your household have any sofas? (No; Yes)
33	Does your household have any irons (electric or charcoal)? (No; Yes)
32	Can the female head/spouse read a short sentence from a card in Kiswahali or English? (Cannot read at all, or blind or visually impaired; Able to read only parts of sentence in Kiswahili; Able to read whole sentence in Kiswahili; No female head/spouse; Able to read whole sentence or only parts of sentence in English)
32	What is the highest grade completed by the female head/spouse? (None, pre-primary or nursury, adult education, or primary 1 to 3; Primary 4 to 6; Primary 7 or 8, training after primary, or secondary form I, II, or III; No female head/spouse; Secondary form IV, training after secondary, form V or VI, training after form VI, diploma, other course, university year 1 or more, Master's, or PhD)
29	Does your household have any fans? (No; Yes)
29	How many other households share toilet facilities with your household? (No toilet/bush/field; None; One; Two to five; More than five)

Table 2 (cont.): Poverty indicators

Uncertainty coefficient	Indicator (Responses ordered starting with those linked with higher poverty likelihoods)
28	What is the main source of drinking water for your household in the dry season? (Tubewell/borehole; Unprotected dug well; Unprotected spring; Protected spring; Surface water (river, dam, lake, pond, stream, canal, irrigation channel), rainwater collection, or other; Protected dug well; Public tap/standpipe, cart with small tank/drum, or tanker truck; Piped water to yard/plot; Piped water into dwelling, or bottled water)
28	What was the main status of the male head/spouse in the past 12 months? (Self-employed without employees; Unpaid helper in family business, housewife/househusband, student, long-term unemployed, retired, too young to work; too old to work, disabled, or other; No male head/spouse; Self-employed with employees; Employee)
27	In the past 12 months, how many household members had as their main status self-employed (with or without employees)? (None; One; Two or more)
27	Does your household have any refrigerators, freezers, or fridge-freezers? (No; Yes)
23	What is the main measure taken by your household to ensure the safety of drinking water? (None; Strain through a cloth, treat with chemicals, boil, or other; Use water filter, or bottled water)
21	Does your household have any tables? (No; Yes)
16	Does your household have any lanterns? (No; Yes)
16	Can the (oldest) female head/spouse read and write a short sentence in Kiswahali, English, Kiswahali and English, or any other language? (No female head/spouse; No; Only Kiswahili; Only English, Kiswahili and English, or other language)
16	Does your household have any books (not for school)? (No; Yes)
15	Is there a place for hand-washing with soap and water present, and where is it located? (Interviewer, please observe) (No toilet/bush/field; No; Yes—near to the latrine; Yes—near the kitchen; Yes—other location)
14	In the past 12 months, was the main status of the female head/spouse employee or self-employed (with or without employees)? (Yes; No; No female head/spouse)

Table 2 (cont.): Poverty indicators

Uncertainty coefficient	Indicator (Responses ordered starting with those linked with higher poverty likelihoods)
14	Does your household have any radios, cassette/tape recorders, or hi-fi systems? (No; Yes)
13	What is the marital status of the female head/spouse? (Polygamous married; Living together; Separated; Monogamous married; Widowed; Divorced; No female head/spouse; Single, never-married)
11	What is the marital status of the male head/spouse? (Polygamous married; Living together; Monogamous married; No male head/spouse; Separated, divorced, or widowed; Single, never-married)
11	Does your household have any radios? (No; Yes)
11	How many rooms are used for sleeping in this household? (One; Two; Three; Four or more)
10	Does your household have any sewing machines? (No; Yes)
10	Does your household have any electric or gas stoves (including microwaves)? (No; Yes)
09	Does your household have any cassette/tape recorders or hi-fi systems? (No; Yes)
08	Does your household have any water heaters? (No; Yes)
07	Does your household have any personal computers/laptops (with or without internet)? (No; Yes)
07	Does your household have any bicycles, motorcycles, motorbikes, scooters, mopeds, *bajaji*, diesel cars, or other cars? (No; Only bicycle; Motorcycle, motorbike, scooter, moped, *bajaji*, but no car (regardless of bicycles); Car (regardless of any others))
06	What is the structure of household headship? (Female head/spouse only; Male head/spouse only; Both male and female heads/spouses)
06	Does your household have any beds? (No; Yes)
05	Does your household have any diesel cars or other cars? (No; Yes)
05	Does your household have any mosquito nets? (No; Yes)
04	What type of storage container is used to collect water from the source? (Bucket without lid, or traditional clay pot without cover; Drums—metal/plastic; Jerry can; Traditional clay pot with cover; Bucket with lid, overhead tank, underground tank, or other)
02	Does your household have any firewood or coal stoves? (Yes; No)

Table 2 (cont.): Poverty indicators

Uncertainty coefficient	Indicator (Responses ordered starting with those linked with higher poverty likelihoods)
02	Does your household have any cooking pots? (No; Yes)
02	Does your household have any bicycles? (No; Yes)
02	Does your household have any motorcycles, motorbikes, scooters, mopeds, or *bajaji*? (No; Yes)
02	Does your household have any chairs? (No; Yes)
01	In the past 12 months, was the main status of the male head/spouse employee or self-employed (with or without employees)? (Yes; No; No male head/spouse)

Source: 2011/12 HBS and 100% of the national poverty line

Tables for
100% of the National Poverty Line

(and tables pertaining to all poverty lines)

Table 3 (100% of the national line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	82.3
15–19	62.1
20–24	51.2
25–29	40.3
30–34	32.9
35–39	20.2
40–44	14.0
45–49	10.9
50–54	6.6
55–59	4.1
60–64	2.2
65–69	1.3
70–74	1.0
75–79	1.0
80–84	1.0
85–89	0.0
90–94	0.0
95–100	0.0

Table 4 (100% of the 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	0	÷	0	=	100.0
5–9	202	÷	202	=	100.0
10–14	1,251	÷	1,521	=	82.3
15–19	2,443	÷	3,935	=	62.1
20–24	3,843	÷	7,499	=	51.2
25–29	4,076	÷	10,104	=	40.3
30–34	3,929	÷	11,924	=	32.9
35–39	2,363	÷	11,705	=	20.2
40–44	1,711	÷	12,177	=	14.0
45–49	985	÷	9,027	=	10.9
50–54	537	÷	8,154	=	6.6
55–59	242	÷	5,902	=	4.1
60–64	102	÷	4,593	=	2.2
65–69	70	÷	5,466	=	1.3
70–74	29	÷	2,926	=	1.0
75–79	19	÷	1,922	=	1.0
80–84	19	÷	1,873	=	1.0
85–89	0	÷	1,026	=	0.0
90–94	0	÷	39	=	0.0
95–100	0	÷	6	=	0.0

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

Table 5 (100% of the national line): Errors (average differences between estimated and true poverty likelihoods for households by score range), with confidence intervals, from 1,000 bootstraps of n = 16,384, 2011/12 scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	0.0	0.0	0.0	0.0
5–9	0.0	0.0	0.0	0.0
10–14	+18.0	5.3	6.4	8.3
15–19	–7.3	5.2	5.5	6.1
20–24	+6.9	2.5	3.0	4.0
25–29	–4.4	3.3	3.5	4.1
30–34	+5.9	1.8	2.2	2.8
35–39	–4.1	2.9	3.1	3.4
40–44	–1.1	1.5	1.8	2.3
45–49	–0.8	1.7	2.0	2.8
50–54	+1.4	1.0	1.2	1.7
55–59	–6.6	4.5	4.7	5.2
60–64	+0.7	0.7	0.9	1.1
65–69	–0.8	0.9	1.1	1.4
70–74	–0.5	0.9	1.1	1.4
75–79	+0.9	0.1	0.1	0.1
80–84	+0.5	0.5	0.6	0.8
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

Table 6 (100% of the national line): Errors (average differences between estimated poverty rates and true values) for a group at a point in time by sample size, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2011/12 scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+0.1	65.5	68.6	79.0
4	–0.9	36.1	41.3	55.1
8	–0.5	28.5	32.7	41.5
16	–0.2	19.7	23.1	28.9
32	+0.1	14.4	16.3	20.8
64	0.0	10.3	11.9	15.1
128	+0.1	7.0	8.4	11.2
256	–0.1	5.2	6.0	7.9
512	–0.1	3.6	4.4	5.9
1,024	–0.1	2.5	3.0	4.1
2,048	0.0	1.8	2.1	2.8
4,096	0.0	1.3	1.5	2.0
8,192	0.0	0.9	1.1	1.5
16,384	0.0	0.7	0.8	1.0

Table 7 (National poverty lines): Errors (average differences between estimates and true values) for estimated poverty rates of a group of households at a point in time, precision, and the α factor for precision, 2011/12 scorecard applied to the validation sample

	Poverty lines			
		National poverty lines		
	Food	**100%**	**150%**	**200%**
Error (estimate minus true value)	–1.0	0.0	–0.2	+0.4
Precision of difference	0.5	0.7	0.7	0.6
α factor for precision	1.42	1.20	1.04	0.95

Results pertain to the 2011/12 scorecard applied to the 2011/12 validation sample.
Errors (differences between estimates and true values) are displayed in units of percentage points.
Precision is measured as 90-percent confidence intervals of estimates in units of ± percentage points.
Error 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.

Table 7 (International 2005 and 2011 PPP poverty lines): Errors (average differences between estimates and true values) for estimated poverty rates of a group of households at a point in time, precision, and the α factor for precision, 2011/12 scorecard applied to the validation sample

	Poverty lines							
	2005 PPP poverty lines				2011 PPP poverty lines			
	$1.25	**$2.00**	**$2.50**	**$5.00**	**$1.90**	**$3.10**	**$3.80**	**$4.00**
Error (estimate minus true value)	–0.1	–0.8	0.0	+0.2	0.0	–0.2	–0.1	+0.1
Precision of difference	0.7	0.6	0.5	0.2	0.7	0.5	0.5	0.4
α factor for precision	1.10	0.94	0.91	0.70	1.05	0.90	0.88	0.84

Results pertain to the 2011/12 scorecard applied to the 2011/12 validation sample.

Errors (differences between estimates and true values) are displayed in units of percentage points.

Precision is measured as 90-percent confidence intervals of estimates in units of ± percentage points.

Error 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.

Table 7 (Relative and percentile-based lines): Errors (average differences between estimates and true values) for estimated poverty rates of a group of households at a point in time, precision, and the α factor for precision, 2011/12 scorecard applied to the validation sample

	Poverty lines					
	Poorest half	Percentile-based lines				
	<100% Natl.	20th	40th	50th	60th	80th
Error (estimate minus true value)	–0.5	+0.2	–0.3	–1.3	–1.3	–1.8
Precision of difference	0.5	0.6	0.7	0.7	0.6	0.5
α factor for precision	1.30	1.23	1.09	1.04	0.99	0.86

Results pertain to the 2011/12 scorecard applied to the 2011/12 validation sample.

Errors (differences between estimates and true values) are displayed in units of percentage points.

Precision is measured as 90-percent confidence intervals of estimates in units of ± percentage points.

Error 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.

Table 8 (All poverty lines): Possible targeting outcomes

True poverty status	Targeting segment: Targeted	Targeting segment: Non-targeted
Below poverty line	**Inclusion** Below poverty line correctly targeted	**Undercoverage** Below poverty line mistakenly non-targeted
Above poverty line	**Leakage** Above poverty line mistakenly targeted	**Exclusion** Above poverty line correctly non-targeted

Table 9 (100% of the national line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2011/12 scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.0	21.5	0.0	78.5	78.5	–100.0
≤9	0.2	21.3	0.0	78.5	78.7	–98.1
≤14	1.2	20.3	0.5	78.0	79.2	–86.3
≤19	3.9	17.6	1.8	76.7	80.6	–55.7
≤24	7.1	14.4	6.1	72.4	79.5	–5.8
≤29	11.5	10.0	11.7	66.8	78.3	+45.4
≤34	15.0	6.5	20.2	58.3	73.2	+5.8
≤39	17.6	3.9	29.3	49.2	66.8	–36.4
≤44	19.3	2.2	39.8	38.8	58.1	–85.1
≤49	20.3	1.2	47.8	30.7	51.0	–122.6
≤54	20.8	0.7	55.5	23.0	43.8	–158.2
≤59	21.2	0.3	61.0	17.6	38.7	–183.8
≤64	21.3	0.2	65.5	13.1	34.3	–204.7
≤69	21.4	0.1	70.8	7.7	29.1	–229.6
≤74	21.5	0.0	73.7	4.8	26.3	–243.0
≤79	21.5	0.0	75.6	2.9	24.4	–251.9
≤84	21.5	0.0	77.4	1.1	22.6	–260.5
≤89	21.5	0.0	78.5	0.0	21.5	–265.3
≤94	21.5	0.0	78.5	0.0	21.5	–265.5
≤100	21.5	0.0	78.5	0.0	21.5	–265.5

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

Table 10 (100% of the 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), 2011/12 scorecard applied to the 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	0.0	100.0	0.0	Only poor targeted
≤9	0.2	100.0	0.9	Only poor targeted
≤14	1.7	71.0	5.7	2.4:1
≤19	5.7	68.2	18.0	2.1:1
≤24	13.2	53.8	32.9	1.2:1
≤29	23.3	49.5	53.6	1.0:1
≤34	35.2	42.5	69.6	0.7:1
≤39	46.9	37.5	81.9	0.6:1
≤44	59.1	32.7	89.9	0.5:1
≤49	68.1	29.8	94.3	0.4:1
≤54	76.2	27.3	96.7	0.4:1
≤59	82.1	25.8	98.6	0.3:1
≤64	86.7	24.5	99.1	0.3:1
≤69	92.2	23.2	99.7	0.3:1
≤74	95.1	22.6	99.9	0.3:1
≤79	97.1	22.1	99.9	0.3:1
≤84	98.9	21.7	100.0	0.3:1
≤89	100.0	21.5	100.0	0.3:1
≤94	100.0	21.5	100.0	0.3:1
≤100	100.0	21.5	100.0	0.3:1

Tables for the Food Poverty Line

Table 3 (Food): 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	39.5
10–14	36.6
15–19	29.9
20–24	21.0
25–29	13.3
30–34	10.4
35–39	4.4
40–44	2.8
45–49	1.5
50–54	1.3
55–59	0.6
60–64	0.6
65–69	0.4
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Table 5 (Food line): Errors (average differences between estimated and true poverty likelihoods for households by score range), with confidence intervals, from 1,000 bootstraps of $n = 16{,}384$, 2011/12 scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	0.0	0.0	0.0	0.0
5–9	–26.6	20.0	21.0	22.2
10–14	+4.9	5.1	6.2	8.5
15–19	–4.4	3.9	4.2	5.2
20–24	+3.5	2.2	2.6	3.3
25–29	–3.8	2.8	3.1	3.4
30–34	+2.1	1.0	1.2	1.7
35–39	–3.4	2.3	2.5	2.7
40–44	–2.2	1.7	1.9	2.1
45–49	+0.7	0.3	0.4	0.6
50–54	–0.5	0.6	0.7	1.0
55–59	–4.5	3.3	3.5	4.0
60–64	+0.4	0.2	0.2	0.3
65–69	–0.8	0.9	1.0	1.2
70–74	–0.9	0.9	1.0	1.3
75–79	0.0	0.0	0.0	0.0
80–84	0.0	0.0	0.0	0.0
85–89	0.0	0.0	0.0	0.0
90–94	0.0	0.0	0.0	0.0
95–100	0.0	0.0	0.0	0.0

Table 6 (Food line): Errors (average differences between estimated poverty rates and true values) for a group at a point in time by sample size, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2011/12 scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	–1.6	53.8	59.8	66.9
4	–1.2	24.9	30.4	40.7
8	–1.2	19.6	22.7	31.6
16	–1.2	14.0	16.9	20.4
32	–1.1	10.5	12.1	15.5
64	–1.0	7.5	9.1	11.0
128	–0.9	5.3	6.1	8.7
256	–1.0	3.8	4.6	6.0
512	–1.0	2.6	3.3	4.2
1,024	–1.0	1.9	2.3	2.9
2,048	–1.0	1.3	1.7	2.1
4,096	–1.0	0.9	1.1	1.4
8,192	–1.0	0.7	0.8	1.0
16,384	–1.0	0.5	0.6	0.8

Table 9 (Food line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2011/12 scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.0	7.4	0.0	92.6	92.6	–100.0
≤9	0.1	7.3	0.1	92.5	92.6	–95.7
≤14	0.7	6.8	1.1	91.5	92.2	–68.0
≤19	2.0	5.5	3.7	88.9	90.8	+2.4
≤24	3.1	4.4	10.1	82.5	85.5	–35.7
≤29	4.7	2.7	18.5	74.0	78.7	–149.1
≤34	5.8	1.7	29.4	63.1	68.9	–295.2
≤39	6.6	0.9	40.3	52.2	58.8	–441.6
≤44	7.0	0.5	52.1	40.5	47.5	–599.6
≤49	7.1	0.4	61.0	31.6	38.6	–719.5
≤54	7.2	0.2	69.0	23.6	30.8	–826.9
≤59	7.4	0.1	74.8	17.8	25.1	–904.6
≤64	7.4	0.1	79.4	13.2	20.6	–966.0
≤69	7.4	0.0	84.8	7.8	15.2	–1,038.9
≤74	7.4	0.0	87.7	4.9	12.3	–1,077.9
≤79	7.4	0.0	89.6	2.9	10.4	–1,103.7
≤84	7.4	0.0	91.5	1.1	8.5	–1,128.9
≤89	7.4	0.0	92.5	0.0	7.5	–1,142.7
≤94	7.4	0.0	92.5	0.0	7.5	–1,143.2
≤100	7.4	0.0	92.6	0.0	7.4	–1,143.3

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

Table 10 (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), 2011/12 scorecard applied to the 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	0.0	100.0	0.0	Only poor targeted
≤9	0.2	58.2	1.6	1.4:1
≤14	1.7	38.2	8.9	0.6:1
≤19	5.7	34.7	26.4	0.5:1
≤24	13.2	23.2	41.0	0.3:1
≤29	23.3	20.3	63.4	0.3:1
≤34	35.2	16.4	77.4	0.2:1
≤39	46.9	14.0	88.3	0.2:1
≤44	59.1	11.8	93.9	0.1:1
≤49	68.1	10.4	95.2	0.1:1
≤54	76.2	9.5	97.3	0.1:1
≤59	82.1	9.0	98.9	0.1:1
≤64	86.7	8.5	99.2	0.1:1
≤69	92.2	8.1	99.8	0.1:1
≤74	95.1	7.8	100.0	0.1:1
≤79	97.1	7.7	100.0	0.1:1
≤84	98.9	7.5	100.0	0.1:1
≤89	100.0	7.4	100.0	0.1:1
≤94	100.0	7.4	100.0	0.1:1
≤100	100.0	7.4	100.0	0.1:1

Tables for
150% of the National Poverty Line

Table 3 (150% of the national line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	93.1
15–19	89.2
20–24	84.5
25–29	77.4
30–34	68.0
35–39	58.1
40–44	42.5
45–49	40.2
50–54	29.2
55–59	24.2
60–64	13.5
65–69	8.6
70–74	5.9
75–79	5.9
80–84	2.6
85–89	2.2
90–94	0.0
95–100	0.0

Table 5 (150% of the national line): Errors (average differences between estimated and true poverty likelihoods for households by score range), with confidence intervals, from 1,000 bootstraps of n = 16,384, 2011/12 scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	0.0	0.0	0.0	0.0
5–9	0.0	0.0	0.0	0.0
10–14	–4.0	2.7	2.8	2.9
15–19	–2.6	2.3	2.4	3.0
20–24	+3.7	1.8	2.1	3.0
25–29	–1.6	1.7	2.0	2.5
30–34	+3.7	2.0	2.4	3.5
35–39	–4.6	3.3	3.5	3.9
40–44	–0.5	1.9	2.3	3.0
45–49	+2.1	2.5	2.9	3.8
50–54	+5.1	2.4	2.9	3.6
55–59	–14.0	8.6	8.9	9.4
60–64	+0.2	2.5	2.9	3.9
65–69	+1.6	1.5	1.8	2.5
70–74	+2.2	1.3	1.6	2.0
75–79	+0.1	2.6	3.1	4.0
80–84	+0.6	1.3	1.6	2.0
85–89	+0.9	1.4	1.5	1.9
90–94	0.0	0.0	0.0	0.0
95–100	0.0	0.0	0.0	0.0

Table 6 (150% of the national line): Errors (average differences between estimated poverty rates and true values) for a group at a point in time by sample size, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2011/12 scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	–1.5	69.4	76.6	88.6
4	–0.8	40.6	46.8	57.1
8	–0.8	29.2	34.6	44.6
16	–0.6	21.3	24.7	32.9
32	–0.3	14.4	17.9	22.6
64	–0.5	9.9	12.1	15.8
128	–0.2	7.4	9.0	11.3
256	–0.2	5.3	6.1	8.2
512	–0.3	3.8	4.5	6.2
1,024	–0.3	2.6	3.2	4.2
2,048	–0.3	1.8	2.1	2.9
4,096	–0.3	1.4	1.6	2.1
8,192	–0.2	0.9	1.1	1.5
16,384	–0.2	0.7	0.8	1.0

Table 9 (150% of the national line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2011/12 scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.0	48.2	0.0	51.8	51.8	–100.0
≤9	0.2	48.0	0.0	51.8	52.0	–99.2
≤14	1.7	46.6	0.1	51.7	53.4	–93.0
≤19	5.3	42.9	0.4	51.4	56.7	–77.3
≤24	11.1	37.1	2.0	49.7	60.9	–49.6
≤29	18.9	29.3	4.3	47.5	66.4	–12.5
≤34	26.9	21.3	8.3	43.5	70.4	+28.8
≤39	34.2	14.1	12.7	39.0	73.2	+68.1
≤44	39.6	8.6	19.5	32.3	71.9	+59.7
≤49	43.0	5.2	25.1	26.7	69.7	+48.0
≤54	45.1	3.2	31.2	20.6	65.7	+35.3
≤59	46.9	1.4	35.3	16.5	63.3	+26.8
≤64	47.5	0.8	39.3	12.5	60.0	+18.5
≤69	47.9	0.3	44.3	7.5	55.4	+8.1
≤74	48.1	0.2	47.1	4.7	52.8	+2.4
≤79	48.2	0.1	48.9	2.9	51.0	–1.4
≤84	48.2	0.0	50.7	1.0	49.3	–5.2
≤89	48.2	0.0	51.7	0.0	48.3	–7.3
≤94	48.2	0.0	51.8	0.0	48.2	–7.4
≤100	48.2	0.0	51.8	0.0	48.2	–7.4

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

Table 10 (150% of the 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), 2011/12 scorecard applied to the 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	0.0	100.0	0.0	Only poor targeted
≤9	0.2	100.0	0.4	Only poor targeted
≤14	1.7	95.8	3.4	22.6:1
≤19	5.7	93.4	11.0	14.2:1
≤24	13.2	84.6	23.1	5.5:1
≤29	23.3	81.4	39.3	4.4:1
≤34	35.2	76.5	55.8	3.3:1
≤39	46.9	72.9	70.8	2.7:1
≤44	59.1	67.1	82.1	2.0:1
≤49	68.1	63.2	89.2	1.7:1
≤54	76.2	59.1	93.4	1.4:1
≤59	82.1	57.0	97.2	1.3:1
≤64	86.7	54.7	98.4	1.2:1
≤69	92.2	52.0	99.3	1.1:1
≤74	95.1	50.5	99.7	1.0:1
≤79	97.1	49.6	99.9	1.0:1
≤84	98.9	48.7	100.0	1.0:1
≤89	100.0	48.2	100.0	0.9:1
≤94	100.0	48.2	100.0	0.9:1
≤100	100.0	48.2	100.0	0.9:1

Tables for
200% of the National Poverty Line

Table 3 (200% of the national line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	100.0
15–19	98.6
20–24	93.8
25–29	93.8
30–34	87.4
35–39	79.4
40–44	67.9
45–49	63.7
50–54	51.2
55–59	43.8
60–64	31.8
65–69	28.1
70–74	19.5
75–79	16.8
80–84	7.3
85–89	7.3
90–94	7.3
95–100	0.0

Table 5 (200% of the national line): Errors (average differences between estimated and true poverty likelihoods for households by score range), with confidence intervals, from 1,000 bootstraps of n = 16,384, 2011/12 scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	0.0	0.0	0.0	0.0
5–9	0.0	0.0	0.0	0.0
10–14	0.0	0.0	0.0	0.0
15–19	+2.8	1.5	1.8	2.4
20–24	–1.5	1.2	1.3	1.6
25–29	+2.7	1.3	1.5	2.0
30–34	+3.1	1.5	1.8	2.3
35–39	–4.2	2.9	3.0	3.3
40–44	–1.7	1.9	2.2	2.7
45–49	+8.1	2.6	3.1	3.8
50–54	–4.5	3.6	3.8	4.1
55–59	–10.2	6.7	7.1	7.6
60–64	+3.2	3.4	4.2	5.1
65–69	+6.3	2.7	3.3	4.9
70–74	+7.7	2.5	3.1	4.0
75–79	+7.1	3.2	3.8	4.7
80–84	+1.6	2.3	2.8	3.7
85–89	+4.7	1.8	2.1	2.6
90–94	+7.3	0.0	0.0	0.0
95–100	0.0	0.0	0.0	0.0

Table 6 (200% of the national line): Errors (average differences between estimated poverty rates and true values) for a group at a point in time by sample size, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2011/12 scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+0.8	67.8	79.6	90.9
4	+0.5	35.8	41.9	56.1
8	0.0	25.2	29.4	41.3
16	+0.3	17.7	21.2	27.4
32	+0.5	12.9	14.9	19.6
64	+0.2	8.9	10.6	14.3
128	+0.5	6.5	7.6	10.2
256	+0.5	4.6	5.3	7.3
512	+0.4	3.3	3.9	5.2
1,024	+0.4	2.3	2.8	3.7
2,048	+0.4	1.6	1.9	2.3
4,096	+0.4	1.2	1.3	1.7
8,192	+0.4	0.8	1.0	1.3
16,384	+0.4	0.6	0.7	0.9

Table 9 (200% of the national line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2011/12 scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.0	66.4	0.0	33.6	33.6	–100.0
≤9	0.2	66.2	0.0	33.6	33.8	–99.4
≤14	1.7	64.7	0.0	33.6	35.3	–94.8
≤19	5.5	60.9	0.1	33.5	39.0	–83.2
≤24	12.6	53.8	0.6	33.0	45.6	–61.2
≤29	21.9	44.5	1.4	32.2	54.1	–32.0
≤34	32.1	34.3	3.1	30.5	62.6	+1.3
≤39	41.8	24.6	5.1	28.6	70.4	+33.6
≤44	50.5	15.9	8.6	25.0	75.5	+65.0
≤49	55.8	10.6	12.3	21.3	77.1	+81.5
≤54	60.1	6.3	16.1	17.5	77.6	+75.7
≤59	63.0	3.4	19.1	14.5	77.5	+71.2
≤64	64.3	2.1	22.4	11.2	75.5	+66.2
≤69	65.6	0.8	26.6	7.0	72.6	+59.9
≤74	66.0	0.4	29.1	4.5	70.5	+56.2
≤79	66.2	0.2	30.8	2.8	69.0	+53.6
≤84	66.3	0.0	32.6	1.0	67.4	+50.9
≤89	66.4	0.0	33.6	0.0	66.4	+49.4
≤94	66.4	0.0	33.6	0.0	66.4	+49.4
≤100	66.4	0.0	33.6	0.0	66.4	+49.4

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

Table 10 (200% of the 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), 2011/12 scorecard applied to the 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	0.0	100.0	0.0	Only poor targeted
≤9	0.2	100.0	0.3	Only poor targeted
≤14	1.7	100.0	2.6	Only poor targeted
≤19	5.7	97.7	8.3	41.8:1
≤24	13.2	95.7	19.0	22.1:1
≤29	23.3	94.1	33.0	15.9:1
≤34	35.2	91.2	48.3	10.4:1
≤39	46.9	89.2	63.0	8.3:1
≤44	59.1	85.5	76.0	5.9:1
≤49	68.1	81.9	84.0	4.5:1
≤54	76.2	78.8	90.5	3.7:1
≤59	82.1	76.7	94.9	3.3:1
≤64	86.7	74.1	96.9	2.9:1
≤69	92.2	71.1	98.8	2.5:1
≤74	95.1	69.4	99.5	2.3:1
≤79	97.1	68.2	99.8	2.1:1
≤84	98.9	67.1	99.9	2.0:1
≤89	100.0	66.4	100.0	2.0:1
≤94	100.0	66.4	100.0	2.0:1
≤100	100.0	66.4	100.0	2.0:1

Tables for
the $1.25/day 2005 PPP Poverty Line

Table 3 ($1.25/day 2005 PPP line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	86.7
15–19	77.5
20–24	71.9
25–29	59.6
30–34	48.0
35–39	35.2
40–44	23.5
45–49	20.2
50–54	13.8
55–59	7.3
60–64	4.2
65–69	1.7
70–74	1.7
75–79	1.7
80–84	1.3
85–89	0.1
90–94	0.0
95–100	0.0

Table 5 ($1.25/day 2005 PPP line): Errors (average differences between estimated and true poverty likelihoods for households by score range), with confidence intervals, from 1,000 bootstraps of n = 16,384, 2011/12 scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	0.0	0.0	0.0	0.0
5–9	0.0	0.0	0.0	0.0
10–14	+0.4	3.7	4.5	5.6
15–19	–9.9	6.0	6.3	6.5
20–24	+9.5	2.6	3.1	3.8
25–29	–2.6	2.4	2.6	3.4
30–34	+2.2	2.0	2.4	3.4
35–39	–5.8	3.9	4.1	4.6
40–44	+0.5	1.7	2.0	2.5
45–49	+1.8	2.2	2.6	3.3
50–54	+5.5	1.2	1.5	2.0
55–59	–5.5	3.9	4.2	4.6
60–64	–2.1	2.1	2.4	2.9
65–69	–1.2	1.2	1.3	1.7
70–74	0.0	1.0	1.2	1.6
75–79	–0.9	1.9	2.2	2.9
80–84	+0.8	0.5	0.6	0.8
85–89	+0.1	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

Table 6 ($1.25/day 2005 PPP line): Errors (average differences between estimated poverty rates and true values) for a group at a point in time by sample size, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2011/12 scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	–0.5	68.0	75.9	86.7
4	–0.6	40.0	46.7	57.3
8	–0.4	28.5	33.2	40.2
16	–0.6	21.9	26.1	32.0
32	–0.2	14.3	16.8	22.2
64	–0.3	10.2	12.4	15.8
128	–0.1	7.1	8.9	11.6
256	–0.2	5.2	6.1	8.2
512	–0.3	3.7	4.5	5.8
1,024	–0.2	2.6	3.1	3.9
2,048	–0.2	1.9	2.2	2.9
4,096	–0.2	1.3	1.6	2.0
8,192	–0.1	0.9	1.1	1.5
16,384	–0.1	0.7	0.8	1.2

Table 9 ($1.25/day 2005 PPP line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2011/12 scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.0	32.2	0.0	67.8	67.8	–100.0
≤9	0.2	32.0	0.0	67.8	68.0	–98.7
≤14	1.5	30.7	0.2	67.6	69.1	–90.0
≤19	4.9	27.3	0.7	67.0	72.0	–67.2
≤24	9.5	22.7	3.6	64.1	73.7	–29.6
≤29	15.6	16.7	7.7	60.1	75.7	+20.5
≤34	21.3	10.9	13.9	53.9	75.2	+56.9
≤39	25.9	6.3	21.0	46.8	72.7	+34.8
≤44	28.7	3.5	30.3	37.5	66.2	+5.9
≤49	30.3	1.9	37.8	30.0	60.2	–17.4
≤54	31.1	1.1	45.1	22.7	53.8	–40.0
≤59	31.7	0.5	50.4	17.3	49.0	–56.6
≤64	32.0	0.3	54.8	13.0	45.0	–70.0
≤69	32.1	0.1	60.1	7.7	39.8	–86.5
≤74	32.2	0.0	63.0	4.8	37.0	–95.4
≤79	32.2	0.0	64.9	2.9	35.1	–101.3
≤84	32.2	0.0	66.7	1.1	33.3	–107.0
≤89	32.2	0.0	67.7	0.0	32.3	–110.2
≤94	32.2	0.0	67.8	0.0	32.2	–110.3
≤100	32.2	0.0	67.8	0.0	32.2	–110.4

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

Table 10 ($1.25/day 2005 PPP 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), 2011/12 scorecard applied to the 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	0.0	100.0	0.0	Only poor targeted
≤9	0.2	100.0	0.6	Only poor targeted
≤14	1.7	87.5	4.7	7.0:1
≤19	5.7	86.9	15.3	6.6:1
≤24	13.2	72.4	29.6	2.6:1
≤29	23.3	66.9	48.3	2.0:1
≤34	35.2	60.5	66.1	1.5:1
≤39	46.9	55.2	80.4	1.2:1
≤44	59.1	48.7	89.2	0.9:1
≤49	68.1	44.5	94.0	0.8:1
≤54	76.2	40.8	96.6	0.7:1
≤59	82.1	38.6	98.4	0.6:1
≤64	86.7	36.8	99.2	0.6:1
≤69	92.2	34.8	99.7	0.5:1
≤74	95.1	33.8	99.8	0.5:1
≤79	97.1	33.2	99.9	0.5:1
≤84	98.9	32.6	100.0	0.5:1
≤89	100.0	32.2	100.0	0.5:1
≤94	100.0	32.2	100.0	0.5:1
≤100	100.0	32.2	100.0	0.5:1

Tables for
the $2.00/day 2005 PPP Poverty Line

Table 3 ($2.00/day 2005 PPP line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	100.0
15–19	99.9
20–24	93.0
25–29	92.4
30–34	84.3
35–39	77.2
40–44	64.3
45–49	54.3
50–54	45.4
55–59	38.7
60–64	25.1
65–69	21.3
70–74	13.8
75–79	12.8
80–84	3.1
85–89	2.8
90–94	0.0
95–100	0.0

Table 5 ($2.00/day 2005 PPP line): Errors (average differences between estimated and true poverty likelihoods for households by score range), with confidence intervals, from 1,000 bootstraps of n = 16,384, 2011/12 scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	0.0	0.0	0.0	0.0
5–9	0.0	0.0	0.0	0.0
10–14	0.0	0.0	0.0	0.0
15–19	+4.1	1.5	1.8	2.4
20–24	–1.0	1.1	1.3	1.8
25–29	+1.9	1.3	1.5	2.0
30–34	–1.2	1.5	1.7	2.2
35–39	–5.6	3.6	3.8	3.9
40–44	+0.1	2.0	2.3	2.9
45–49	+1.7	2.5	3.1	4.0
50–54	–2.8	2.7	3.1	4.2
55–59	–12.2	7.8	8.1	8.7
60–64	+2.7	3.2	3.7	4.6
65–69	+6.6	2.4	2.8	3.6
70–74	+4.2	2.3	2.7	3.6
75–79	+4.5	3.0	3.5	4.3
80–84	+0.8	1.4	1.7	2.1
85–89	+1.4	1.4	1.6	2.0
90–94	0.0	0.0	0.0	0.0
95–100	0.0	0.0	0.0	0.0

Table 6 ($2.00/day 2005 PPP line): Errors (average differences between estimated poverty rates and true values) for a group at a point in time by sample size, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2011/12 scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	–0.1	69.3	79.6	89.6
4	–1.2	36.2	42.5	54.7
8	–1.6	25.2	29.5	41.4
16	–1.3	18.1	22.1	28.0
32	–0.9	13.4	15.8	20.3
64	–1.1	9.3	10.8	15.1
128	–0.8	6.2	7.4	10.4
256	–0.8	4.7	5.4	6.9
512	–0.8	3.3	3.8	4.8
1,024	–0.8	2.4	2.8	3.8
2,048	–0.8	1.6	1.9	2.5
4,096	–0.8	1.2	1.4	1.8
8,192	–0.8	0.8	0.9	1.2
16,384	–0.8	0.6	0.7	0.9

Table 9 ($2.00/day 2005 PPP line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2011/12 scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.0	63.3	0.0	36.7	36.7	–100.0
≤9	0.2	63.1	0.0	36.7	36.9	–99.4
≤14	1.7	61.6	0.0	36.7	38.4	–94.6
≤19	5.5	57.8	0.1	36.6	42.1	–82.3
≤24	12.5	50.8	0.7	36.0	48.5	–59.5
≤29	21.7	41.6	1.6	35.1	56.8	–29.0
≤34	32.0	31.3	3.2	33.5	65.5	+6.1
≤39	41.6	21.7	5.3	31.4	73.0	+39.8
≤44	49.7	13.6	9.4	27.3	76.9	+71.8
≤49	54.5	8.8	13.6	23.1	77.5	+78.5
≤54	58.2	5.1	18.1	18.7	76.9	+71.5
≤59	60.8	2.5	21.4	15.3	76.1	+66.3
≤64	61.8	1.5	24.9	11.8	73.6	+60.6
≤69	62.7	0.6	29.5	7.2	69.9	+53.3
≤74	63.0	0.2	32.1	4.6	67.7	+49.3
≤79	63.2	0.1	33.8	2.9	66.1	+46.5
≤84	63.3	0.0	35.7	1.0	64.3	+43.7
≤89	63.3	0.0	36.7	0.0	63.3	+42.1
≤94	63.3	0.0	36.7	0.0	63.3	+42.0
≤100	63.3	0.0	36.7	0.0	63.3	+42.0

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

Table 10 ($2.00/day 2005 PPP 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), 2011/12 scorecard applied to the 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	0.0	100.0	0.0	Only poor targeted
≤9	0.2	100.0	0.3	Only poor targeted
≤14	1.7	100.0	2.7	Only poor targeted
≤19	5.7	97.7	8.7	41.8:1
≤24	13.2	95.0	19.7	18.9:1
≤29	23.3	93.2	34.2	13.6:1
≤34	35.2	90.9	50.5	10.0:1
≤39	46.9	88.7	65.7	7.9:1
≤44	59.1	84.1	78.4	5.3:1
≤49	68.1	80.0	86.1	4.0:1
≤54	76.2	76.3	92.0	3.2:1
≤59	82.1	74.0	96.0	2.8:1
≤64	86.7	71.3	97.7	2.5:1
≤69	92.2	68.0	99.0	2.1:1
≤74	95.1	66.3	99.6	2.0:1
≤79	97.1	65.1	99.9	1.9:1
≤84	98.9	64.0	100.0	1.8:1
≤89	100.0	63.3	100.0	1.7:1
≤94	100.0	63.3	100.0	1.7:1
≤100	100.0	63.3	100.0	1.7:1

Tables for
the $2.50/day 2005 PPP Poverty Line

Table 3 ($2.50/day 2005 PPP line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	100.0
15–19	99.9
20–24	97.0
25–29	96.7
30–34	95.0
35–39	89.5
40–44	79.6
45–49	75.7
50–54	61.2
55–59	56.9
60–64	45.6
65–69	36.8
70–74	27.1
75–79	23.2
80–84	8.1
85–89	8.1
90–94	7.6
95–100	0.0

Table 5 ($2.50/day 2005 PPP line): Errors (average differences between estimated and true poverty likelihoods for households by score range), with confidence intervals, from 1,000 bootstraps of n = 16,384, 2011/12 scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	0.0	0.0	0.0	0.0
5–9	0.0	0.0	0.0	0.0
10–14	0.0	0.0	0.0	0.0
15–19	+3.8	1.5	1.8	2.4
20–24	–0.4	0.8	1.0	1.4
25–29	0.0	0.8	1.0	1.3
30–34	+3.2	1.2	1.4	1.9
35–39	–1.4	1.3	1.5	1.9
40–44	–3.4	2.4	2.6	2.7
45–49	+4.3	2.5	2.9	3.7
50–54	–4.4	3.4	3.7	4.0
55–59	–6.2	4.7	4.9	5.5
60–64	–4.7	4.2	4.5	5.6
65–69	+10.0	2.9	3.4	4.9
70–74	+4.4	3.8	4.7	6.1
75–79	+2.6	4.6	5.4	6.8
80–84	–2.4	3.9	4.5	5.7
85–89	+3.7	2.3	2.8	3.6
90–94	+7.6	0.0	0.0	0.0
95–100	0.0	0.0	0.0	0.0

Table 6 ($2.50/day 2005 PPP line): Errors (average differences between estimated poverty rates and true values) for a group at a point in time by sample size, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2011/12 scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+1.4	67.0	76.4	94.4
4	+0.2	30.7	37.2	48.1
8	–0.1	20.8	25.6	35.9
16	–0.2	15.2	18.4	24.0
32	0.0	11.2	13.1	17.5
64	–0.3	7.6	9.2	12.6
128	–0.1	5.5	6.4	9.0
256	0.0	4.0	4.8	6.1
512	+0.1	2.9	3.4	4.2
1,024	0.0	2.0	2.4	3.2
2,048	0.0	1.4	1.7	2.2
4,096	0.0	1.0	1.1	1.5
8,192	0.0	0.7	0.8	1.1
16,384	0.0	0.5	0.6	0.8

Table 9 ($2.50/day 2005 PPP line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2011/12 scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.0	74.9	0.0	25.1	25.1	–100.0
≤9	0.2	74.7	0.0	25.1	25.3	–99.5
≤14	1.7	73.2	0.0	25.1	26.8	–95.4
≤19	5.5	69.3	0.1	25.0	30.6	–85.0
≤24	12.9	62.0	0.3	24.8	37.7	–65.3
≤29	22.6	52.2	0.6	24.5	47.2	–38.7
≤34	33.7	41.2	1.5	23.7	57.4	–8.0
≤39	44.4	30.5	2.5	22.6	67.0	+21.9
≤44	54.6	20.3	4.5	20.6	75.2	+51.8
≤49	61.1	13.8	7.0	18.1	79.2	+72.5
≤54	66.4	8.5	9.9	15.2	81.6	+86.8
≤59	69.9	5.0	12.2	12.9	82.8	+83.7
≤64	71.9	3.0	14.9	10.3	82.1	+80.2
≤69	73.5	1.4	18.7	6.4	79.9	+75.0
≤74	74.3	0.6	20.9	4.2	78.5	+72.1
≤79	74.6	0.2	22.4	2.7	77.4	+70.1
≤84	74.8	0.1	24.1	1.0	75.8	+67.8
≤89	74.9	0.0	25.1	0.0	74.9	+66.5
≤94	74.9	0.0	25.1	0.0	74.9	+66.5
≤100	74.9	0.0	25.1	0.0	74.9	+66.5

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

Table 10 ($2.50/day 2005 PPP 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), 2011/12 scorecard applied to the 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	0.0	100.0	0.0	Only poor targeted
≤9	0.2	100.0	0.3	Only poor targeted
≤14	1.7	100.0	2.3	Only poor targeted
≤19	5.7	98.1	7.4	50.9:1
≤24	13.2	97.7	17.2	42.8:1
≤29	23.3	97.4	30.2	36.9:1
≤34	35.2	95.8	45.0	23.0:1
≤39	46.9	94.7	59.3	17.7:1
≤44	59.1	92.4	72.9	12.2:1
≤49	68.1	89.7	81.6	8.7:1
≤54	76.2	87.0	88.6	6.7:1
≤59	82.1	85.1	93.4	5.7:1
≤64	86.7	82.9	96.0	4.8:1
≤69	92.2	79.7	98.2	3.9:1
≤74	95.1	78.0	99.2	3.6:1
≤79	97.1	76.9	99.7	3.3:1
≤84	98.9	75.6	99.9	3.1:1
≤89	100.0	74.9	100.0	3.0:1
≤94	100.0	74.9	100.0	3.0:1
≤100	100.0	74.9	100.0	3.0:1

Tables for
the $5.00/day 2005 PPP Poverty Line

Table 3 ($5.00/day 2005 PPP line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	100.0
15–19	100.0
20–24	99.7
25–29	99.5
30–34	99.5
35–39	99.5
40–44	98.5
45–49	97.3
50–54	94.2
55–59	93.4
60–64	86.1
65–69	79.9
70–74	78.7
75–79	71.3
80–84	44.2
85–89	38.3
90–94	30.6
95–100	0.0

Table 5 ($5.00/day 2005 PPP line): Errors (average differences between estimated and true poverty likelihoods for households by score range), with confidence intervals, from 1,000 bootstraps of n = 16,384, 2011/12 scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	0.0	0.0	0.0	0.0
5–9	0.0	0.0	0.0	0.0
10–14	0.0	0.0	0.0	0.0
15–19	0.0	0.0	0.0	0.0
20–24	+0.3	0.4	0.5	0.6
25–29	–0.5	0.3	0.3	0.3
30–34	+0.1	0.4	0.4	0.6
35–39	+1.4	0.6	0.7	1.0
40–44	–0.2	0.4	0.5	0.7
45–49	–1.3	0.8	0.9	0.9
50–54	–1.1	1.0	1.1	1.4
55–59	+2.0	1.5	1.9	2.4
60–64	–2.1	2.0	2.5	3.6
65–69	+5.7	3.4	4.0	5.4
70–74	+5.2	4.4	5.3	7.1
75–79	+2.6	5.5	6.4	8.5
80–84	+5.2	5.3	6.1	7.8
85–89	–16.5	11.6	12.4	13.5
90–94	–0.4	33.2	37.4	50.0
95–100	0.0	0.0	0.0	0.0

Table 6 ($5.00/day 2005 PPP line): Errors (average differences between estimated poverty rates and true values) for a group at a point in time by sample size, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2011/12 scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	0.0	39.5	54.3	79.8
4	+0.4	15.0	20.0	31.3
8	+0.3	9.7	12.7	18.1
16	+0.3	6.8	8.4	11.7
32	+0.2	4.7	5.9	7.4
64	+0.2	3.3	4.1	5.0
128	+0.2	2.5	3.0	3.8
256	+0.2	1.7	2.1	2.7
512	+0.2	1.2	1.4	2.0
1,024	+0.2	0.9	1.0	1.3
2,048	+0.2	0.6	0.7	0.9
4,096	+0.2	0.4	0.5	0.7
8,192	+0.2	0.3	0.4	0.5
16,384	+0.2	0.2	0.3	0.3

Table 9 ($5.00/day 2005 PPP line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2011/12 scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.0	93.4	0.0	6.6	6.6	–100.0
≤9	0.2	93.2	0.0	6.6	6.8	–99.6
≤14	1.7	91.7	0.0	6.6	8.3	–96.3
≤19	5.7	87.8	0.0	6.6	12.2	–87.9
≤24	13.1	80.3	0.0	6.5	19.6	–71.9
≤29	23.2	70.2	0.0	6.5	29.7	–50.3
≤34	35.1	58.4	0.1	6.5	41.5	–24.8
≤39	46.6	46.8	0.3	6.3	52.9	+0.1
≤44	58.6	34.8	0.5	6.1	64.7	+25.9
≤49	67.4	26.0	0.7	5.9	73.3	+45.0
≤54	75.1	18.4	1.2	5.4	80.5	+61.9
≤59	80.4	13.0	1.7	4.8	85.2	+74.0
≤64	84.4	9.1	2.4	4.2	88.5	+83.1
≤69	88.6	4.9	3.7	2.9	91.4	+93.4
≤74	90.9	2.6	4.3	2.3	93.1	+95.4
≤79	92.2	1.3	4.9	1.7	93.8	+94.8
≤84	92.9	0.5	6.0	0.6	93.5	+93.6
≤89	93.4	0.0	6.5	0.0	93.4	+93.0
≤94	93.4	0.0	6.6	0.0	93.4	+93.0
≤100	93.4	0.0	6.6	0.0	93.4	+93.0

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

Table 10 ($5.00/day 2005 PPP 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), 2011/12 scorecard applied to the 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	0.0	100.0	0.0	Only poor targeted
≤9	0.2	100.0	0.2	Only poor targeted
≤14	1.7	100.0	1.8	Only poor targeted
≤19	5.7	100.0	6.1	Only poor targeted
≤24	13.2	99.7	14.0	298.9:1
≤29	23.3	99.8	24.8	529.2:1
≤34	35.2	99.7	37.5	359.5:1
≤39	46.9	99.4	49.9	161.2:1
≤44	59.1	99.2	62.7	127.6:1
≤49	68.1	99.0	72.1	97.7:1
≤54	76.2	98.5	80.3	64.4:1
≤59	82.1	97.9	86.0	46.1:1
≤64	86.7	97.2	90.3	35.3:1
≤69	92.2	96.0	94.8	24.2:1
≤74	95.1	95.5	97.2	21.2:1
≤79	97.1	95.0	98.6	18.9:1
≤84	98.9	93.9	99.5	15.5:1
≤89	100.0	93.5	100.0	14.3:1
≤94	100.0	93.4	100.0	14.3:1
≤100	100.0	93.4	100.0	14.3:1

Tables for
the $1.90/day 2011 PPP Poverty Line

Table 3 ($1.90/day 2011 PPP line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	89.0
15–19	85.0
20–24	78.0
25–29	70.6
30–34	57.7
35–39	47.2
40–44	31.2
45–49	28.5
50–54	18.8
55–59	11.9
60–64	5.8
65–69	3.7
70–74	3.1
75–79	2.9
80–84	1.3
85–89	1.2
90–94	0.0
95–100	0.0

Table 5 ($1.90/day 2011 PPP line): Errors (average differences between estimated and true poverty likelihoods for households by score range), with confidence intervals, from 1,000 bootstraps of n = 16,384, 2011/12 scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	0.0	0.0	0.0	0.0
5–9	0.0	0.0	0.0	0.0
10–14	–5.1	3.6	3.9	4.3
15–19	–5.4	3.7	3.9	4.2
20–24	+0.7	2.1	2.6	3.2
25–29	+0.7	1.9	2.3	3.0
30–34	+2.4	2.0	2.4	3.4
35–39	–2.3	2.3	2.7	3.4
40–44	–1.3	1.9	2.2	3.1
45–49	+3.4	2.3	2.7	3.6
50–54	+7.5	1.5	1.7	2.3
55–59	–9.3	6.1	6.3	6.7
60–64	–2.5	2.4	2.6	3.2
65–69	+0.5	1.2	1.4	1.8
70–74	+0.8	1.1	1.3	1.7
75–79	+0.4	1.9	2.2	2.9
80–84	+0.8	0.5	0.6	0.8
85–89	+1.2	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

Table 6 ($1.90/day 2011 PPP line): Errors (average differences between estimated poverty rates and true values) for a group at a point in time by sample size, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2011/12 scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	–1.5	64.6	79.6	86.1
4	–1.1	39.9	45.0	57.2
8	–0.7	27.6	32.4	41.7
16	–0.6	20.5	24.3	30.6
32	–0.2	14.3	16.1	20.4
64	–0.4	9.8	11.7	15.9
128	+0.1	7.3	8.6	11.8
256	0.0	5.0	6.3	7.7
512	–0.2	3.7	4.4	5.9
1,024	–0.1	2.6	3.0	3.8
2,048	–0.1	1.8	2.2	2.9
4,096	0.0	1.3	1.6	2.0
8,192	0.0	0.9	1.1	1.4
16,384	0.0	0.7	0.8	1.1

Table 9 ($1.90/day 2011 PPP line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2011/12 scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.0	39.4	0.0	60.6	60.6	–100.0
≤9	0.2	39.2	0.0	60.6	60.8	–99.0
≤14	1.6	37.8	0.1	60.5	62.1	–91.5
≤19	5.2	34.2	0.5	60.1	65.3	–72.5
≤24	10.8	28.6	2.4	58.2	69.0	–39.3
≤29	17.7	21.7	5.5	55.1	72.8	+4.1
≤34	24.5	14.8	10.7	50.0	74.5	+51.6
≤39	30.2	9.2	16.7	44.0	74.2	+57.7
≤44	34.3	5.1	24.8	35.8	70.1	+37.1
≤49	36.4	3.0	31.7	28.9	65.4	+19.6
≤54	37.6	1.7	38.6	22.0	59.6	+2.0
≤59	38.7	0.7	43.5	17.1	55.8	–10.4
≤64	39.0	0.3	47.7	12.9	52.0	–21.1
≤69	39.2	0.2	53.0	7.6	46.9	–34.5
≤74	39.3	0.1	55.8	4.8	44.1	–41.7
≤79	39.4	0.0	57.7	2.9	42.3	–46.5
≤84	39.4	0.0	59.6	1.1	40.4	–51.2
≤89	39.4	0.0	60.6	0.0	39.4	–53.8
≤94	39.4	0.0	60.6	0.0	39.4	–53.9
≤100	39.4	0.0	60.6	0.0	39.4	–53.9

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

Table 10 ($1.90/day 2011 PPP 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), 2011/12 scorecard applied to the 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	0.0	100.0	0.0	Only poor targeted
≤9	0.2	100.0	0.5	Only poor targeted
≤14	1.7	93.2	4.1	13.7:1
≤19	5.7	91.3	13.1	10.5:1
≤24	13.2	81.8	27.3	4.5:1
≤29	23.3	76.2	45.0	3.2:1
≤34	35.2	69.7	62.3	2.3:1
≤39	46.9	64.5	76.8	1.8:1
≤44	59.1	58.1	87.1	1.4:1
≤49	68.1	53.5	92.5	1.1:1
≤54	76.2	49.4	95.6	1.0:1
≤59	82.1	47.1	98.2	0.9:1
≤64	86.7	45.0	99.1	0.8:1
≤69	92.2	42.5	99.6	0.7:1
≤74	95.1	41.3	99.9	0.7:1
≤79	97.1	40.6	99.9	0.7:1
≤84	98.9	39.8	100.0	0.7:1
≤89	100.0	39.4	100.0	0.6:1
≤94	100.0	39.4	100.0	0.6:1
≤100	100.0	39.4	100.0	0.6:1

Tables for
the $3.10/day 2011 PPP Poverty Line

Table 3 ($3.10/day 2011 PPP line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	100.0
15–19	99.9
20–24	96.3
25–29	94.5
30–34	91.7
35–39	84.4
40–44	71.8
45–49	66.1
50–54	54.0
55–59	45.5
60–64	33.7
65–69	28.9
70–74	21.2
75–79	20.1
80–84	6.3
85–89	6.3
90–94	6.3
95–100	0.0

Table 5 ($3.10/day 2011 PPP line): Errors (average differences between estimated and true poverty likelihoods for households by score range), with confidence intervals, from 1,000 bootstraps of n = 16,384, 2011/12 scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	0.0	0.0	0.0	0.0
5–9	0.0	0.0	0.0	0.0
10–14	0.0	0.0	0.0	0.0
15–19	+4.1	1.5	1.8	2.4
20–24	+0.2	1.0	1.1	1.6
25–29	+0.1	1.0	1.2	1.5
30–34	+3.1	1.3	1.6	2.1
35–39	–3.3	2.3	2.5	2.7
40–44	–1.8	1.9	2.2	2.9
45–49	+1.0	2.5	2.8	3.7
50–54	–2.5	2.7	3.2	4.1
55–59	–12.6	7.9	8.1	8.9
60–64	–1.8	3.7	4.3	5.5
65–69	+9.4	2.5	3.2	4.4
70–74	+8.3	2.6	3.1	4.2
75–79	+9.6	3.2	3.9	5.0
80–84	–1.5	3.5	4.2	5.2
85–89	+4.1	1.6	2.0	2.4
90–94	+6.3	0.0	0.0	0.0
95–100	0.0	0.0	0.0	0.0

Table 6 ($3.10/day 2011 PPP line): Errors (average differences between estimated poverty rates and true values) for a group at a point in time by sample size, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2011/12 scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+0.7	63.1	81.4	91.5
4	+0.2	34.8	40.0	54.8
8	–0.4	23.9	28.4	35.7
16	–0.6	17.3	19.7	25.7
32	–0.2	12.3	14.4	18.1
64	–0.5	8.5	10.3	13.6
128	–0.4	6.0	7.1	9.7
256	–0.2	4.4	5.0	6.4
512	–0.2	3.0	3.5	4.7
1,024	–0.2	2.0	2.5	3.3
2,048	–0.2	1.5	1.8	2.3
4,096	–0.2	1.0	1.2	1.6
8,192	–0.2	0.7	0.9	1.2
16,384	–0.2	0.5	0.7	0.9

Table 9 ($3.10/day 2011 PPP line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2011/12 scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.0	69.7	0.0	30.3	30.3	–100.0
≤9	0.2	69.5	0.0	30.3	30.5	–99.4
≤14	1.7	68.0	0.0	30.3	32.0	–95.1
≤19	5.5	64.2	0.1	30.2	35.7	–84.0
≤24	12.7	57.0	0.5	29.8	42.5	–62.9
≤29	22.3	47.4	1.0	29.3	51.5	–34.7
≤34	33.0	36.7	2.2	28.1	61.0	–2.2
≤39	43.2	26.5	3.7	26.6	69.8	+29.3
≤44	52.5	17.2	6.6	23.8	76.3	+60.1
≤49	58.4	11.3	9.7	20.6	79.0	+81.5
≤54	62.9	6.8	13.3	17.0	79.9	+80.9
≤59	66.0	3.7	16.1	14.2	80.2	+76.9
≤64	67.6	2.1	19.1	11.2	78.8	+72.5
≤69	68.8	0.9	23.4	6.9	75.7	+66.4
≤74	69.3	0.4	25.8	4.5	73.8	+63.0
≤79	69.5	0.2	27.5	2.8	72.3	+60.5
≤84	69.7	0.0	29.3	1.0	70.7	+58.0
≤89	69.7	0.0	30.3	0.0	69.7	+56.6
≤94	69.7	0.0	30.3	0.0	69.7	+56.5
≤100	69.7	0.0	30.3	0.0	69.7	+56.5

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

Table 10 ($3.10/day 2011 PPP 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), 2011/12 scorecard applied to the 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	0.0	100.0	0.0	Only poor targeted
≤9	0.2	100.0	0.3	Only poor targeted
≤14	1.7	100.0	2.5	Only poor targeted
≤19	5.7	97.7	7.9	41.8:1
≤24	13.2	96.5	18.2	27.8:1
≤29	23.3	95.7	31.9	22.1:1
≤34	35.2	93.7	47.3	14.8:1
≤39	46.9	92.2	62.0	11.8:1
≤44	59.1	88.9	75.3	8.0:1
≤49	68.1	85.8	83.8	6.0:1
≤54	76.2	82.5	90.3	4.7:1
≤59	82.1	80.4	94.8	4.1:1
≤64	86.7	77.9	97.0	3.5:1
≤69	92.2	74.6	98.7	2.9:1
≤74	95.1	72.9	99.5	2.7:1
≤79	97.1	71.7	99.8	2.5:1
≤84	98.9	70.4	99.9	2.4:1
≤89	100.0	69.7	100.0	2.3:1
≤94	100.0	69.7	100.0	2.3:1
≤100	100.0	69.7	100.0	2.3:1

Tables for
the $3.80/day 2011 PPP Poverty Line

Table 3 ($3.80/day 2011 PPP line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	100.0
15–19	99.9
20–24	97.4
25–29	97.4
30–34	96.7
35–39	92.5
40–44	85.4
45–49	81.8
50–54	69.7
55–59	63.1
60–64	50.2
65–69	44.7
70–74	34.4
75–79	33.0
80–84	11.4
85–89	11.4
90–94	7.6
95–100	0.0

Table 5 ($3.80/day 2011 PPP line): Errors (average differences between estimated and true poverty likelihoods for households by score range), with confidence intervals, from 1,000 bootstraps of n = 16,384, 2011/12 scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	0.0	0.0	0.0	0.0
5–9	0.0	0.0	0.0	0.0
10–14	0.0	0.0	0.0	0.0
15–19	+2.6	1.3	1.6	2.1
20–24	0.0	0.8	1.0	1.4
25–29	+0.7	0.8	1.0	1.3
30–34	+4.7	1.2	1.4	1.8
35–39	–1.4	1.2	1.3	1.6
40–44	–4.9	3.0	3.2	3.4
45–49	–0.2	1.9	2.3	2.9
50–54	–2.3	2.3	2.8	3.8
55–59	–7.2	5.1	5.2	5.9
60–64	–5.5	4.6	4.9	5.6
65–69	+12.1	3.1	3.6	5.1
70–74	+5.1	4.1	4.9	6.9
75–79	+6.7	4.8	5.7	7.6
80–84	–1.2	4.1	4.8	5.9
85–89	+6.0	2.6	3.1	4.2
90–94	+7.6	0.0	0.0	0.0
95–100	0.0	0.0	0.0	0.0

Table 6 ($3.80/day 2011 PPP line): Errors (average differences between estimated poverty rates and true values) for a group at a point in time by sample size, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2011/12 scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+1.3	65.8	72.9	93.0
4	+0.5	27.6	36.3	46.5
8	–0.1	19.1	23.1	32.2
16	0.0	13.6	17.1	21.7
32	+0.1	10.1	12.0	15.3
64	–0.4	6.7	8.2	11.6
128	–0.2	4.9	6.1	8.0
256	–0.1	3.6	4.4	5.7
512	–0.1	2.6	3.1	3.8
1,024	–0.1	1.9	2.2	3.0
2,048	–0.1	1.3	1.6	1.8
4,096	–0.1	0.9	1.1	1.4
8,192	–0.1	0.7	0.8	1.0
16,384	–0.1	0.5	0.5	0.7

Table 9 ($3.80/day 2011 PPP line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2011/12 scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.0	79.0	0.0	21.0	21.0	–100.0
≤9	0.2	78.8	0.0	21.0	21.2	–99.5
≤14	1.7	77.2	0.0	21.0	22.8	–95.6
≤19	5.6	73.4	0.1	21.0	26.6	–85.8
≤24	12.9	66.1	0.2	20.8	33.7	–67.0
≤29	22.7	56.3	0.5	20.5	43.2	–41.8
≤34	33.8	45.1	1.3	19.7	53.5	–12.6
≤39	44.9	34.1	2.0	19.0	63.9	+16.2
≤44	55.8	23.1	3.2	17.8	73.6	+45.5
≤49	63.1	15.9	5.0	16.0	79.1	+66.1
≤54	68.9	10.1	7.4	13.6	82.5	+83.8
≤59	72.8	6.1	9.3	11.7	84.6	+88.2
≤64	75.1	3.9	11.6	9.4	84.5	+85.3
≤69	77.1	1.8	15.1	5.9	83.1	+80.9
≤74	78.2	0.8	17.0	4.1	82.2	+78.5
≤79	78.7	0.3	18.4	2.6	81.3	+76.7
≤84	78.9	0.1	20.1	1.0	79.9	+74.6
≤89	79.0	0.0	21.0	0.0	79.0	+73.4
≤94	79.0	0.0	21.0	0.0	79.0	+73.4
≤100	79.0	0.0	21.0	0.0	79.0	+73.4

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

Table 10 ($3.80/day 2011 PPP 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), 2011/12 scorecard applied to the 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	0.0	100.0	0.0	Only poor targeted
≤9	0.2	100.0	0.3	Only poor targeted
≤14	1.7	100.0	2.2	Only poor targeted
≤19	5.7	98.8	7.1	85.3:1
≤24	13.2	98.1	16.3	52.4:1
≤29	23.3	97.6	28.8	41.5:1
≤34	35.2	96.2	42.9	25.3:1
≤39	46.9	95.7	56.8	22.3:1
≤44	59.1	94.5	70.7	17.3:1
≤49	68.1	92.6	79.9	12.5:1
≤54	76.2	90.3	87.2	9.3:1
≤59	82.1	88.7	92.2	7.8:1
≤64	86.7	86.6	95.1	6.5:1
≤69	92.2	83.6	97.7	5.1:1
≤74	95.1	82.2	99.0	4.6:1
≤79	97.1	81.1	99.6	4.3:1
≤84	98.9	79.7	99.9	3.9:1
≤89	100.0	79.0	100.0	3.8:1
≤94	100.0	79.0	100.0	3.8:1
≤100	100.0	79.0	100.0	3.8:1

Tables for
the $4.00/day 2011 PPP Poverty Line

Table 3 ($4.00/day 2011 PPP line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	100.0
15–19	99.9
20–24	98.3
25–29	98.0
30–34	97.7
35–39	94.1
40–44	88.9
45–49	84.1
50–54	72.2
55–59	65.3
60–64	53.8
65–69	46.8
70–74	37.2
75–79	35.7
80–84	13.4
85–89	13.4
90–94	7.6
95–100	0.0

Table 5 ($4.00/day 2011 PPP line): Errors (average differences between estimated and true poverty likelihoods for households by score range), with confidence intervals, from 1,000 bootstraps of n = 16,384, 2011/12 scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	0.0	0.0	0.0	0.0
5–9	0.0	0.0	0.0	0.0
10–14	0.0	0.0	0.0	0.0
15–19	+2.6	1.3	1.6	2.1
20–24	+0.7	0.8	1.0	1.3
25–29	+0.8	0.8	0.9	1.1
30–34	+3.8	1.0	1.2	1.5
35–39	–0.3	1.0	1.2	1.5
40–44	–3.2	2.1	2.2	2.4
45–49	+1.5	1.9	2.3	3.0
50–54	–5.2	3.7	3.9	4.4
55–59	–7.7	5.3	5.5	6.1
60–64	–5.0	4.2	4.5	5.5
65–69	+9.7	3.1	3.7	5.4
70–74	+5.5	4.1	5.0	6.8
75–79	+8.5	4.9	5.9	7.5
80–84	–1.0	4.2	4.9	6.1
85–89	+8.0	2.6	3.1	4.2
90–94	+7.6	0.0	0.0	0.0
95–100	0.0	0.0	0.0	0.0

Table 6 ($4.00/day 2011 PPP line): Errors (average differences between estimated poverty rates and true values) for a group at a point in time by sample size, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2011/12 scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+1.7	68.6	75.9	92.4
4	+1.0	27.3	35.7	47.8
8	+0.2	18.4	22.7	29.9
16	+0.2	12.6	15.0	21.8
32	+0.2	9.4	11.5	15.3
64	–0.1	6.4	7.9	10.7
128	0.0	4.9	5.7	8.0
256	+0.1	3.4	4.1	5.4
512	+0.1	2.5	3.0	3.7
1,024	+0.1	1.7	2.1	2.8
2,048	+0.1	1.2	1.4	1.8
4,096	+0.1	0.8	1.0	1.3
8,192	+0.1	0.6	0.7	0.9
16,384	+0.1	0.4	0.5	0.7

Table 9 ($4.00/day 2011 PPP line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2011/12 scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.0	80.6	0.0	19.4	19.4	–100.0
≤9	0.2	80.4	0.0	19.4	19.6	–99.5
≤14	1.7	78.8	0.0	19.4	21.2	–95.7
≤19	5.6	75.0	0.1	19.4	25.0	–86.0
≤24	12.9	67.6	0.2	19.2	32.2	–67.6
≤29	22.8	57.8	0.5	19.0	41.8	–42.8
≤34	34.1	46.5	1.1	18.3	52.4	–14.0
≤39	45.2	35.4	1.7	17.7	62.9	+14.3
≤44	56.3	24.3	2.8	16.6	72.9	+43.2
≤49	63.6	17.0	4.5	14.9	78.5	+63.5
≤54	69.7	10.9	6.6	12.9	82.5	+81.1
≤59	73.9	6.7	8.3	11.2	85.0	+89.7
≤64	76.3	4.3	10.4	9.0	85.3	+87.0
≤69	78.5	2.0	13.7	5.8	84.3	+83.0
≤74	79.7	0.9	15.5	4.0	83.7	+80.8
≤79	80.2	0.3	16.8	2.6	82.8	+79.1
≤84	80.5	0.1	18.5	1.0	81.5	+77.1
≤89	80.6	0.0	19.4	0.0	80.6	+75.9
≤94	80.6	0.0	19.4	0.0	80.6	+75.9
≤100	80.6	0.0	19.4	0.0	80.6	+75.9

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

Table 10 ($4.00/day 2011 PPP 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), 2011/12 scorecard applied to the 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	0.0	100.0	0.0	Only poor targeted
≤9	0.2	100.0	0.3	Only poor targeted
≤14	1.7	100.0	2.1	Only poor targeted
≤19	5.7	98.8	6.9	85.3:1
≤24	13.2	98.4	16.1	60.5:1
≤29	23.3	98.0	28.3	48.4:1
≤34	35.2	96.9	42.3	30.8:1
≤39	46.9	96.3	56.1	26.0:1
≤44	59.1	95.3	69.8	20.1:1
≤49	68.1	93.4	78.9	14.2:1
≤54	76.2	91.4	86.5	10.6:1
≤59	82.1	89.9	91.7	8.9:1
≤64	86.7	88.0	94.7	7.3:1
≤69	92.2	85.2	97.5	5.7:1
≤74	95.1	83.8	98.9	5.2:1
≤79	97.1	82.7	99.6	4.8:1
≤84	98.9	81.3	99.9	4.4:1
≤89	100.0	80.6	100.0	4.2:1
≤94	100.0	80.6	100.0	4.1:1
≤100	100.0	80.6	100.0	4.1:1

Tables for
the Poverty Line Marking the Poorest Half of People
below 100% of the National Poverty Line

Table 3 (Poorest half below 100% of 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	80.6
10–14	49.2
15–19	42.6
20–24	25.6
25–29	20.6
30–34	16.9
35–39	8.1
40–44	5.4
45–49	3.1
50–54	1.7
55–59	0.5
60–64	0.5
65–69	0.3
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Table 5 (Poorest half below 100% of national line): Errors (average differences between estimated and true poverty likelihoods for households by score range), with confidence intervals, from 1,000 bootstraps of n = 16,384, 2011/12 scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	0.0	0.0	0.0	0.0
5–9	+8.5	13.2	16.3	21.5
10–14	+8.4	5.6	6.5	8.5
15–19	–11.2	7.3	7.6	8.2
20–24	–3.2	2.8	3.0	3.9
25–29	–2.7	2.3	2.5	3.0
30–34	+5.9	1.2	1.4	1.8
35–39	–2.1	1.8	1.9	2.2
40–44	–0.5	1.2	1.4	1.9
45–49	+0.3	0.7	0.8	1.1
50–54	–0.1	0.6	0.7	1.0
55–59	–0.3	0.5	0.6	0.8
60–64	+0.5	0.0	0.0	0.0
65–69	+0.3	0.0	0.0	0.0
70–74	–0.9	0.9	1.0	1.3
75–79	0.0	0.0	0.0	0.0
80–84	0.0	0.0	0.0	0.0
85–89	0.0	0.0	0.0	0.0
90–94	0.0	0.0	0.0	0.0
95–100	0.0	0.0	0.0	0.0

Table 6 (Poorest half below 100% of national line): Errors (average differences between estimated poverty rates and true values) for a group at a point in time by sample size, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2011/12 scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	–0.2	52.5	58.6	73.0
4	–0.6	28.2	34.7	45.5
8	–0.6	21.3	26.3	34.6
16	–0.9	15.3	18.4	22.9
32	–0.5	10.9	13.2	17.0
64	–0.5	8.3	9.8	12.1
128	–0.5	5.7	6.7	9.0
256	–0.6	4.0	4.7	6.6
512	–0.5	2.9	3.4	4.6
1,024	–0.5	2.0	2.4	3.1
2,048	–0.5	1.4	1.7	2.2
4,096	–0.5	1.0	1.2	1.6
8,192	–0.5	0.7	0.9	1.2
16,384	–0.5	0.5	0.6	0.8

Table 9 (Poorest half below 100% of national line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2011/12 scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.0	10.4	0.0	89.6	89.6	–100.0
≤9	0.1	10.3	0.1	89.5	89.6	–96.7
≤14	0.8	9.6	0.9	88.6	89.5	–75.7
≤19	2.8	7.7	2.9	86.7	89.4	–19.3
≤24	4.7	5.7	8.5	81.1	85.8	+19.0
≤29	7.0	3.5	16.3	73.3	80.3	–55.9
≤34	8.4	2.0	26.7	62.8	71.3	–156.1
≤39	9.4	1.0	37.4	52.1	61.6	–258.5
≤44	9.9	0.5	49.1	40.4	50.4	–370.5
≤49	10.2	0.2	57.9	31.7	41.9	–454.1
≤54	10.4	0.1	65.9	23.7	34.0	–530.9
≤59	10.4	0.0	71.7	17.8	28.2	–586.9
≤64	10.4	0.0	76.3	13.2	23.7	–630.9
≤69	10.4	0.0	81.8	7.8	18.2	–683.2
≤74	10.4	0.0	84.7	4.9	15.3	–711.0
≤79	10.4	0.0	86.6	2.9	13.4	–729.4
≤84	10.4	0.0	88.5	1.1	11.5	–747.4
≤89	10.4	0.0	89.5	0.0	10.5	–757.2
≤94	10.4	0.0	89.6	0.0	10.4	–757.6
≤100	10.4	0.0	89.6	0.0	10.4	–757.6

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

Table 10 (Poorest half below 100% of 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), 2011/12 scorecard applied to the 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	0.0	100.0	0.0	Only poor targeted
≤9	0.2	68.8	1.3	2.2:1
≤14	1.7	47.0	7.8	0.9:1
≤19	5.7	48.9	26.5	1.0:1
≤24	13.2	35.7	45.0	0.6:1
≤29	23.3	30.0	66.8	0.4:1
≤34	35.2	24.0	80.8	0.3:1
≤39	46.9	20.2	90.5	0.3:1
≤44	59.1	16.8	95.1	0.2:1
≤49	68.1	15.0	98.0	0.2:1
≤54	76.2	13.6	99.2	0.2:1
≤59	82.1	12.7	99.8	0.1:1
≤64	86.7	12.0	99.8	0.1:1
≤69	92.2	11.3	99.8	0.1:1
≤74	95.1	11.0	100.0	0.1:1
≤79	97.1	10.8	100.0	0.1:1
≤84	98.9	10.6	100.0	0.1:1
≤89	100.0	10.4	100.0	0.1:1
≤94	100.0	10.4	100.0	0.1:1
≤100	100.0	10.4	100.0	0.1:1

Tables for
the First-Quintile (20th-Percentile) Poverty Line

Table 3 (First-quintile (20th-percentile) 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	76.6
15–19	53.2
20–24	38.3
25–29	28.0
30–34	23.2
35–39	12.5
40–44	7.9
45–49	5.4
50–54	3.4
55–59	0.6
60–64	0.6
65–69	0.4
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Table 5 (First-quintile (20th-percentile) line): Errors (average differences between estimated and true poverty likelihoods for households by score range), with confidence intervals, from 1,000 bootstraps of n = 16,384, 2011/12 scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	0.0	0.0	0.0	0.0
5–9	0.0	0.0	0.0	0.0
10–14	+25.1	5.6	6.7	8.3
15–19	–10.5	6.8	7.2	7.8
20–24	+0.7	2.6	3.0	4.0
25–29	–5.4	3.7	4.0	4.4
30–34	+7.8	1.4	1.6	2.1
35–39	–1.2	1.5	1.7	2.3
40–44	–0.7	1.4	1.6	2.0
45–49	–0.1	1.1	1.3	1.6
50–54	+0.3	0.8	1.0	1.3
55–59	–0.3	0.5	0.6	0.8
60–64	+0.6	0.0	0.0	0.0
65–69	–0.8	0.8	1.0	1.1
70–74	–0.9	0.9	1.0	1.3
75–79	0.0	0.0	0.0	0.0
80–84	0.0	0.0	0.0	0.0
85–89	0.0	0.0	0.0	0.0
90–94	0.0	0.0	0.0	0.0
95–100	0.0	0.0	0.0	0.0

Table 6 (First-quintile (20th-percentile) line): Errors (average differences between estimated poverty rates and true values) for a group at a point in time by sample size, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2011/12 scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+0.4	57.5	62.9	85.6
4	–0.5	30.8	37.7	50.7
8	–0.3	24.4	30.1	37.6
16	0.0	17.3	20.1	26.8
32	+0.1	12.2	14.2	18.4
64	+0.2	9.1	10.5	14.0
128	+0.1	6.2	7.4	9.4
256	0.0	4.3	5.1	7.2
512	+0.1	3.1	3.7	4.8
1,024	+0.1	2.2	2.7	3.5
2,048	+0.2	1.6	1.8	2.3
4,096	+0.2	1.1	1.3	1.7
8,192	+0.2	0.8	0.9	1.2
16,384	+0.2	0.6	0.7	0.8

Table 9 (First-quintile (20th-percentile) line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2011/12 scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.0	14.4	0.0	85.5	85.5	–100.0
≤9	0.2	14.2	0.0	85.5	85.7	–97.2
≤14	1.0	13.4	0.6	84.9	85.9	–81.1
≤19	3.4	11.0	2.2	83.3	86.7	–37.5
≤24	5.9	8.5	7.2	78.3	84.3	+31.9
≤29	9.2	5.2	14.0	71.5	80.7	+3.0
≤34	11.3	3.1	23.8	61.7	73.0	–64.8
≤39	12.8	1.6	34.1	51.5	64.3	–135.7
≤44	13.6	0.8	45.4	40.1	53.7	–214.4
≤49	14.1	0.4	54.0	31.5	45.6	–273.6
≤54	14.3	0.1	61.9	23.6	38.0	–328.3
≤59	14.4	0.1	67.7	17.8	32.2	–368.7
≤64	14.4	0.1	72.3	13.2	27.6	–400.5
≤69	14.4	0.0	77.7	7.8	22.2	–438.1
≤74	14.4	0.0	80.7	4.9	19.3	–458.2
≤79	14.4	0.0	82.6	2.9	17.4	–471.5
≤84	14.4	0.0	84.4	1.1	15.5	–484.5
≤89	14.4	0.0	85.5	0.0	14.5	–491.6
≤94	14.4	0.0	85.5	0.0	14.5	–491.9
≤100	14.4	0.0	85.5	0.0	14.4	–491.9

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

Table 10 (First-quintile (20th-percentile) 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), 2011/12 scorecard applied to the 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	0.0	100.0	0.0	Only poor targeted
≤9	0.2	100.0	1.4	Only poor targeted
≤14	1.7	60.7	7.2	1.5:1
≤19	5.7	60.3	23.6	1.5:1
≤24	13.2	45.1	41.1	0.8:1
≤29	23.3	39.6	63.7	0.7:1
≤34	35.2	32.2	78.5	0.5:1
≤39	46.9	27.3	88.6	0.4:1
≤44	59.1	23.0	94.1	0.3:1
≤49	68.1	20.7	97.4	0.3:1
≤54	76.2	18.8	99.2	0.2:1
≤59	82.1	17.5	99.6	0.2:1
≤64	86.7	16.6	99.6	0.2:1
≤69	92.2	15.6	99.9	0.2:1
≤74	95.1	15.2	100.0	0.2:1
≤79	97.1	14.9	100.0	0.2:1
≤84	98.9	14.6	100.0	0.2:1
≤89	100.0	14.5	100.0	0.2:1
≤94	100.0	14.4	100.0	0.2:1
≤100	100.0	14.4	100.0	0.2:1

Tables for
the Second-Quintile (40^{th}-Percentile) Poverty Line

Table 3 (Second-quintile (40^{th}-percentile) 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	89.0
15–19	78.8
20–24	71.8
25–29	58.0
30–34	48.4
35–39	36.3
40–44	23.2
45–49	17.9
50–54	12.1
55–59	6.0
60–64	1.8
65–69	1.1
70–74	1.1
75–79	1.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Table 5 (Second-quintile (40th-percentile) line): Errors (average differences between estimated and true poverty likelihoods for households by score range), with confidence intervals, from 1,000 bootstraps of n = 16,384, 2011/12 scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	0.0	0.0	0.0	0.0
5–9	0.0	0.0	0.0	0.0
10–14	+2.9	3.7	4.5	5.6
15–19	–9.3	5.7	5.9	6.2
20–24	+5.0	2.3	2.7	3.7
25–29	–2.1	2.1	2.5	3.3
30–34	+3.1	1.9	2.3	3.3
35–39	–6.7	4.4	4.5	4.9
40–44	+0.3	1.7	2.1	2.8
45–49	+4.1	1.8	2.1	2.9
50–54	+5.0	1.2	1.4	2.0
55–59	–7.3	4.9	5.2	5.5
60–64	–0.4	1.0	1.1	1.6
65–69	–1.4	1.3	1.4	1.7
70–74	–0.3	1.0	1.1	1.4
75–79	–1.6	1.9	2.2	2.9
80–84	0.0	0.0	0.0	0.0
85–89	0.0	0.0	0.0	0.0
90–94	0.0	0.0	0.0	0.0
95–100	0.0	0.0	0.0	0.0

Table 6 (Second-quintile (40th-percentile) line): Errors (average differences between estimated poverty rates and true values) for a group at a point in time by sample size, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2011/12 scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	–0.6	67.4	76.9	87.5
4	–1.1	39.0	44.6	54.3
8	–0.5	28.2	33.3	39.9
16	–0.6	20.8	24.6	30.5
32	–0.2	14.3	17.2	20.9
64	–0.4	10.1	12.4	15.7
128	–0.2	7.1	8.7	11.6
256	–0.4	5.1	6.0	7.8
512	–0.5	3.6	4.4	5.8
1,024	–0.4	2.5	2.9	3.9
2,048	–0.4	1.8	2.1	2.7
4,096	–0.3	1.3	1.5	2.1
8,192	–0.3	1.0	1.1	1.5
16,384	–0.3	0.7	0.8	1.1

Table 9 (Second-quintile (40th-percentile) line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2011/12 scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.0	31.2	0.0	68.7	68.7	–100.0
≤9	0.2	31.0	0.0	68.7	68.9	–98.7
≤14	1.5	29.7	0.2	68.5	70.0	–89.7
≤19	4.9	26.3	0.7	68.0	73.0	–66.1
≤24	9.8	21.5	3.4	65.3	75.1	–26.6
≤29	15.7	15.5	7.5	61.2	76.9	+24.8
≤34	21.2	10.0	14.0	54.8	76.0	+55.3
≤39	25.9	5.4	21.0	47.8	73.6	+32.9
≤44	28.6	2.6	30.4	38.3	66.9	+2.6
≤49	29.8	1.5	38.3	30.5	60.2	–22.5
≤54	30.4	0.8	45.8	23.0	53.4	–46.5
≤59	30.9	0.3	51.2	17.6	48.5	–63.8
≤64	31.0	0.2	55.7	13.1	44.1	–78.2
≤69	31.2	0.1	61.0	7.7	38.9	–95.3
≤74	31.2	0.0	63.9	4.8	36.0	–104.5
≤79	31.2	0.0	65.8	2.9	34.2	–110.6
≤84	31.2	0.0	67.7	1.1	32.3	–116.6
≤89	31.2	0.0	68.7	0.0	31.3	–119.9
≤94	31.2	0.0	68.7	0.0	31.2	–120.0
≤100	31.2	0.0	68.7	0.0	31.2	–120.0

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

Table 10 (Second-quintile (40th-percentile) 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), 2011/12 scorecard applied to the 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	0.0	100.0	0.0	Only poor targeted
≤9	0.2	100.0	0.6	Only poor targeted
≤14	1.7	86.9	4.8	6.6:1
≤19	5.7	87.3	15.8	6.9:1
≤24	13.2	74.2	31.2	2.9:1
≤29	23.3	67.6	50.3	2.1:1
≤34	35.2	60.3	68.0	1.5:1
≤39	46.9	55.2	82.9	1.2:1
≤44	59.1	48.5	91.6	0.9:1
≤49	68.1	43.7	95.3	0.8:1
≤54	76.2	39.9	97.5	0.7:1
≤59	82.1	37.7	99.1	0.6:1
≤64	86.7	35.8	99.4	0.6:1
≤69	92.2	33.8	99.7	0.5:1
≤74	95.1	32.8	99.9	0.5:1
≤79	97.1	32.2	100.0	0.5:1
≤84	98.9	31.6	100.0	0.5:1
≤89	100.0	31.3	100.0	0.5:1
≤94	100.0	31.2	100.0	0.5:1
≤100	100.0	31.2	100.0	0.5:1

Tables for the Median (50^{th}-Percentile) Poverty Line

Table 3 (Median (50th-percentile) 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	90.1
15–19	88.1
20–24	80.1
25–29	73.7
30–34	60.8
35–39	48.9
40–44	33.0
45–49	27.4
50–54	17.6
55–59	8.6
60–64	4.4
65–69	1.3
70–74	1.3
75–79	1.3
80–84	1.3
85–89	0.1
90–94	0.0
95–100	0.0

Table 5 (Median (50th-percentile) line): Errors (average differences between estimated and true poverty likelihoods for households by score range), with confidence intervals, from 1,000 bootstraps of n = 16,384, 2011/12 scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	0.0	0.0	0.0	0.0
5–9	0.0	0.0	0.0	0.0
10–14	–3.8	3.0	3.3	3.6
15–19	–2.3	2.2	2.4	3.0
20–24	–0.3	2.0	2.3	3.3
25–29	–2.5	2.1	2.2	2.7
30–34	+3.3	2.0	2.4	3.4
35–39	–3.4	2.8	3.0	3.5
40–44	–1.2	1.9	2.3	3.0
45–49	+0.1	2.3	2.8	3.8
50–54	+4.9	1.7	1.9	2.4
55–59	–17.4	10.2	10.5	10.9
60–64	–2.3	2.2	2.5	3.0
65–69	–1.3	1.2	1.3	1.7
70–74	–0.6	1.0	1.2	1.7
75–79	–3.8	3.2	3.5	4.1
80–84	+1.1	0.3	0.4	0.5
85–89	+0.1	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

Table 6 (Median (50th-percentile) line): Errors (average differences between estimated poverty rates and true values) for a group at a point in time by sample size, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2011/12 scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	–1.2	63.9	81.2	87.8
4	–1.7	39.3	45.1	55.0
8	–2.1	28.4	32.8	42.1
16	–1.9	21.1	24.5	31.2
32	–1.3	14.7	17.1	20.8
64	–1.6	9.9	11.7	16.0
128	–1.3	7.3	8.8	11.1
256	–1.4	5.2	6.1	7.8
512	–1.4	3.7	4.4	5.7
1,024	–1.4	2.6	3.1	3.9
2,048	–1.4	1.8	2.2	2.8
4,096	–1.3	1.3	1.6	2.1
8,192	–1.3	0.9	1.1	1.4
16,384	–1.3	0.7	0.8	1.0

Table 9 (Median (50th-percentile) line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2011/12 scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.0	40.5	0.0	59.5	59.5	–100.0
≤9	0.2	40.3	0.0	59.5	59.7	–99.0
≤14	1.6	38.9	0.1	59.4	61.0	–91.8
≤19	5.1	35.3	0.5	59.0	64.1	–73.3
≤24	11.0	29.4	2.1	57.4	68.5	–40.2
≤29	18.5	22.0	4.8	54.8	73.3	+3.2
≤34	25.6	14.9	9.6	49.9	75.5	+50.2
≤39	31.5	8.9	15.3	44.2	75.7	+62.1
≤44	35.7	4.8	23.4	36.1	71.8	+42.2
≤49	37.8	2.7	30.3	29.2	67.1	+25.2
≤54	38.9	1.5	37.3	22.2	61.2	+7.8
≤59	40.0	0.5	42.2	17.4	57.3	–4.2
≤64	40.2	0.3	46.5	13.0	53.2	–15.0
≤69	40.3	0.1	51.9	7.7	48.0	–28.2
≤74	40.4	0.1	54.7	4.8	45.2	–35.2
≤79	40.5	0.0	56.6	2.9	43.4	–39.8
≤84	40.5	0.0	58.5	1.1	41.5	–44.4
≤89	40.5	0.0	59.5	0.0	40.5	–47.0
≤94	40.5	0.0	59.5	0.0	40.5	–47.1
≤100	40.5	0.0	59.5	0.0	40.5	–47.1

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

Table 10 (Median (50th-percentile) 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), 2011/12 scorecard applied to the 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	0.0	100.0	0.0	Only poor targeted
≤9	0.2	100.0	0.5	Only poor targeted
≤14	1.7	92.6	3.9	12.5:1
≤19	5.7	90.8	12.7	9.9:1
≤24	13.2	83.9	27.3	5.2:1
≤29	23.3	79.6	45.7	3.9:1
≤34	35.2	72.8	63.3	2.7:1
≤39	46.9	67.3	77.9	2.1:1
≤44	59.1	60.4	88.1	1.5:1
≤49	68.1	55.5	93.4	1.2:1
≤54	76.2	51.1	96.2	1.0:1
≤59	82.1	48.7	98.8	0.9:1
≤64	86.7	46.4	99.3	0.9:1
≤69	92.2	43.7	99.7	0.8:1
≤74	95.1	42.5	99.8	0.7:1
≤79	97.1	41.7	100.0	0.7:1
≤84	98.9	40.9	100.0	0.7:1
≤89	100.0	40.5	100.0	0.7:1
≤94	100.0	40.5	100.0	0.7:1
≤100	100.0	40.5	100.0	0.7:1

Tables for
the Third-Quintile (60th-Percentile) Poverty Line

Table 3 (Third-quintile (60^{th}-percentile) 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	98.6
15–19	95.0
20–24	87.6
25–29	83.9
30–34	74.2
35–39	62.4
40–44	45.4
45–49	39.4
50–54	26.8
55–59	18.8
60–64	10.3
65–69	5.5
70–74	2.6
75–79	2.6
80–84	1.4
85–89	0.1
90–94	0.0
95–100	0.0

Table 5 (Third-quintile (60th-percentile) line): Errors (average differences between estimated and true poverty likelihoods for households by score range), with confidence intervals, from 1,000 bootstraps of n = 16,384, 2011/12 scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	0.0	0.0	0.0	0.0
5–9	0.0	0.0	0.0	0.0
10–14	+0.1	1.0	1.1	1.4
15–19	+1.8	1.8	2.1	2.9
20–24	+1.7	1.7	2.0	2.8
25–29	+0.2	1.5	1.8	2.3
30–34	+2.1	1.8	2.2	2.9
35–39	–7.1	4.6	4.7	5.1
40–44	–3.3	2.6	2.9	3.2
45–49	+1.4	2.4	3.0	3.9
50–54	+2.1	2.4	2.7	3.6
55–59	–15.4	9.3	9.5	9.8
60–64	0.0	2.4	2.8	3.5
65–69	–1.4	1.7	2.1	2.7
70–74	+0.4	1.1	1.3	1.7
75–79	–2.5	2.6	3.1	3.9
80–84	+0.1	1.2	1.3	1.7
85–89	+0.1	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

Table 6 (Third-quintile (60th-percentile) line): Errors (average differences between estimated poverty rates and true values) for a group at a point in time by sample size, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2011/12 scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	0.0	67.4	82.6	89.9
4	–1.4	39.8	46.4	58.1
8	–1.5	28.2	32.8	42.5
16	–1.3	20.4	24.3	30.3
32	–1.1	14.6	17.4	22.3
64	–1.4	9.7	12.1	15.0
128	–1.3	6.9	8.0	10.7
256	–1.3	5.0	6.0	7.4
512	–1.3	3.6	4.1	5.4
1,024	–1.4	2.6	3.1	3.9
2,048	–1.4	1.8	2.1	2.6
4,096	–1.3	1.3	1.5	2.0
8,192	–1.3	0.9	1.1	1.5
16,384	–1.3	0.6	0.8	1.0

Table 9 (Third-quintile (60th-percentile) line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2011/12 scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.0	49.8	0.0	50.2	50.2	–100.0
≤9	0.2	49.6	0.0	50.2	50.4	–99.2
≤14	1.7	48.1	0.0	50.1	51.8	–93.2
≤19	5.4	44.4	0.3	49.9	55.3	–77.9
≤24	11.7	38.1	1.4	48.8	60.5	–50.1
≤29	20.1	29.7	3.2	47.0	67.1	–13.0
≤34	28.8	21.0	6.4	43.8	72.5	+28.4
≤39	36.6	13.2	10.2	39.9	76.6	+67.7
≤44	42.5	7.3	16.6	33.6	76.0	+66.7
≤49	45.6	4.2	22.5	27.7	73.3	+54.9
≤54	47.5	2.3	28.8	21.4	68.9	+42.2
≤59	48.9	0.9	33.2	17.0	65.9	+33.3
≤64	49.3	0.5	37.4	12.8	62.1	+24.9
≤69	49.6	0.2	42.6	7.6	57.2	+14.5
≤74	49.7	0.1	45.4	4.8	54.5	+8.8
≤79	49.8	0.0	47.3	2.9	52.7	+5.1
≤84	49.8	0.0	49.1	1.1	50.9	+1.4
≤89	49.8	0.0	50.1	0.0	49.8	–0.7
≤94	49.8	0.0	50.2	0.0	49.8	–0.8
≤100	49.8	0.0	50.2	0.0	49.8	–0.8

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

Table 10 (Third-quintile (60th-percentile) 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), 2011/12 scorecard applied to the 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	0.0	100.0	0.0	Only poor targeted
≤9	0.2	100.0	0.4	Only poor targeted
≤14	1.7	97.5	3.4	39.2:1
≤19	5.7	94.9	10.8	18.5:1
≤24	13.2	89.0	23.5	8.1:1
≤29	23.3	86.3	40.3	6.3:1
≤34	35.2	81.7	57.8	4.5:1
≤39	46.9	78.1	73.6	3.6:1
≤44	59.1	71.9	85.3	2.6:1
≤49	68.1	67.0	91.6	2.0:1
≤54	76.2	62.2	95.3	1.6:1
≤59	82.1	59.5	98.2	1.5:1
≤64	86.7	56.9	99.0	1.3:1
≤69	92.2	53.8	99.6	1.2:1
≤74	95.1	52.2	99.8	1.1:1
≤79	97.1	51.3	99.9	1.1:1
≤84	98.9	50.3	100.0	1.0:1
≤89	100.0	49.8	100.0	1.0:1
≤94	100.0	49.8	100.0	1.0:1
≤100	100.0	49.8	100.0	1.0:1

Tables for
the Fourth-Quintile (80th-Percentile) Poverty Line

Table 3 (Fourth-quintile (80th-percentile) 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	98.6
15–19	95.0
20–24	87.6
25–29	83.9
30–34	74.2
35–39	62.4
40–44	45.4
45–49	39.4
50–54	26.8
55–59	18.8
60–64	10.3
65–69	5.5
70–74	2.6
75–79	2.6
80–84	1.4
85–89	0.1
90–94	0.0
95–100	0.0

Table 5 (Fourth-quintile (80th-percentile) line): Errors (average differences between estimated and true poverty likelihoods for households by score range), with confidence intervals, from 1,000 bootstraps of n = 16,384, 2011/12 scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	0.0	0.0	0.0	0.0
5–9	0.0	0.0	0.0	0.0
10–14	+0.1	1.0	1.1	1.4
15–19	+1.8	1.8	2.1	2.9
20–24	+1.7	1.7	2.0	2.8
25–29	+0.2	1.5	1.8	2.3
30–34	+2.1	1.8	2.2	2.9
35–39	–7.1	4.6	4.7	5.1
40–44	–3.3	2.6	2.9	3.2
45–49	+1.4	2.4	3.0	3.9
50–54	+2.1	2.4	2.7	3.6
55–59	–15.4	9.3	9.5	9.8
60–64	0.0	2.4	2.8	3.5
65–69	–1.4	1.7	2.1	2.7
70–74	+0.4	1.1	1.3	1.7
75–79	–2.5	2.6	3.1	3.9
80–84	+0.1	1.2	1.3	1.7
85–89	+0.1	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

Table 6 (Fourth-quintile (80th-percentile) line): Errors (average differences between estimated poverty rates and true values) for a group at a point in time by sample size, with confidence intervals, for 1,000 bootstraps of various sample sizes, 2011/12 scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	0.0	67.4	82.6	89.9
4	–1.4	39.8	46.4	58.1
8	–1.5	28.2	32.8	42.5
16	–1.3	20.4	24.3	30.3
32	–1.1	14.6	17.4	22.3
64	–1.4	9.7	12.1	15.0
128	–1.3	6.9	8.0	10.7
256	–1.3	5.0	6.0	7.4
512	–1.3	3.6	4.1	5.4
1,024	–1.4	2.6	3.1	3.9
2,048	–1.4	1.8	2.1	2.6
4,096	–1.3	1.3	1.5	2.0
8,192	–1.3	0.9	1.1	1.5
16,384	–1.3	0.6	0.8	1.0

Table 9 (Fourth-quintile (80th-percentile) line): Percentages of households by cut-off score and targeting classification, along with the hit rate and BPAC, 2011/12 scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.0	71.8	0.0	28.2	28.2	–100.0
≤9	0.2	71.6	0.0	28.2	28.4	–99.4
≤14	1.7	70.1	0.0	28.2	29.9	–95.2
≤19	5.6	66.2	0.1	28.1	33.7	–84.3
≤24	12.9	58.9	0.3	27.9	40.8	–63.7
≤29	22.7	49.1	0.6	27.6	50.2	–36.0
≤34	33.7	38.1	1.5	26.7	60.4	–4.0
≤39	44.5	27.3	2.4	25.8	70.3	+27.3
≤44	54.5	17.3	4.5	23.6	78.1	+58.1
≤49	60.8	11.0	7.2	20.9	81.7	+79.5
≤54	65.6	6.2	10.6	17.6	83.2	+85.3
≤59	68.6	3.2	13.5	14.7	83.3	+81.2
≤64	70.1	1.7	16.6	11.5	81.6	+76.8
≤69	71.1	0.7	21.1	7.1	78.1	+70.6
≤74	71.4	0.4	23.7	4.5	75.9	+67.0
≤79	71.7	0.1	25.3	2.8	74.5	+64.7
≤84	71.8	0.0	27.1	1.1	72.8	+62.2
≤89	71.8	0.0	28.1	0.0	71.8	+60.8
≤94	71.8	0.0	28.2	0.0	71.8	+60.8
≤100	71.8	0.0	28.2	0.0	71.8	+60.8

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

Table 10 (Fourth-quintile (80th-percentile) 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), 2011/12 scorecard applied to the 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	0.0	100.0	0.0	Only poor targeted
≤9	0.2	100.0	0.3	Only poor targeted
≤14	1.7	100.0	2.4	Only poor targeted
≤19	5.7	98.8	7.8	85.3:1
≤24	13.2	98.0	18.0	48.3:1
≤29	23.3	97.4	31.6	38.2:1
≤34	35.2	95.8	47.0	23.0:1
≤39	46.9	94.9	62.0	18.8:1
≤44	59.1	92.3	75.9	11.9:1
≤49	68.1	89.3	84.7	8.4:1
≤54	76.2	86.1	91.4	6.2:1
≤59	82.1	83.5	95.6	5.1:1
≤64	86.7	80.8	97.6	4.2:1
≤69	92.2	77.1	99.0	3.4:1
≤74	95.1	75.1	99.5	3.0:1
≤79	97.1	73.9	99.9	2.8:1
≤84	98.9	72.6	100.0	2.6:1
≤89	100.0	71.8	100.0	2.5:1
≤94	100.0	71.8	100.0	2.5:1
≤100	100.0	71.8	100.0	2.5:1