

Simple Poverty Scorecard[®] Poverty-Assessment Tool Palestine (West Bank and Gaza Strip)

Mark Schreiner

11 June 2014

This document is at SimplePovertyScorecard.com.

Abstract

The Simple Poverty Scorecard-brand poverty-assessment tool uses 10 low-cost indicators from the 2011 Palestine (West Bank and Gaza Strip) Expenditure and Consumption Survey to estimate the likelihood that a household has consumption below a given poverty line. Field workers can collect responses in about ten minutes. The scorecard's accuracy is reported for a range of poverty lines. The scorecard is a practical way for pro-poor programs in Palestine (West Bank and Gaza Strip) to measure poverty rates, to track changes in poverty rates over time, and to segment clients for differentiated treatment.

Version note

This paper uses 2011 data and Palestine's new definition of *poverty*. It replaces Schreiner (2010a and 2008c), which use 2007 and 2005 data and an older definition of *poverty*. The new 2011 scorecard here and the new definition should be used from now on. Existing users can still measure change over time using "legacy" poverty definitions with a baseline from the old 2005 or 2007 scorecard and a follow-up from the new 2011 scorecard.

Acknowledgements

This paper is funded by the United Nations Relief and Works Agency for Palestine Refugees in the Near East. Data are from the Palestinian Central Bureau of Statistics. Thanks go to Alex Pollock and Niveen Zein-Eldeen. "Simple Poverty Scorecard" is a Registered Trademark of Microfinance Risk Management, L.L.C. for its brand of poverty-assessment tools. Copyright © 2017 Microfinance Risk Management.

Simple Poverty Scorecard[®] Poverty-Assessment Tool

Interview ID: _____	<u>Name</u>	<u>Identifier</u>
Interview date: _____	Participant: _____	_____
Country: <u>PSE</u>	Field agent: _____	_____
Scorecard: <u>003</u>	Service point: _____	_____
Sampling wgt.: _____	Number of household members: _____	

Indicator	Response	Points	Score
1. How many household members are there?	A. Eight or more	0	
	B. Seven	8	
	C. Six	9	
	D. Five	13	
	E. Four	20	
	F. One, two, or three	26	
2. How many household members are self-employed without employees, are unpaid employees, or work for an irregular wage?	A. Two or more	0	
	B. One	4	
	C. None	9	
3. During the past week, was the male head/spouse employed for one hour or more?	A. No	0	
	B. No male head/spouse	5	
	C. Yes	10	
4. During the past week, was the female head/spouse employed for one hour or more?	A. No	0	
	B. Yes	4	
	C. No female head/spouse	9	
5. How many rooms does the residence have (excluding kitchens and bathrooms)?	A. One	0	
	B. Two	6	
	C. Three	9	
	D. Four	13	
	E. Five or more	16	
6. What is the main source of energy for heating the residence?	A. None, coal/firewood, solar, or other	0	
	B. Gas, kerosene, or electricity	4	
7. Does the household have a solar boiler?	A. No	0	
	B. Yes	5	
8. Does the household have a filter?	A. No	0	
	B. Yes	4	
9. Does the household have a vacuum cleaner?	A. No	0	
	B. Yes	10	
10. Does the household have a computer?	A. No	0	
	B. Yes	7	

**Look-up table for converting scores
to old-definition (legacy) poverty likelihoods**

Score	Poverty likelihood (%)					
	National poverty lines				Intl. 2005 PPP	
	Deep	100%	150%	200%	\$3.75	\$5.00
0–4	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
10–14	92.2	94.0	98.4	98.5	77.5	94.0
15–19	70.6	86.1	96.9	97.2	53.7	86.1
20–24	58.0	76.1	96.2	97.2	50.2	75.0
25–29	42.5	65.7	88.9	95.9	34.5	61.2
30–34	37.7	57.3	82.3	95.2	24.7	53.7
35–39	24.3	45.7	75.9	91.3	18.2	40.8
40–44	15.7	34.2	69.1	83.7	9.2	27.4
45–49	10.7	22.9	61.3	78.1	6.1	20.5
50–54	2.2	9.2	40.0	63.3	0.8	7.1
55–59	2.2	8.3	33.0	57.1	0.7	5.7
60–64	0.5	2.7	19.1	40.9	0.2	2.1
65–69	0.1	1.1	16.7	37.5	0.1	1.1
70–74	0.1	0.9	14.6	27.1	0.1	0.9
75–79	0.0	0.0	8.0	13.4	0.0	0.0
80–84	0.0	0.0	2.4	6.1	0.0	0.0
85–89	0.0	0.0	1.5	6.1	0.0	0.0
90–94	0.0	0.0	0.0	5.1	0.0	0.0
95–100	0.0	0.0	0.0	0.0	0.0	0.0

There is no legacy median poverty line.

Note on measuring changes in poverty rates over time using old-definition (legacy) lines with old 2005 or 2007 and new 2011 scorecards

This paper uses data from the 2011 PECS and Palestine's new definition of *poverty*. It replaces Schreiner (2010a and 2008c), which uses data from the 2005 and 2007 PECS and an older definition of *poverty*. The new 2011 scorecard here should be used from now on.

Some organizations in Palestine already use the old 2007 or 2005 scorecards. Even after switching to the new 2011 scorecard here, these legacy users can still measure changes in poverty rates over time with existing baseline estimates from the old 2007 or 2005 scorecards and follow-up estimates from the new 2011 scorecard. This is possible because the new 2011 scorecard is calibrated not only to the new definition of *poverty* but also to the older definition. Hybrid estimates of change that splice together estimates from the two scorecards are valid as long as they use an old-definition (legacy) poverty line. The appendix of this paper is a step-by-step guide to calculating hybrid estimates of change. The appendix also gives instructions on how to find a “grand” estimate of change that adds a hybrid estimate of change with an estimate of change based solely on new-definition poverty lines.

To sum up, both first-time and legacy users should use the new 2011 scorecard and the new definition of *poverty*. Looking forward, this establishes a baseline with the best, most policy-relevant poverty lines. Looking backward, legacy users can still salvage existing estimates when measuring change in poverty rates over time.

Simple Poverty Scorecard[®] Poverty-Assessment Tool Palestine (West Bank and Gaza Strip)

1. Introduction

This paper presents the Simple Poverty Scorecard brand poverty-assessment tool. Pro-poor programs in Palestine¹ (West Bank and Gaza Strip) can use it to estimate the likelihood that a household has consumption below a given poverty line, to estimate a population's poverty rate at a point in time, to track changes in a population's poverty rate over time, and to segment participants for differentiated treatment.

The direct approach to poverty measurement via surveys is difficult and costly. As a case in point, the 2011 Palestine Expenditure and Consumption Survey (PECS) runs 56 pages. Female enumerators visited households eight to ten times over a period of four weeks, helping the household complete a consumption diary (Palstinian Central Bureau of Statistics, 2012a) and applying a consumption module with almost 800 items, asking questions such as: “In the past week, how much long-grain rice did the household eat? How much was this rice worth? Now then, in the past week, how much short-grain rice did the household eat? . . .”

In comparison, the indirect approach via the scorecard is simple, quick, and inexpensive. It uses ten verifiable indicators (such as “How many rooms does the

¹ *Palestine* is used here as shorthand for the “occupied Palestinian territories” or the “West Bank and Gaza Strip”.

residence have (excluding kitchens and bathrooms)?” and “Does the household have a solar boiler?”) to get a score that is highly correlated with poverty status as measured by the exhaustive PECS 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). Estimates 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 live in households whose per-adult-equivalent or per-capita consumption is below a given poverty line, for example, the Millennium Development Goals’ \$1.25/day per-person line at 2005 purchase-power parity (PPP). USAID microenterprise partners in Palestine can use scoring with the median poverty line to report how many of their participants are “very poor”.³ Scoring can also be used to measure net movement across

² The Simple Poverty Scorecard tool for Palestine 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 expenditure is below the highest of the new-definition \$1.25/day 2005 PPP line (NIS3.68 in Palestine in 2011, Figure 1) or the new-definition median line that divides people in households below Palestine’s new-definition national poverty line into two equal-size groups (NIS11.87).

a poverty line over time. In all these cases, the scorecard provides a consumption-based, objective tool with known accuracy. While consumption surveys are costly even for governments, some local pro-poor organizations may be able to implement an inexpensive poverty-assessment tool to help with poverty monitoring and (if desired) targeting.

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

Beyond its simplicity and transparency, the scorecard’s technical approach is innovative in how it associates scores with poverty likelihoods, in the extent of its accuracy tests, and in how it derives formulas for standard errors. Although the

accuracy tests are simple and commonplace in statistical practice and in the for-profit field of credit-risk scoring, they have rarely been applied to poverty-assessment tools.

The scorecard is constructed with data from the 2011 PECS from the Palestinian Central Bureau of Statistics (PCBS). 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 both the West Bank and the Gaza Strip

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 group of households at a point in time. This is the average poverty likelihood of households in the group.

Third, the scorecard can estimate changes in the poverty rate for a group of households (or for two independent samples of households that are both representative of the same population) between two points in time. This is the baseline/follow-up change in the average poverty likelihood of the group(s).

The scorecard can also be used to target services to different segments of participants. To help managers choose an appropriate targeting cut-off for their

purposes, this paper reports several measures of targeting accuracy for a range of possible cut-offs.

This paper presents a single scorecard whose indicators and points are derived from household consumption data from the 2011 PECS and Palestine’s new-definition national poverty line. Scores from this one scorecard are calibrated—again using data from the 2011 PECS—to poverty likelihoods for 10 new-definition poverty lines and six legacy lines based on an old definition of *poverty*.⁴

The scorecard is calibrated to both the new and old definition of poverty status with data from the 2011 PECS under the assumption that this paper has appropriately updated the old-definition poverty lines to account for changes in prices from 2007 to 2011 and from 2005 to 2011.

The scorecard is constructed and calibrated using half of the data from the 2011 PECS. The other half is used to validate the scorecard’s accuracy for estimating households’ poverty likelihoods, for estimating groups’ poverty rates at a point in time, and for targeting.

All three of scoring’s estimates are *unbiased*. That is, they match the true value on average in repeated samples when constructed from (and applied to) a single, unchanging population at a point in time. Like all predictive models, the specific

⁴ Section 2 below discusses the new and old definitions of *poverty*.

scorecard here is constructed from a single sample and so misses the mark to some unknown extent when applied to a different population or when applied after 2011.⁵

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

When applied to the validation sample with 1,000 bootstraps of $n = 16,384$, the average average difference between scorecard estimates of groups' poverty rates and the true rates at a point in time with the new-definition national line is +1.3 percentage points. The average absolute average difference across all 10 new-definition poverty lines is about 0.9 percentage points, and the maximum absolute average difference is 3.1 percentage points. These differences are due to sampling variation rather than bias; the average difference for a given poverty line would be zero if the whole 2011 PECS were to be repeatedly redrawn and divided into sub-samples before repeating the entire process of constructing and validating scorecards.

⁵ Important examples include nationally representative samples at a different point in time or sub-groups that are not nationally representative (Tarozzi and Deaton, 2009).

The 90-percent confidence intervals for these estimates are ± 0.6 percentage points or less. For $n = 1,024$, the 90-percent intervals are ± 2.5 percentage points or less.

When the old-definition lines are applied to the validation sample with 1,000 bootstraps of $n = 16,384$, the average absolute average difference across the six lines is about 1.3 percentage points, and the maximum absolute average difference is 2.9 percentage points. The 90-percent confidence intervals are ± 0.6 percentage points or less. For $n = 1,024$, the 90-percent intervals are ± 1.9 percentage points or less.

Section 2 below describes data and definitions of *poverty*. Sections 3 and 4 describe scorecard construction and offer guidelines for use in practice. Sections 5 and 6 detail the estimation of households' poverty likelihoods and of groups' poverty rates at a point in time. Section 7 discusses estimating changes in poverty rates over time, and Section 8 covers targeting. Section 9 places the scorecard here in the context of a similar exercise for Palestine. The last section is a summary.

2. Data and poverty lines

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

2.1 Data

The scorecard is based on data from the 4,317 households in the 2011 PECS fielded by the PCBS from January 15, 2011 to January 14, 2012. This is Palestine's most recent, available national consumption survey.

For the purposes of the scorecard, the households in the 2011 PECS are randomly divided into two sub-samples:

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

2.2 Poverty rates

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

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 equal-weighted average of poverty statuses (or estimated poverty likelihoods) across households with participants. In the example here, this is $\frac{1 \cdot 1 + 1 \cdot 0}{1 + 1} = \frac{1}{2} = 0.5 = 50$ percent. In the “1 · 1” term in the numerator, the first “1” is the first household’s weight, and the second “1” is the first household’s poverty status (poor). In the “1 · 0” term in the numerator, the “1” is the second household’s weight, and the “0” is the second household’s poverty status (non-poor). The “1 + 1” in the denominator is the sum of the weights of the two households. Each household has a weight of one (1) 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 average of poverty statuses for households with participants, or $\frac{3 \cdot 1 + 4 \cdot 0}{3 + 4} = \frac{3}{7} = 0.43 = 43$ percent. In the “3 · 1” term in the numerator, the “3” is the first household’s weight because it has three members, and the “1” is its poverty status (poor). In the “4 · 0” term in the numerator, the “4” is

the second household’s weight because it has four members, and the zero is its poverty status (non-poor). The “3 + 4” in the denominator is the sum of the weights of the two households. A household’s weight is its number of members because the unit of analysis is the household member.

As a final example—one that pertains to what is likely the most common situation in practice—a program counts as *participants* only those household members with whom it deals with directly. For the example here, this means that some—but not all—household members are counted. The person-level rate is now the participant-weighted average of the poverty statuses of households with participants, or

$$\frac{1 \cdot 1 + 2 \cdot 0}{1 + 2} = \frac{1}{3} = 0.33 = 33 \text{ percent.}$$

The first “1” in the “1 · 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 · 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 the weights are the number of relevant units in the household. When reporting, organizations should explain who is counted as a *participant* and why.

Figure 1 reports poverty lines and poverty rates (for both new and old definitions of *poverty*) for both households and people for all of Palestine and for the West Bank and the Gaza Strip in 2005, 2007, and 2011. Person-level poverty rates are included in Figure 1 because these are the rates reported by governments and used in most policy discussions. Household-level poverty rates are also reported because—as discussed above—household-level poverty likelihoods can be straightforwardly converted into poverty rates for other units of analysis. This is also the reason why the scorecard is constructed, calibrated, and validated with household weights.

2.3 Poverty status, consumption, poverty lines, and poverty rates

Poverty status is whether a household is poor or non-poor. In Palestine, poverty status is determined by whether per-adult-equivalent or per-capita aggregate household consumption is less than a poverty line. Thus, *poverty status* has two aspects: a measure of household consumption, and a poverty line.

Consumption is measured in the same way in the 2005, 2007, and 2011 PECS. The components of consumption are discussed in Schreiner (2010a).

In terms of poverty lines, Palestine has a “deep” line (the observed consumption on food, clothing, and housing for households at the 30th percentile of total consumption in 1995, updated over time for inflation, see World Bank, 2004) and a “national” line

(the deep line plus the observed consumption on health care, education, transportation, personal care, and residential utilities for households at the 30th percentile in 1995).

The old definition of *poverty* used for the previous two scorecards for Palestine in Schreiner (2010a) with data from the 2007 PECS and in Schreiner (2008c) with data from the 2005 PECS applied a single value for the deep line and a single value for the national line to all households in Palestine. This old definition of *poverty* used by the two previous scorecards differs from Palestine’s official old definition of *poverty* because the scorecard authors were unaware of the correct official definition, in part because the official definition is not fully and clearly documented.

The new definition of *poverty* used in this paper with data from the 2011 PECS matches Palestine’s official new definition of *poverty*. The official new definition differs from the official old definition in that it adjusts both the deep and national lines for cost-of-living differences between the West Bank and the Gaza Strip.⁶

For the deep line and the national line under the official definitions (both old and new), a household is *poor* if its consumption per adult-equivalent is below the given line. In all years, the number of adult equivalents in a household is defined as $(\text{Adults} + 0.46 \cdot \text{Children})^{0.89}$, where a *child* is as anyone 17-years-old or younger.⁷

⁶ This seems to be the only documented difference between the old and new official definitions of *poverty*.

⁷ This formula (and the fact that lines are per-adult-equivalent) was unknown to Schreiner (2010b) and Schreiner (2008c). The sole documentation of this aspect of the lines appears to be World Bank (2004).

The old definition of *poverty* used in the two previous scorecards had a deep poverty line of NIS7.18 per person (not per adult equivalent) per day in 2005 and NIS10.33 in 2007 (Figure 1). The old-definition national line used in the two previous scorecards was NIS10.51 per person per day in 2005 and NIS12.92 in 2007.

For 2011, the old-definition poverty lines are updated from their 2007 values by multiplying them by the change in the all-Palestine consumer price index (CPI).⁸ The factor is $102.88 \div 85.36 = 1.2052$, so the old-definition deep line used with the new 2011 scorecard is $\text{NIS}10.33 \times 1.2052 = \text{NIS}12.46$ per person (not per adult equivalent) per day, and the old-definition national line used with the new 2011 scorecard is $\text{NIS}12.92 \times 1.2052 = \text{NIS}15.58$.⁹

In 2011, the published, official, new-definition poverty rate for the deep line for Palestine as a whole is 12.9 percent of people (PCBS, 2012b). The poverty rate

⁸ pcbs.gov.ps/site/lang__en/507/site/695/default.aspx, retrieved 2 June 2014.

⁹ Again, the old-definition poverty lines used by Schreiner (2010a) with the 2007 PECS and by Schreiner (2008c) with the 2005 PECS differ from Palestine's official old-definition lines because the correct official definitions were not known to those authors. Thus, the old-definition lines used in previous scorecards are incorrectly in per-capita terms and are incorrectly based on the single poverty line for PCBS' representative household. Consequently, the resulting old-definition poverty rates calculated from the 2005 and 2007 data do not match the official, published old-definition poverty rates reported by the PCBS for that same data. The scorecard estimates based on the mistaken old-definition lines are not completely useless; rather, they relate to a different definition of *poverty* than do the official estimates. For compatibility with estimates based on the old 2007 and 2005 scorecards, the same mistakes are repeated in the old-definition (legacy) poverty lines used with the new 2011 scorecard. The mistakes are corrected in the new-definition lines to be used from now on.

calculated here from the data is 12.4 percent for people and 9.5 percent for households (Figure 1).¹⁰

For the new-definition national line (sometimes called here “100% of the new-definition national line”), the published poverty rate for Palestine as a whole in 2011 is 25.8 percent of people (PCBS, 2012b). With the data received for this project, the poverty rate for this same line is 25.4 percent of people and 20.3 percent of households (Figure 1).

¹⁰ The cause of the discrepancy between the official (12.9 percent) and calculated (12.4 percent) person-level poverty rates is unknown. In communication with Niveen Zein-Eldeen, the PCBS said that the algorithm in this paper is correct but that the published figures differ because they include not only the West Bank and the Gaza Strip but also Jerusalem Area J1. But this explanation does not fit other evidence. PCBS (2012b) does not report a poverty rate for Jerusalem Area J1 in 2011, and footnote 34 in PCBS (2012a) says that “Jerusalem Area J1 is not included in the data”. The data received for this project from PCBS has no households from Jerusalem Area J1. Furthermore, the weighted average of the published person-level rates for the new-definition national line in the West Bank (17.8 percent) and Gaza Strip (38.8 percent) is, when using the person-level weights in the data (0.6182 for the West Bank and 0.3818 for the Gaza Strip), the same as the published all-Palestine rate of 25.8 percent. Finally, poverty rates in Jerusalem Area J1 are low (in the range of 1–5 percent for the national line, PCBS, 2010), so including them in the all-Palestine average calculated here would only increase the discrepancy with the published rate. Thus, the discrepancy probably stems from undocumented differences in the consumption data used to derive the official estimate versus in the data provided to this project.

Using data from the 2011 PECS, the new 2011 scorecard is constructed with the new-definition national line and calibrated both to new-definition lines and to old-definition lines. This permits existing users of the old 2007 scorecard in Schreiner (2010a) to switch to the new 2011 scorecard here without having to start over from scratch when measuring change over time, and it also allows the use of the correct, higher-quality new-definition lines from now on.

Because local, pro-poor programs in Palestine may want to use different or various poverty lines, this paper calibrates scores from its single scorecard to poverty likelihoods for ten new-definition lines and six old-definition (legacy) lines:

- New-definition deep
- 100% of new-definition national
- 150% of new-definition national
- 200% of new-definition national
- New-definition median
- New-definition \$1.25/day 2005 PPP
- New-definition \$2.00/day 2005 PPP
- New-definition \$2.50/day 2005 PPP
- New-definition \$3.75/day 2005 PPP
- New-definition \$5.00/day 2005 PPP
- Old-definition deep
- 100% of the old-definition national
- 150% of the old-definition national
- 200% of the old-definition national
- Old-definition \$3.75/day 2005 PPP
- Old-definition \$5.00/day 2005 PPP

The new-definition median line is defined as the median per-capita consumption of people (not households) in a given poverty-line region who are below that region's new-definition national line (Schreiner, 2014; United States Congress, 2004).

The new-definition and old-definition 2005 PPP lines are derived from:

- 2005 PPP exchange rate of NIS2.310 per \$1.00 (Sun and Swanson, 2009)
- Consumer Price Index for Palestine of:¹¹
 - Average in 2005: 80.70
 - Average in 2007: 85.36
 - Average in 2011: 102.88
- Average new-definition national line in 2011 for all of Palestine (Figure 1): NIS24.84
- New-definition national line in 2011 for the West Bank (NIS26.01) and for the Gaza Strip (NIS22.95) (Figure 1)
- Old-definition \$3.75/day 2005 PPP line in 2007: NIS9.23 (Figure 1)

The old-definition \$3.75/day line for 2011 is the old-definition \$3.75/day line for 2007, increased by the change in the all-Palestine CPI between 2007 and 2011, which is $\text{NIS}9.23 \times (102.88 \div 85.36) = \text{NIS}11.12$ (Figure 1). This line applies to all of Palestine. The old-definition \$5.00/day line for 2011 is the old-definition \$3.75/day line for 2011, multiplied by $(5 \div 3.75)$.

Using the formula in Sillers (2006), the all-Palestine new-definition \$1.25/day 2005 PPP line for 2011 is:

$$\begin{aligned} & (\text{2005 PPP exchange rate}) \cdot \$1.25 \cdot \left(\frac{\text{CPI}_{2011}}{\text{CPI}_{2005}} \right) = \\ & \left(\frac{\text{NIS}2.310}{\$1.00} \right) \cdot \$1.25 \cdot \left(\frac{102.88}{80.70} \right) = \text{NIS}3.68. \end{aligned}$$

¹¹ pcbs.gov.ps/site/lang__en/507/site/695/default.aspx, retrieved 2 June 2014.

These line applies to Palestine as a whole.¹² This average new-definition \$1.25/day line is adjusted for cost-of-living differences across the West Bank and the Gaza Strip by multiplying it by a given poverty-line region’s new-definition national line and then dividing it by the average all-Palestine new-definition national line. For the example of the West Bank, this is the all-Palestine new-definition \$1.25/day line (NIS3.68) multiplied by new-definition national line for the West Bank (NIS26.01), divided by the average all-Palestine new-definition national line (NIS24.84), or $3.68 \times (26.01 \div 24.84) = \text{NIS}3.85$ (Figure 1).

USAID microenterprise partners in Palestine who use the scorecard to report the share of their participants who are “very poor” to USAID should use the new-definition median line. This is because USAID defines the “very poor” as those people in households whose consumption is below the highest of two lines:

- New-definition \$1.25/day 2005 PPP (NIS3.68, Figure 1)
- New-definition median line (NIS11.87).

¹² The latest person-level poverty rates for the \$1.25/day line from the World Bank’s PovcalNet (iresearch.worldbank.org/PovcalNet, retrieved 1 June 2014) is 0.04 percent for the 2009 PECS. This is close to the 0.2 percent in Figure 1 for the 2011 PECS. Schreiner (2014) argues that the estimate here is to be preferred because PovcalNet does not document its poverty line, its definition of *consumption*, its price adjustment between 2005 and 2009, nor its regional price adjustments.

2.4 Conditions required to measure change over time when combining estimates from two scorecards

2.4.1 Hybrid estimates of change

Hybrid estimates of change over time in a poverty rate for an old-definition (legacy) line are found by splicing together a baseline estimate from the old 2007 scorecard with a follow-up estimate from the new 2011 scorecard. For this to be valid, the old-definition line must be properly adjusted from 2007 prices to 2011 prices. In this paper, this is done in step with the change in the all-Palestine CPI from 2007 to 2011 (85.36 to 102.88). This seems reasonable, and it mirrors how the PCBS (2010) retrospectively adjusts new-definition lines across time.

While prices increased by 20.5 percent from 2007 to 2011, average per-capita consumption increased by about 32 percent, and average per-adult-equivalent consumption increased by about 28 percent. With consumption outpacing inflation, person-level poverty rates in the data for the old-definition national line decreased from 37.4 percent in 2007 (Figure 2 in Schreiner, 2010a) to 29.5 percent in 2011 (Figure 1).

2.4.2 “Grand” estimates of change

Hybrid estimates of change matter now as existing users of the old 2007 scorecard switch to the new 2011 scorecard. In the future, these users may also want to find “grand” estimates of change which add together a hybrid estimate of change based on old-definition lines with a non-hybrid estimate of change based only on new-

definition lines.¹³ Such “grand” estimates of change are valid if the rates of change under both the old and new definitions of *poverty* are the same, even if the level of poverty rates differ under the two definitions.

This “parallel lines” assumption can be checked; PCBS (2010, p. 1) reports estimated poverty rates from the 2004, 2005, 2006, 2007, and 2009 PECS for both the official old-definition and new-definition national lines. For the official old-definition national line, the all-Palestine person-level poverty rate increases between each survey round, from 30.6 percent in 2004, to 33.7 in 2005, to 35.8 percent in 2006. It then jumps up 8.5 percentage points from 2006 to 2007, and finally increases by another 1.4 percentage points from 2007 to 2009.

Except for the big jump 2006/7 jump, the new-definition national line does not follow the same pattern. It stays at 24–25 percent from 2004 to 2006, jumps up about 7 percentage points in 2006/7, and then falls almost 9 percentage points to 21.9 percent in 2007 to 2009. This strongly violates the “parallel lines” assumption; poverty increases a lot from 2004 to 2009 under the old definition, while poverty decreases a little under the new definition.

As discussed above, the old-definition national line used by the scorecard differs from the official old-definition national line in that the scorecard’s definition is in per-capita terms rather than per-adult-equivalent terms. Nevertheless, it seems unlikely that the “parallel lines” assumption would hold.

¹³ See the appendix for step-by-step instructions.

3. Scorecard construction

For Palestine, about 75 candidate indicators are initially prepared in the areas of:

- Household composition (such as the number of members)
- Education (such as school attendance)
- Housing (such as the number of roomss)
- Ownership of durable assets (such as solar boilers or vacuum cleaners)
- Employment (such as whether the male head/spouse is employed)

Figure 2 lists the candidate indicators, ordered by the entropy-based “uncertainty coefficient” (Goodman and Kruskal, 1979) that measures how well a given indicator predicts poverty on its own.¹⁴

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

The scorecard itself is built using 100% of the new-definition 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).

¹⁴ The uncertainty coefficient is not used as a criterion when selecting scorecard indicators; it is just a way to order the candidate indicators in Figure 2.

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

A series of two-indicator scorecards are then built, each based on the one-indicator scorecard selected in the first round, with a second candidate indicator added. The best two-indicator scorecard is then selected, again based on “c” and judgment about how to best balance the non-statistical criteria. These steps are repeated until the scorecard has 10 indicators that work well together.

The final step is to transform the Logit coefficients into non-negative integers such that total scores range from 0 (most likely below a poverty line) to 100 (least likely below a poverty line).

This algorithm is similar to the common R^2 -based stepwise least-squares regression. It differs from naïve stepwise in that the selection of indicators considers both statistical¹⁵ and non-statistical criteria. The non-statistical criteria can improve robustness through time and help ensure that indicators are simple, sensible, and acceptable to users.

¹⁵ The statistical criterion for selecting an indicator is not the p value of its coefficient but rather its contribution to the ranking of households by poverty status.

The single scorecard here applies to all of Palestine. 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 poverty-assessment tools by urban/rural does not improve targeting accuracy much. In general, segmentation may improve the accuracy of estimates of poverty rates (Tarozzi and Deaton, 2009), but segmentation may also increase the risk of overfitting (Haslett, 2012).

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, Palestine’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 weights (non-negative integers, and no arithmetic beyond addition)

The scorecard (and its back-page worksheet) is ready to be photocopied. A field worker using Palestine’s paper scorecard would:

- Record the names and identifiers of the participant, of the field worker, and of the relevant organizational service point
- Record the date that the participant first participated with the organization
- Record the date of the scorecard interview
- Complete the back-page worksheet with each household member’s:
 - First name
 - Age
 - Work status
 - Work type
- Record household size in the scorecard header, and record the responses to first four scorecard indicators based on the back-page worksheet
- Read each of the remaining six questions one-by-one from the scorecard, drawing a circle around the relevant responses and their points, and writing each point value in the far right-hand column
- Add up the points to get a total score
- Implement targeting policy (if any)
- Deliver the paper scorecard to a central office for data entry and filing

Of course, field workers must be trained. The quality of outputs depends on the quality of inputs. If organizations or field workers gather their own data and believe that they have an incentive to exaggerate poverty rates (for example, if 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).¹⁶ IRIS Center (2007a) and Toohig (2008) are useful nuts-and-bolts guides for budgeting, training field workers and supervisors, logistics, sampling, interviewing, piloting, recording data, and controlling quality.

In particular, while collecting scorecard indicators is relatively easier than alternative ways of measuring poverty, it is still absolutely difficult. Training and explicit definitions of terms and concepts in the scorecard are essential, and field workers should scrupulously study and follow the “Guidelines for the Interpretation of Scorecard Indicators” found at the end of this paper, as they are an integral part of the Simple Poverty Scorecard tool.¹⁷

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

¹⁶ If a program does not want field workers to know the points associated with responses, then it can erase the points from the paper scorecard and then apply the points later at a central office. 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. Even if points are hidden, field workers and respondents can apply common sense to guess how response options are linked with poverty.

¹⁷ 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 the PCBS did in the 2011 PECS.

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

In terms of implementation planning, an organization must make choices about:

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

In general, the sampling design should follow from the organization’s goals for the exercise, the questions to be answered, and the budget. The main goal should be to make sure that the sample is representative of a well-defined population and that the scorecard will inform an issue that matters to the organization.

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

- Employees of the organization
- Third parties

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

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

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

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

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

The frequency of application can be:

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

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

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

An example set of choices is illustrated by BRAC and ASA, two microfinance organizations in Bangladesh who each have about 7 million participants and who apply the Simple Poverty Scorecard tool for Bangladesh (Schreiner, 2013a) with a sample of about 25,000. Their design is that all loan officers in a random sample of branches score all participants each time they visit a homestead (about once a year) as part of their standard due diligence prior to loan disbursement. They record responses on paper in the field before sending the forms to a central office to be entered into a database and converted to poverty likelihoods.

5. Estimates of household poverty likelihoods

The sum of scorecard points for a household is called the *score*. For Palestine, 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, cutting the score in half increases the estimated likelihood of being poor, but does not double it.

To get absolute units, scores must be converted to *poverty likelihoods*, that is, probabilities of being below a poverty line. This is done via simple look-up tables. For the example of 100% of the new-definition national line, scores of 30–34 have a poverty likelihood of 48.1 percent, and scores of 35–39 have a poverty likelihood of 40.1 percent (Figure 3).

The poverty likelihood associated with a score varies by poverty line. For example, scores of 30–34 are associated with a poverty likelihood of 48.1 percent for 100% of the new-definition national line but of 50.4 percent for the new-definition \$5.00/day 2005 PPP line.¹⁸

¹⁸ Starting with Figure 3, many figures have 16 versions, one for each of the six old-definition poverty lines and the ten new-definition lines. To keep them straight, they are grouped by poverty line. Single tables pertaining to all lines are placed with the tables for the new-definition deep 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 are below a given poverty line.

For the example of 100% of the new-definition national line (Figure 4), there are 6,426 (normalized) households in the calibration sub-sample with a score of 30–34. Of these, 3,088 (normalized) are below the poverty line. The estimated poverty likelihood associated with a score of 30–34 is then 48.1 percent, because $3,088 \div 6,426 = 0.481$.

To illustrate with 100% of the new-definition national line and a score of 35–39, there are 8,917 (normalized) households in the calibration sample, of whom 3,577 (normalized) are below the line (Figure 4). The poverty likelihood for this score is then $3,577 \div 8,917 = 40.1$ percent.

The same method is used to calibrate scores with estimated poverty likelihoods for the other 16 poverty lines.¹⁹

Figure 5 shows, for all scores—and separately for new-definition/old-definition lines and as well as for per-adult-equivalent/per-capita lines—the likelihood that a given household’s consumption falls in a range demarcated by two adjacent poverty lines.

¹⁹ To ensure that poverty likelihoods always decrease as scores increase, it is sometimes necessary to average likelihoods iteratively across series of adjacent scores before grouping scores into ranges. This preserves unbiasedness, and it keeps users from balking when sampling variation in score ranges with few households leads to higher scores being linked with higher poverty likelihoods.

For the example of new-definition, per-adult-equivalent lines, the daily per-adult-equivalent consumption of a household with a score of 30–34 falls in the following ranges with probability:

- 27.8 percent below the deep line
- 20.2 percent between the deep line and 100% of the national line
- 34.2 percent between 100% and 150% of the national line
- 14.2 percent between 150% and 200% of the national line
- 3.5 percent above 200% of the national line

For new-definition, per-capita lines, the distribution of per-capita consumption for a household with a score of 30–34 is:

- 0.0 percent below \$1.25/day
- 1.1 percent between \$1.25/day and \$2.00/day
- 2.6 percent between \$2.00/day and \$2.50/day
- 21.2 percent between \$2.50/day and \$3.75/day
- 2.5 percent between \$3.75/day and the median line
- 23.1 percent between the median line and \$5.00/day
- 49.6 percent above \$5.00/day

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 for the scorecard 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.*, 2004). Of course, the scorecard here is constructed with both data and judgment. The fact that this paper acknowledges that some choices in scorecard construction—as in

most statistical analysis—are informed by judgment in no way impugns the objectivity of the poverty likelihoods, as this depends on using data in score calibration, not on using data (and nothing else) in scorecard construction.

Although the points in the Palestine 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 that 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. 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.²⁰

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

²⁰ This follows because these estimates of groups' poverty rates are linear functions of the unbiased estimates of households' poverty likelihoods.

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

- Score each household in the validation sample
- Draw a new bootstrap sample *with replacement* from the validation sample
- For each score, compute the true poverty likelihood in the bootstrap sample, that is, the share of households with the score and with consumption below a poverty line
- For each score, record the difference between the estimated poverty likelihood (Figure 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 $n = 16,384$, Figure 6 shows the average difference between estimated and true poverty likelihoods as well as confidence intervals for the differences.

For the 100% of the new-definition national line, the average poverty likelihood across bootstrap samples for scores of 30–34 in the validation sample is too low by 5.2 percentage points. For scores of 34–39, the estimate is too high by 4.3 percentage points.²¹

For 100% of the new-definition national line, the 90-percent confidence interval for the differences for scores of 30–34 is ± 4.0 percentage points (Figure 6). This means that in 900 of 1,000 bootstraps, the difference between the estimate and the true value is between -9.2 and -1.2 percentage points (because $-5.2 - 4.0 = -9.2$, and $-5.2 + 4.0 = -1.2$). In 950 of 1,000 bootstraps (95 percent), the difference is -5.2 ± 4.2 percentage points, and in 990 of 1,000 bootstraps (99 percent), the difference is -5.2 ± 4.5 percentage points.

For some scores, Figure 6 shows large differences between estimated poverty likelihoods and true values. This is because the validation sample is a single sample that—thanks to sampling variation—differs in distribution from the construction/calibration sub-samples and from Palestine’s population. For targeting, however, what matters is less the difference in all score ranges and more the difference in score ranges just above and below the targeting cut-off. This mitigates the effects of

²¹ These differences are not zero, despite the estimator’s unbiasedness, because the scorecard comes from a single sample. The average difference by score range 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.

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.

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 PECS fieldwork in January 2012. That is, it may fit the data from the 2011 PECS so closely that it captures not only some timeless patterns but also some random patterns that, due to sampling variation, show up only in the 2011 PECS but not in the overall population of Palestine. 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 it is applied to non-nationally representative samples.

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

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

6. Estimates of a group's poverty rate at a point in time

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

To illustrate, suppose an organization samples three households on 1 January 2014 and that they have scores of 20, 30, and 40, corresponding to poverty likelihoods of 74.1, 48.1, and 32.0 percent (100% of the new-definition national line, Figure 3). The group's estimated poverty rate is the households' average poverty likelihood of $(74.1 + 48.1 + 32.0) \div 3 = 51.4$ percent.

Be careful; the group's poverty rate is *not* the poverty likelihood associated with the average score. Here, the average score is 30, which corresponds to a poverty likelihood of 48.1 percent. This differs from the 51.4 percent found as the average of the three individual poverty likelihoods associated with each of the three scores. Unlike poverty likelihoods, scores are ordinal symbols, like letters in the alphabet or colors in the spectrum. Because scores are not cardinal numbers, they cannot be added up or averaged across households. Only three operations are valid for scores: conversion to poverty likelihoods, distributional analysis (Schreiner, 2012a), or comparison—if desired—with a cut-off for targeting. The safest rule to follow is: Always use poverty likelihoods, never scores.

6.1 Accuracy of estimated poverty rates at a point in time

For the Palestine scorecard applied to 1,000 bootstraps of $n = 16,384$ from the validation sample, the average difference between the estimated poverty rate at a point in time and the true rate for 100% of the new-definition national line is +1.3 percentage points (Figure 8, summarizing Figure 7 across poverty lines). Across the ten new-definition lines, the maximum absolute average difference is 3.1 percentage points, and the average average absolute difference is about 0.9 percentage points.²² At least part of these differences is due to sampling variation in the division of the 2011 PECS into two sub-samples.

When estimating poverty rates at a point in time, the bias reported in Figure 8 should be subtracted from the average poverty likelihood to make the estimate unbiased. For the Palestine scorecard and the example of the 100% of the new-definition national line, bias is +1.3 percentage points, so the unbiased estimate in the three-household example above is $51.4 - (+1.3) = 50.1$ percent.

In terms of precision, the 90-percent confidence interval for a group's estimated poverty rate at a point in time with $n = 16,384$ is ± 0.6 percentage points or less (Figure 8). This means that in 900 of 1,000 bootstraps of this size, the average estimate (after subtracting off bias) is within 0.6 percentage points of the true value.

²² Across the six old-definition lines, the maximum absolute average difference is 2.9 percentage points, and the average absolute average difference is about 1.3 percentage points (Figure 8).

For example, suppose that the average poverty likelihood in a sample of $n = 16,384$ with the Palestine scorecard and 100% of the new-definition national line is 51.4 percent. Then estimates in 90 percent of such samples would be expected to fall in the range of $51.4 - (+1.3) - 0.5 = 49.6$ percent to $51.4 - (+1.3) + 0.5 = 50.6$ percent, with the most likely true value being the unbiased estimate in the middle of this range ($51.4 - (+1.3) = 50.1$ percent). This is because the original (biased) estimate is 51.4 percent, bias is +1.3 percentage points, and the 90-percent confidence interval for 100% of the new-definition national line is ± 0.5 percentage points (Figure 8).

6.2 Formula for standard errors for estimates of poverty rates

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

To derive a formula for the standard errors of estimated poverty rates at a point in time from indirect measurement via poverty-assessment tools (Schreiner, 2008a), first note that the textbook formula (Cochran, 1977) that relates confidence intervals with standard errors in the case of direct measurement of ratios is $\pm c = \pm z \cdot \sigma$, 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 of $\sqrt{\frac{N - n}{N - 1}}$,

N is the population size, and

n is the sample size.

For example, the direct-measure estimate of Palestine's household-level poverty rate for 100% of the new-definition national line in the 2011 PECS is $\hat{p} = 20.3$ percent (Figure 1). If this estimate came from a sample of $n = 16,384$ households from a population N of 697,827 (an estimate of the number of households in Palestine in

2011),²³ then the finite population correction ϕ is $\sqrt{\frac{697,827 - 16,384}{697,827 - 1}} = 0.9882$, which

can be taken as $\phi = 1$. If the desired confidence level is 90-percent ($z = 1.64$), then the

confidence interval $\pm c$ is $\pm z \cdot \sqrt{\frac{\hat{p} \cdot (1 - \hat{p})}{n}} \cdot \sqrt{\frac{N - n}{N - 1}} = \pm 1.64 \cdot \sqrt{\frac{0.203 \cdot (1 - 0.203)}{16,384}} \cdot 1 =$

± 0.515 percentage points.

The scorecard, however, does not measure poverty directly, so this formula is not applicable. To derive a formula for the Palestine scorecard, consider Figure 7, which reports empirical confidence intervals $\pm c$ for the differences for the scorecard applied to 1,000 bootstrap samples of various sizes from the validation sample. For example, with $n = 16,384$ and 100% of the new-definition national line, the 90-percent confidence interval is ± 0.479 percentage points.²⁴

Thus, the 90-percent confidence interval with $n = 16,384$ is ± 0.479 percentage points for indirect estimates via the Palestine scorecard and ± 0.515 percentage points for direct measurement via the 2011 PECS. The ratio of the two intervals is $0.479 \div 0.515 = 0.93$.

²³ The average household in the 2011 PECS has 6.01 members. In the 2007 census, Palestine's population was 3,719,189 people, and in 2014, its estimated population was 4,550,000 (en.wikipedia.org/wiki/Palestinian_territories, retrieved 3 June 2014). Interpolating gives an estimated 2011 population of $3,719,189 + (2011 - 2007) \div (2014 - 2007) \times (4,550,000 - 3,719,189) = 4,193,938$ people. Dividing by the average household size gives $4,193,938 \div 6.01 = 697,827$ households in 2011.

²⁴ Due to rounding, Figure 7 displays 0.5, not 0.479.

Now consider the same case, but with $n = 8,192$. The confidence interval under direct measurement is $\pm 1.64 \cdot \sqrt{\frac{0.203 \cdot (1 - 0.203)}{8,192}} \cdot 1 = \pm 0.729$ percentage points. The empirical confidence interval with the Palestine scorecard and the new-definition national line (Figure 7) is ± 0.694 percentage points. Thus for $n = 8,192$, the ratio of the two intervals is $0.694 \div 0.729 = 0.95$.

This ratio of 0.95 for $n = 8,192$ is close to the ratio of 0.93 for $n = 16,384$. It turns out that across all sample sizes of 256 or more in Figure 7, the average ratio is 0.93, implying that confidence intervals for indirect estimates of poverty rates via the Palestine scorecard and 100% of the new-definition national poverty line are—for a given sample size—about 7-percent narrower than confidence intervals for direct estimates via the 2011 PECS. This 0.93 appears in Figure 8 as the “ α factor for precision” because if $\alpha = 0.93$, then the formula for confidence intervals c for the Palestine scorecard is $\pm c = \pm z \cdot \alpha \cdot \sigma$. That is, the formula for the standard error σ for point-in-time estimates of poverty rates via scoring is $\alpha \cdot \sqrt{\frac{\hat{p} \cdot (1 - \hat{p})}{n}} \cdot \sqrt{\frac{N - n}{N - 1}}$.

In general, α can be more or less than 1.00. When α is less than 1.00, it means that the scorecard is more precise than direct measurement. This occurs for 13 of 16 poverty lines in Figure 8.

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 \tilde{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

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

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),

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

To illustrate how to use this, suppose the population N is 697,827 (the estimated number of households in Palestine while the 2011 PECS was in the field), suppose $c = 0.03808$, $z = 1.64$ (90-percent confidence), and the relevant poverty line is 100% of the new-definition national line so that the most sensible expected poverty rate \tilde{p} is Palestine’s overall poverty rate for that line in 2011 (20.3 percent at the household level, Figure 1). The α factor is 0.93 (Figure 8). Then the sample-size formula gives

$$n = 697,827 \cdot \left(\frac{1.64^2 \cdot 0.93^2 \cdot 0.203 \cdot (1 - 0.203)}{1.64^2 \cdot 0.93^2 \cdot 0.203 \cdot (1 - 0.203) + 0.03808^2 \cdot (697,827 - 1)} \right) = 260,$$

which is very close to the sample size of 256 observed for these parameters in Figure 7 for 100% of the new-definition national line. Taking the finite population correction factor ϕ as

$$\text{one (1) gives the same answer, as } n = \left(\frac{0.93 \cdot 1.64}{0.03808} \right)^2 \cdot 0.203 \cdot (1 - 0.203) = 260.^{25}$$

²⁵ Although USAID has not specified required confidence levels nor intervals, IRIS Center (2007a and 2007b) says that a sample size of $n = 300$ is sufficient for USAID reporting. USAID microenterprise partners in Palestine should report using the new-definition median line. Given an α factor of 0.93 for this line (Figure 8), an expected before-measurement household-level poverty rate of 9.0 percent (the all-Palestine

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

In practice after the end of fieldwork for the PECS in January 2012, a program would select a poverty line (say, 100% of the new-definition national line), note its participants' population size (say, $N = 10,000$ 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 \tilde{p} (perhaps based on a previous measurement such as the household-level poverty rate for 100% of the new-definition national line for Palestine overall of 20.3 percent in the 2011 PECS in Figure 1), look up α (here, 0.93, Figure 8), assume that the scorecard will still work in the future and for non-nationally representative sub-groups,²⁶ and then compute the required sample size. In this illustration,

$$n = 10,000 \cdot \left(\frac{1.64^2 \cdot 0.93^2 \cdot 0.203 \cdot (1 - 0.203)}{1.64^2 \cdot 0.93^2 \cdot 0.203 \cdot (1 - 0.203) + 0.02^2 \cdot (10,000 - 1)} \right) = 861.$$

household-level rate for the median line in 2011, Figure 1), and a confidence level of 90 percent (so that $z = 1.64$), then $n = 300$ implies a confidence interval of

$$\pm 0.93 \cdot 1.64 \cdot \sqrt{\frac{0.090 \cdot (1 - 0.090)}{300}} = \pm 2.5 \text{ percentage points.}$$

²⁶ This paper reports accuracy for the scorecard applied to the 2011 validation sample, but it cannot test accuracy for later years or for other groups. Performance after January 2012 will resemble that in the 2011 PECS with deterioration over time to the extent that the relationships between indicators and poverty status change.

7. Estimates of changes in poverty rates over time

The change in a group's poverty rate between two points in time is estimated as the change in the average poverty likelihood of the households in the group. This paper does not test estimates of change over time for Palestine,²⁷ so it can only suggest approximate formulas for standard errors. Nevertheless, the relevant concepts are presented here because, in practice, pro-poor organizations in Palestine can apply the scorecard to collect their own data and measure change through time.

7.1 Warning: Change is not impact

Scoring can estimate change. Of course, poverty could get better or worse, and scoring does not indicate what caused change. This point is often forgotten or confused, so it bears repeating: the scorecard simply estimates change, and it does not, in and of itself, indicate the reason for the change. In particular, estimating the impact of participation requires knowing what would have happened to participants if they had not been participants. Knowing this requires either strong assumptions or a control group that resembles participants in all ways except participation. To belabor the point, the scorecard can help estimate the impact of participation only if there is some way to know—or explicit assumptions about—what would have happened in the

²⁷ It is not possible to estimate changes between 2007 and 2011 (or between 2005 and 2011) based on the new 2011 scorecard because the 2005 and 2007 PECS do not include all the indicators in the new 2011 scorecard (they omit owning a filter).

absence of participation. And that information must come from somewhere beyond the scorecard.

7.2 Calculating estimated changes in poverty rates over time

Consider the illustration begun in the previous section. On 1 January 2014, an organization samples three households who score 20, 30, and 40 and so have poverty likelihoods of 74.1, 48.1, and 32.0 percent (100% of the new-definition national line, Figure 3). Adjusting for the known bias of +1.3 percentage points (Figure 8), the group's baseline estimated poverty rate is the households' average poverty likelihood of $[(74.1 + 48.1 + 32.0) \div 3] - (+1.3) = 50.1$ percent.

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

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

By way of illustration, suppose that two years later on 1 January 2016, the organization samples three additional households who are in the same population as the three original households (or suppose that the same three original households are scored a second time) and finds that their scores are now 25, 35, and 45 (poverty likelihoods of 56.4, 40.1, and 23.6 percent, 100% of the new-definition national line, Figure 3).

Adjusting for the known bias, the average poverty likelihood at follow-up is $[(56.4 +$

$40.1 + 23.6) \div 3] - (+1.3) = 38.7$ percent, an improvement of $50.1 - 38.7 = 11.4$ percentage points.²⁸

Thus, about one in nine participants in this hypothetical example crossed the poverty line in 2014/6.²⁹ Among those who started below the line, about one in five ($11.4 \div 50.1 = 22.8$ percent) on net ended up above the line.³⁰

7.3 Accuracy for estimated change in two independent samples

This paper does not measure the accuracy of scorecard estimates of changes in groups' poverty rates over time. In practice, of course, local pro-poor organizations can still use the Palestine scorecard to estimate change. The rest of this section suggests approximate formulas for standard errors that may be used.

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 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,³¹ and α is the average (across a range of bootstrapped sample sizes) of

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 people start above the line and end below it, and vice versa.

³⁰ The scorecard does not reveal the reasons for this change.

the ratio of the observed confidence interval from a scorecard and the theoretical confidence interval under direct measurement.

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

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

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

This α has been measured for 11 countries (Schreiner, 2013a, 2013b, 2012c, 2010b, 2009a, 2009b, 2009c, 2009d; Chen and Schreiner, 2009; and Schreiner and Woller, 2010a and 2010b). The simple average of α across countries—after averaging α across poverty lines and survey years within each country—is 1.15. This is as reasonable a figure as any to use for Palestine.

To illustrate how 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 new-definition national line, $\alpha = 1.15$, $\hat{p} = 0.203$ (the household-level poverty rate in 2011 for 100% of the new-definition national line in Figure 1), and the population N is large enough relative to the expected sample size n

³¹ This means that—for a given level of precision—estimating the change in a poverty rate between two points in time requires four times as many measurements (not twice as many) as does estimating a poverty rate at a point in time.

that the finite population correction ϕ can be taken as one. Then the baseline sample

size is $n = 2 \cdot \left(\frac{1.15 \cdot 1.64}{0.02} \right)^2 \cdot 0.203 \cdot (1 - 0.203) \cdot 1 = 2,878$, and the follow-up sample

size is also 2,878.

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:³²

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

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.

³² See McNemar (1947) and Johnson (2007). John Pezzullo helped find this formula.

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:

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

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

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

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

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

In Peru (the only source of a data-based estimate, Schreiner, 2009e), 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 new-definition national line, the sample will first be scored in 2014 and then again in 2017 ($y = 3$), and the population N is so large

relative to the expected sample size n that the finite population correction ϕ can be taken as one. The pre-baseline poverty rate p_{2011} is taken as 20.3 percent (Figure 1), and α is assumed to be 1.30. Then the baseline sample size is

$$n = 2 \cdot \left(\frac{1.30 \cdot 1.64}{0.02} \right)^2 \cdot \{-0.02 + 0.016 \cdot 3 + 0.47 \cdot [0.203 \cdot (1 - 0.203)]\} \cdot 1 = 2,365. \text{ The}$$

same group of 2,365 households is scored at follow-up as well.

8. Targeting

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

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

Targeting is successful when households truly below a poverty line are targeted (*inclusion*) and when households truly above a poverty line are not targeted (*exclusion*). Of course, no scorecard is perfect, and targeting is unsuccessful when households truly below a poverty line are not targeted (*undercoverage*) or when households truly above a poverty line are targeted (*leakage*).

Figure 9 depicts these four possible targeting outcomes. Targeting accuracy varies by the cut-off score; a higher cut-off has better inclusion (but worse leakage), while a lower cut-off has better exclusion (but worse undercoverage).

Programs should weigh these trade-offs when setting a cut-off. A formal way to do this is to assign net benefits—based on a program’s values and mission—to each of

the four possible targeting outcomes and then to choose the cut-off that maximizes total net benefits (Adams and Hand, 2000; Hoadley and Oliver, 1998).

Figure 10 shows the distribution of households by targeting outcome for Palestine.³³ For an example cut-off of 34 or less, outcomes for 100% of the new-definition national line in the validation sample are:

- Inclusion: 8.6 percent are below the line and correctly targeted
- Undercoverage: 11.7 percent are below the line and mistakenly not targeted
- Leakage: 5.5 percent are above the line and mistakenly targeted
- Exclusion: 74.3 percent are above the line and correctly not targeted

Increasing the cut-off to 39 or less improves inclusion and undercoverage but worsens leakage and exclusion:

- Inclusion: 11.8 percent are below the line and correctly targeted
- Undercoverage: 8.5 percent are below the line and mistakenly not targeted
- Leakage: 11.2 percent are above the line and mistakenly targeted
- Exclusion: 68.5 percent are above the line and correctly not targeted

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

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

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

- Assign benefits and costs to possible outcomes, based on its values and mission
- Tally total net benefits for each cut-off using Figure 10 for a given poverty line
- Select the cut-off with the highest total net benefit

³³ The new 2011 scorecard’s targeting accuracy is reported only for new-definition lines.

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:

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

Figure 10 shows the hit rate for all cut-offs for the Palestine scorecard. For 100% of the new-definition national line in the validation sample, total net benefit is greatest (82.8) for a cut-off of 34 or less, with almost five in six households in Palestine 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, it can reflect this by setting the benefit for inclusion to 2 and the benefit for exclusion to 1. Then the chosen cut-off will maximize $(2 \times \text{Households correctly included}) + (1 \times \text{Households correctly excluded})$.³⁴

³⁴ Figure 10 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

As an alternative to assigning benefits and costs to targeting outcomes and then choosing a cut-off to maximize total net benefits, a program could set a cut-off to achieve a desired poverty rate among targeted households. The third column of Figure 11 (“% targeted HHs who are poor”) shows, for the Palestine 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 new-definition national line, targeting households in the validation sample who score 34 or less would target 14.0 percent of all households (second column) and be associated with a poverty rate among those targeted of 61.1 percent (third column).

Figure 11 also reports two other measures of targeting accuracy. The first is a version of coverage (“% poor HHs who are targeted”). For the example of 100% of the new-definition national line with the validation sample and a cut-off of 34 or less, 42.3 percent of all poor households are covered.

The final targeting measure in Figure 11 is the number of successfully targeted poor households for each non-poor household mistakenly targeted (right-most column). For 100% of the new-definition national line with the validation sample and a cut-off of 34 or less, covering 1.6 poor households means leaking to 1 non-poor household.

targeting inclusion. $BPAC = (Inclusion - |Undercoverage - Leakage|) \times [100 \div (Inclusion + Undercoverage)]$. Schreiner (2014) explains why BPAC is not a useful measure of accuracy.

9. Context of poverty-assessment tools in Palestine

This section discusses an existing poverty-assessment tool³⁵ for Palestine in terms of its goals, methods, definition of *poverty*, data, indicators, cost, bias, and precision. In general, the advantages of the new 2011 scorecard are its:

- Use of data from the latest nationally representative consumption survey
- Accuracy for targeting that is probably similar to that of alternatives
- Reporting of bias and precision from out-of-sample tests, including formulas for standard errors
- Feasibility for local, pro-poor programs, due to its simplicity and transparency

PCBS (2013) use the “poverty mapping” approach (Elbers, Lanjouw, and Lanjouw, 2003) to estimate poverty rates for small localities. The goal is “to clarify the causes and scale of poverty . . . and to raise awareness of the factors related to poverty in Palestine in order to build programs to protect and meet the needs of Palestinian society, especially the most disadvantaged areas” (pp. 11–13).

The PCBS regresses—separately for the West Bank and the Gaza Strip—the logarithm of per-adult-equivalent consumption against indicators found in both the 2009 PECS and the 2007 census. Plugging census data into the two tools produces household-level estimates of consumption and then of poverty status based on 100% of the new-definition national poverty line. Aggregating the household-level estimates gives estimates of poverty rates for small localities that are more precise than would be

³⁵ Schreiner (2010a) reviews an early poverty-mapping exercise for Palestine by Astrup and Dessus (2005) that uses data from the 1997 census and the 1996, 1997, and 1998 PECS. Schreiner (2010a) also reviews IRIS Center’s (2009) “Poverty Assessment Tool”, which is like the scorecard tool (Schreiner, 2014) but uses data from the 2007 PECS.

possible with only the 2009 PECS. Finally, PCBS makes “poverty maps” that quickly show how estimated poverty rates vary across localities in a way that makes sense to lay people.

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

- Build poverty-assessment tools with survey data that is representative of one area (either Palestine as a whole, the West Bank, or the Gaza Strip) and then apply them to other data on groups that are not, in general, representative of the same area
- Estimate poverty rates for groups
- Provide unbiased estimates when their assumptions hold
- Report standard errors (or confidence intervals) for their estimates
- Test accuracy empirically
- Report bias

Strengths of poverty mapping include that it:

- Has formally established theoretical properties
- Can be applied straightforwardly to measures of well-being beyond poverty rates
- Requires data on fewer households for construction and calibration
- Uses only indicators that appear in a census

Strengths of the scorecard include that it:

- Uses simple, verifiable indicators that are quick and inexpensive to collect
- Is more transparent in terms of both construction and application
- Surfaces estimates of poverty likelihoods for individual households
- Reduces overfitting by selecting indicators with statistical and non-statistical criteria
- Reports simple formulas for standard errors
- Seeks to be useful in practice and so aims to be transparent and thus understandable for non-specialists

The basic difference between the two approaches is that poverty mapping seeks to help governments to target pro-poor policies, while the scorecard seeks to help local

pro-poor organizations to manage their social performance.³⁶ On a technical level, PCBS estimates consumption directly, whereas the scorecard estimates poverty likelihoods.³⁷

The indicators in the PCBS tools for the West Bank and the Gaza Strip are:

- Demographics:
 - Number of household members (and its square)
 - Number of household members who are adult females
 - Number of household members who are working-age males
 - Share of household members who are 17-years-old or younger
 - Whether any household member is disabled (interacted with governorate)
 - Refugee status (interacted with governorate)
- Education of household members:
 - Whether any completed secondary school
 - Whether any completed more than secondary school
- Residence:
 - Locality type (interacted with governorate and an asset index)
 - Tenancy status
 - Electrification
 - Number of household members per room
 - Location in Gaza City, or Rafah
 - Governorate
- Ownership of consumer durables:
 - Asset index (interacted with governorate)
 - Car
- At the level of primary-sampling units, average share of household heads who work:

³⁶ Another apparent difference is that the developers of poverty mapping (Elbers, Lanjouw, and Lanjouw, 2003; Demombynes *et al.*, 2004) say that poverty mapping is too inaccurate to be used for targeting at the household level. In contrast, Schreiner (2008b) supports household-level targeting as a legitimate, potentially useful application of the scorecard. In Elbers *et al.* (2007), the developers of poverty mapping seem to take a small step away from their original position.

³⁷ Haslett and Jones (2006, p. 61) note that “the benefits of [poverty mapping] accrue when interest is in several non-linear functions of the same target variable [such as the consumption-based poverty gap] . . . or in distributional properties. If only a single measure were of interest, it might be worthwhile to consider direct modelling of this. For example, small-area estimates of poverty incidence could be derived by estimating a logistic regression model for incidence in the survey data”. This is precisely what the scorecard does.

- Part-time
- In finance
- In manufacturing
- In “other”
- In agriculture (interacted with governorate)
- In commerce (interacted with governorate)
- In construction (interacted with governorate)
- In trade and real estate (interacted with governorate)
- In unemployment (interacted with governorate)
- Full-time (interacted with governorate and tenancy in residence)

Local, pro-poor organizations in Palestine could not apply the PCBS tools. The indicators are complex, the make-up of the asset index is not reported, and census averages at the level of the primary sampling unit are not available.

While PCBS reports standard errors for estimated poverty rates at the governorate level, they do not report standard-error formula nor the sample sizes that would allow the derivation of such a formula. Thus, the precision of their estimates cannot be compared with those of the scorecard here.

PCBS (p. 22) compares governorate-level estimates of poverty rates from the 2009 PECS and their poverty-mapping exercise, saying that “the results are remarkably consistent with the poverty rates derived from PECS, with all model predictions lying within survey confidence intervals”. This may not be surprising, as the average governorate’s 95-percent confidence interval for PECS-based estimates is ± 10.8 percentage points. Furthermore, the PCBS tools may be overfit due to their many interaction terms, some of which may act as fixed effects for a governorate. The average absolute bias across governorates for the PCBS tools with the 2007 census data and 100% of the new-definition national line is 3.7 percentage points.³⁸

³⁸ There is no comparable figure for the new 2011 scorecard here because the 2011 PECS database received from the PCBS does not include the governorate.

10. Conclusion

Pro-poor programs in Palestine (West Bank and Gaza Strip) the scorecard here to estimate the likelihood that a household has consumption below a given poverty line, to estimate the poverty rate of a group of households at a point in time, and to estimate changes in the poverty rate of a group of households between two points in time. The scorecard can also be used for targeting.

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

The scorecard is constructed with data from Palestine's 2011 PECS. It replaces earlier scorecards based on the 2005 and 2007 PECS by Schreiner (2010a and 2008c). From now on, only the new 2011 scorecard should be used because it is based on the latest data and because it is calibrated to Palestine's new definition of *poverty*. The new scorecard is also calibrated to six old-definition (legacy) poverty lines that follow the definition of *poverty* used by Schreiner (2010a) with the old 2005 and 2007 scorecards. Estimated poverty likelihoods from the three scorecards are compatible for the old-definition (legacy) lines, so existing users can—if desired—estimate changes over time with a baseline from the old 2005 or old 2007 scorecards and a follow-up from the new 2011 scorecard.

The new 2011 scorecard is constructed with half of the 2011 PECS data, calibrated to ten new-definition poverty lines and six old-definition lines, and tested on the other half of the 2011 data.

Bias and precision are reported for estimates of households' poverty likelihoods, groups' poverty rates at a point in time, and changes in groups' poverty rates over time. Of course, the scorecard's estimates of changes are not the same as estimates of program impact. Targeting accuracy is also reported.

When the scorecard is applied to the validation sample, the maximum absolute average difference between estimates versus true poverty rates for groups of households at a point in time across all 16 new-definition and old-definition poverty lines is 3.1 percentage points. Across the 16 lines, the average absolute average difference is about 1.0 percentage points. Unbiased estimates may be had by subtracting the known bias from the original estimates. For $n = 16,384$ and 90-percent confidence, the precision of these differences is ± 0.6 percentage points or better.

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

Although the statistical technique is innovative, and although technical accuracy is important, the design of the scorecard focuses on transparency and ease-of-use. After all, accuracy is irrelevant if an organization feels so daunted by a scorecard's complexity or its cost that it does not even try to use it. For this reason, the scorecard is kept simple, using ten indicators that are simple, 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 adoption by helping managers understand and trust scoring and by allowing non-specialists to generate scores quickly in the field.

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

References

- Adams, Niall M.; and David J. Hand. (2000) “Improving the Practice of Classifier Performance Assessment”, *Neural Computation*, Vol. 12, pp. 305–311.
- Astrup, Claus; and Sébastien Dessus. (2005) “Targeting the Poor beyond Gaza or the West Bank: The Geography of Poverty in the Palestinian Territories”, *Région et Développement*, No. 21, pp. 173–197.
- Baesens, Bart; Van Gestel, Tony; Viaene, Stijn; Stepanova, Maria; Suykens, Johan A. K.; and Jan Vanthienen. (2003) “Benchmarking State-of-the-Art Classification Algorithms for Credit Scoring”, *Journal of the Operational Research Society*, Vol. 54, pp. 627–635.
- Caire, Dean. (2004) “Building Credit Scorecards for Small-Business Lending in Developing Markets”, microfinance.com/English/Papers/Scoring_SMEs_Hybrid.pdf, retrieved 4 June 2014.
- ; and Mark Schreiner. (2012) “Cross-Tab Weighting for Credit Scorecards in Developing Markets”, dean_caire@hotmail.com.
- Camacho, Adriana; and Emily Conover. (2011) “Manipulation of Social-Program Eligibility”, *American Economic Journal: Economic Policy*, Vol. 3, No. 2, pp. 41–65.
- Chen, Shiyuan; and Mark Schreiner. (2009) “Simple Poverty Scorecard Poverty-Assessment Tool: Vietnam”, SimplePovertyScorecard.com/VNM_2006_ENG.pdf, retrieved 7 July 2016.
- Coady, David; Grosh, Margaret; and John Hoddinott. (2004) *Targeting of Transfers in Developing Countries*, hdl.handle.net/10986/14902, retrieved 13 May 2016.
- Cochran, William G. (1977) *Sampling Techniques, Third Edition*.
- Dawes, Robyn M. (1979) “The Robust Beauty of Improper Linear Models in Decision-Making”, *American Psychologist*, Vol. 34, No. 7, pp. 571–582.
- Demombynes, Gabriel; Elbers, Chris; Lanjouw, Jenny; Lanjouw, Peter; Mistiaen, Johan; and Berk Özler. (2004) “Producing an Improved Geographic Profile of Poverty: Methodology and Evidence from Three Developing Countries”, pp. 154–176 in Anthony Shorrocks and Rolph van der Hoeven (eds.) *Growth, Inequality, and Poverty*.

- Elbers, Chris; Fujii, Tomoki; Lanjouw, Peter; Özler, Berk; and Wesley Yin. (2007) “Poverty Alleviation through Geographic Targeting: How Much Does Disaggregation Help?”, *Journal of Development Economics*, Vol. 83, pp. 198–213.
- ; Lanjouw, Jean O.; and Peter Lanjouw. (2003) “Micro-Level Estimation of Poverty and Inequality”, *Econometrica*, Vol. 71, No. 1, pp. 355–364.
- Friedman, Jerome H. (1997) “On Bias, Variance, 0–1 Loss, and the Curse-of-Dimensionality”, *Data Mining and Knowledge Discovery*, Vol. 1, pp. 55–77.
- Fuller, Rob. (2006) “Measuring the Poverty of Microfinance Clients in Haiti”, microfinance.com/English/Papers/Scoring_Poverty_Haiti_Fuller.pdf, retrieved 4 June 2014.
- 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/W90WN57PDO, retrieved 4 June 2014.
- 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.
- ; and Geoffrey Jones. (2006) “Small-Area Estimation of Poverty, Caloric Intake, and Malnutrition in Nepal”, un.org.np/node/10501, retrieved 4 June 2014.
- Hoadley, Bruce; and Robert M. Oliver. (1998) “Business Measures of Scorecard Benefit”, *IMA Journal of Mathematics Applied in Business and Industry*, Vol. 9, pp. 55–64.
- IRIS Center. (2009) “Poverty Assessment Tool Accuracy Submission: USAID/IRIS Tool for West Bank”, 30 September, povertytools.org/countries/West_Bank/USAID_PAT_West_Bank.pdf, retrieved 4 June 2014.
- (2007a) “Manual for the Implementation of USAID Poverty Assessment Tools”, povertytools.org/training_documents/Manuals/USAID_PAT_Manual_Eng.pdf, retrieved 4 June 2014.

- (2007b) “Introduction to Sampling for the Implementation of PATs”, povertytools.org/training_documents/Sampling/Introduction_Sampling.pdf, retrieved 4 June 2014.
- (2005) “Notes on Assessment and Improvement of Tool Accuracy”, povertytools.org/other_documents/AssessingImproving_Accuracy.pdf, retrieved 4 June 2014.
- Johnson, Glenn. (2007) “Lesson 3: Two-Way Tables—Dependent Samples”, onlinecourses.science.psu.edu/stat504/node/96, retrieved 4 June 2014.
- Kolesar, Peter; and Janet L. Showers. (1985) “A Robust Credit-Screening Model Using Categorical Data”, *Management Science*, Vol. 31, No. 2, pp. 124–133.
- Lovie, Alexander D.; and Patricia Lovie. (1986) “The Flat-Maximum Effect and Linear Scoring Models for Prediction”, *Journal of Forecasting*, Vol. 5, No. 3, 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.
- Matul, Michal; and Sean Kline. (2003) “Scoring Change: Prizma’s Approach to Assessing Poverty”, Microfinance Centre for Central and Eastern Europe and the New Independent States Spotlight Note No. 4, mfc.org.pl/sites/mfc.org.pl/files/spotlight4.PDF, retrieved 4 June 2014.
- McNemar, Quinn. (1947) “Note on the Sampling Error of the Difference between Correlated Proportions or Percentages”, *Psychometrika*, Vol. 17, pp. 153–157.
- Myers, James H.; and Edward W. Forgy. (1963) “The Development of Numerical Credit-Evaluation Systems”, *Journal of the American Statistical Association*, Vol. 58, No. 303, pp. 779–806.
- Narayan, Ambar; and Nobuo Yoshida. (2005) “Proxy-Means Tests for Targeting Welfare Benefits in Sri Lanka”, World Bank Report No. SASPR-7, documents.worldbank.org/curated/en/2005/07/6209268/proxy-means-test-targeting-welfare-benefits-sri-lanka, retrieved 4 June 2014.
- 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.

- Palestinian Central Bureau of Statistics. (2013) “Atlas of Poverty in Palestine”,
pcbs.gov.ps/Portals/_PCBS/Downloads/book1987.pdf, retrieved 4 June 2014.
- (2012a) “Living Standards in the Palestinian Territory: Expenditure,
Consumption, and Poverty in 2011”,
pcbs.gov.ps/Portals/_PCBS/Downloads/Book1899.pdf, retrieved 4 June 2014.
- (2012b) “Press Release: Levels of Living and Poverty in the Palestinian Territory,
2011”, [82.213.38.42/Portals/_pcbs/PressRelease/ Living_2011e.pdf](http://82.213.38.42/Portals/_pcbs/PressRelease/Living_2011e.pdf),
retrieved 4 June 2014.
- (2010) “Main Poverty Characteristics in the Palestinian Territory 2004–9,
according to the Revised Poverty-Rate Estimation Methodology”,
sesrtcic.org/files/nso/1031.pdf, retrieved 4 June 2014.
- SAS Institute Inc. (2004) “The LOGISTIC Procedure: Rank Correlation of Observed
Responses and Predicted Probabilities”, in *SAS/STAT User’s Guide, Version 9*,
[support.sas.com/documentation/cdl/en/statug/63033/HTML/
default/viewer.htm#statug_logistic_sect035.htm](http://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_logistic_sect035.htm), retrieved 4 June 2014.
- Schreiner, Mark. (2014) “How Do the the Simple Poverty Scorecard and the PAT
Differ?”, microfinance.com/English/Papers/Scorecard_versus_PAT.pdf,
retrieved 4 June 2014.
- (2013a) “Simple Poverty Scorecard Poverty-Assessment Tool: Bangladesh”,
SimplePovertyScorecard.com/BGD_2010_ENG.pdf, retrieved 7 July 2016.
- (2013b) “Simple Poverty Scorecard Poverty-Assessment Tool: Nicaragua”,
SimplePovertyScorecard.com/NIC_2009_ENG.pdf, retrieved 7 July 2016.
- (2012a) “An Expert-Based Poverty Scorecard for Rural China”,
microfinance.com/English/Papers/Scoring_Poverty_China_EN.pdf, retrieved
4 June 2014.
- (2012b) “Simple Poverty Scorecard Poverty-Assessment Tool: Colombia”,
SimplePovertyScorecard.com/COL_2009_ENG.pdf, retrieved 7 July 2016.
- (2012c) “Simple Poverty Scorecard Poverty-Assessment Tool: Peru”,
SimplePovertyScorecard.com/PER_2010_ENG.pdf, retrieved 7 July 2016.
- (2010a) “Simple Poverty Scorecard Poverty-Assessment Tool: Palestine”,
SimplePovertyScorecard.com/PSE_2007_ENG.pdf, retrieved 7 July 2016.

- (2010b) “Simple Poverty Scorecard Poverty-Assessment Tool: Honduras”, SimplePovertyScorecard.com/GTM_2006_ENG.pdf, retrieved 7 July 2016.
- (2009a) “Simple Poverty Scorecard Poverty-Assessment Tool: Philippines”, SimplePovertyScorecard.com/PHL_2002_ENG.pdf, retrieved 7 July 2016.
- (2009b) “Simple Poverty Scorecard Poverty-Assessment Tool: Pakistan”, SimplePovertyScorecard.com/PAK_2005_ENG.pdf, retrieved 7 July 2016.
- (2009c) “Simple Poverty Scorecard Poverty-Assessment Tool: Bolivia”, SimplePovertyScorecard.com/BOL_2007_ENG.pdf, retrieved 7 July 2016.
- (2009d) “Simple Poverty Scorecard Poverty-Assessment Tool: Mexico”, SimplePovertyScorecard.com/MEX_2008_ENG.pdf, retrieved 7 July 2016.
- (2009e) “Simple Poverty Scorecard Poverty-Assessment Tool: Peru”, SimplePovertyScorecard.com/PER_2007_ENG.pdf, retrieved 7 July 2016.
- (2008a) “Simple Poverty Scorecard Poverty-Assessment Tool: Peru”, SimplePovertyScorecard.com/PER_2003_ENG.pdf, retrieved 7 July 2016.
- (2008b) “Simple Poverty Scorecard Poverty-Assessment Tool: Ecuador”, SimplePovertyScorecard.com/ECU_2005_ENG.pdf, retrieved 7 July 2016.
- (2008c) “Simple Poverty Scorecard Poverty-Assessment Tool: Palestine”, SimplePovertyScorecard.com/PSE_2005_ENG.pdf, retrieved 7 July 2016.
- (2006) “Is One Simple Poverty Scorecard Poverty-Assessment Tool Enough for India?”, microfinance.com/English/Papers/Scoring_Poverty_India_Segments.pdf, retrieved 4 June 2014.
- (2005a) “La Herramienta del Índice de Calificación de la Pobreza™: México”, SimplePovertyScorecard.com/MEX_2002_SPA.pdf, retrieved 7 July 2016.
- (2005b) “IRIS Questions on the Simple Poverty Scorecard Poverty-Assessment Tool”, microfinance.com/English/Papers/Scoring_Poverty_Response_to_IRIS.pdf, retrieved 4 June 2014
- (2002) *Scoring: The Next Breakthrough in Microfinance?* CGAP Occasional Paper No. 7, microfinance.com/English/Papers/Scoring_Breakthrough_CGAP.pdf, retrieved 13 May 2016.

- ; Matul, Michal; Pawlak, Ewa; and Sean Kline. (2004) “Poverty Scoring: Lessons from a Microlender in Bosnia-Herzegovina”, microfinance.com/English/Papers/Scoring_Poverty_in_BiH_Short.pdf, retrieved 4 June 2014.
- ; and Gary Woller. (2010a) “Simple Poverty Scorecard Poverty-Assessment Tool: Ghana”, SimplePovertyScorecard.com/GHA_2005_ENG.pdf, retrieved 7 July 2016.
- ; and Gary Woller. (2010b) “Simple Poverty Scorecard Poverty-Assessment Tool: Guatemala”, SimplePovertyScorecard.com/GTM_2006_ENG.pdf, retrieved 7 July 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 4 June 2014.
- Sillers, Don. (2006) “National and International Poverty Lines: An Overview”, pdf.usaid.gov/pdf_docs/Pnadh069.pdf, retrieved 4 June 2014.
- 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.
- Sun, Changqing; and Eric Swanson. (2009) “Estimation of PPPs for Non-Benchmark Economies for the 2005 ICP Round”, *ICP Bulletin*, Vol. 6, No. 1, pp. 20–23.
- 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.
- Toohig, Jeff. (2008) “PPI Pilot Training Guide”, microfinancegateway.org/gm/document-1.1.6364/PPITrainingGuide.pdf, retrieved 4 June 2014.
- United States Congress. (2004) “Microenterprise Results and Accountability Act of 2004 (HR 3818 RDS)”, November 20, smith4nj.com/laws/108-484.pdf, retrieved 4 June 2014.
- Wainer, Howard. (1978) “On the Sensitivity of Regression and Regressors”, *Psychological Bulletin*, Vol. 85, pp. 267–73.

World Bank. (2012) *Targeting Poor and Vulnerable Households in Indonesia*,
ausaid.gov.au/Publications/Pages/report-targeting-poor-households-indonesia.aspx, retrieved 4 June 2014.

----- (2004) “Deep Palestinian Poverty in the Midst of Economic Crisis”,
go.worldbank.org/KZNKZZBYX0, retrieved 4 June 2014.

Zeller, Manfred. (2004) “Review of Poverty Assessment Tools”,
pdf.usaid.gov/pdf_docs/PNADH120.pdf, retrieved 4 June 2014.

Calculating a Hybrid Estimate of Change in Poverty Rates over Time

This appendix describes a step-by-step process that allows existing users of the old 2005 or 2007 scorecard to calculate hybrid estimates of changes in poverty rates over time. The process allows legacy users to salvage past applications of the old 2005 or 2007 scorecard, and it also allows all users to make on-going estimates of change based on current and future applications of the new 2011 scorecard.

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

- The “past” (only the old 2005 or 2007 scorecards, considering only old-definition (legacy) lines),
- “Now” (only the new 2011 scorecard, considering both old-definition (legacy) lines and new-definition lines), and
- The “future” (only the new 2011 scorecard, considering only new-definition lines).

The procedure is as follows:

1. Select a old-definition (legacy) poverty line from among those supported in this paper (deep; 100%, 150%, or 200% of national; \$3.75/day, or \$5.00/day 2005 PPP)³⁹
2. Estimate a baseline poverty rate for the given old-definition (legacy) line:
 - a. Retrieve (from a paper file, spreadsheet, or database) the poverty likelihood for the given old-definition (legacy) line for each household in the representative sample of a given population to whom the old 2005 or 2007 scorecard has already been applied in the past. These likelihoods are based on the look-up tables for the given old-definition (legacy) line in Schreiner, 2010a or 2008c (not the look-up tables in this paper)
 - b. Average the households’ poverty likelihoods to get an estimate of their baseline poverty rate for the given old-definition (legacy) line

³⁹ Of course, this entire process can be repeated for additional poverty lines, using the same sample of households and their scores.

3. Estimate a follow-up poverty rate for the given old-definition (legacy) line:
 - a. Apply the new 2011 scorecard to a representative sample of the same population to which the old 2005 or 2007 scorecard was originally applied⁴⁰
 - b. Add up the score for each household for the new 2011 scorecard
 - c. Convert each household's score to poverty likelihoods using the look-up tables for the given old-definition (legacy) line in this paper (not the look-up tables in Schreiner, 2010a). In this paper, the old-definition (legacy) lines are explicitly labeled as such
 - d. Average the households' poverty likelihoods to get an estimate of their follow-up poverty rate for the given old-definition (legacy) line

4. Find hybrid estimates of change for the given old-definition (legacy) line:
 - a. For a given old-definition (legacy) line, the estimated change is the estimated follow-up poverty rate (3d) minus the estimated baseline poverty rate (2b)
 - b. The estimated change relative to the share of participants who were under the given old-definition (legacy) line at baseline is the estimated change (4a) divided by the estimated baseline poverty rate (2b)
 - c. The estimated net number of participants who crossed from below the given old-definition (legacy) line to above it since baseline is the negative of the change (4a) expressed as a proportion, multiplied by the number of participants at baseline

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

5. Select a new-definition poverty line from among those supported in this paper (deep; 100%, 150%, or 200% of the national line; the median line; or \$1.25/day, \$2.00/day, \$2.50/day, \$3.75/day, or \$5.00/day 2005 PPP)

⁴⁰ What matters is that the sample be representative of the same population as that to which the old 2005 or 2007 scorecard was originally applied. The households interviewed “now” need not be the exact same as those interviewed in the “past” (although they may be). Also, keep in mind that a census is a special case of a sample.

6. Estimate a baseline poverty rate for the given new-definition line:
 - a. In addition to the sample of households to which the new 2011 scorecard was applied in (3a), apply the new 2011 scorecard to samples of households that are representative of any additional populations of interest⁴¹
 - b. Add up (or retrieve from 3b) the score for each household to which the new 2011 scorecard was applied in (3a) and each household that is part of a samples from an additional population of interest in (6a)
 - c. Convert each household's score to a poverty likelihood using the look-up tables for the given new-definition line in this paper (not the look-up tables in Schreiner, 2010a, none of which pertain to new-definition lines)
 - d. For the sample of households to which the new 2011 scorecard was applied in 3a (and separately for any households in samples that are representative of any additional populations of interest in 6a), average the households' poverty likelihoods to get an estimate of their baseline poverty rates for the given new-definition line

From this point on, all estimates of change are based solely on new-definition lines:

7. Select the new-definition poverty line for which a baseline poverty rate has been estimated in 6d
8. Estimate follow-up poverty rates for the given new-definition line:
 - a. Apply the new 2011 scorecard to a representative sample of the same population to which the new 2011 scorecard was originally applied (3a), as well as any additional populations represented in (6a)
 - b. Add up the score for each household to which the new 2011 scorecard has just been applied

⁴¹ The instructions from this step on apply to all users of the new 2011 scorecard who want to measure change over time, not just legacy users who have estimated a hybrid measure of change. New users (as well as legacy users who want the scorecard to address business questions for which the relevant population was not already sampled in the “past”) will prepare for this step by defining a business question, specifying a relevant population, and drawing a new sample from the relevant population.

- c. Convert each household's score to a poverty likelihood using the look-up tables for the given new-definition line in this paper (not the look-up tables in Schreiner, 2010a, none of which pertain to new-definition lines)
 - d. For a sample representing a given population, average the households' poverty likelihoods to get an estimate of their follow-up poverty rates for the given new-definition line
9. Find the (non-hybrid) estimate of change for the given new-definition line:
- a. For a given new-definition line, the estimated change is the estimated follow-up poverty rate (8d) minus the estimated baseline poverty rate (6d)
 - b. The estimated change relative to the share of participants who were under the given new-definition line at baseline is the change (9a) divided by the estimated baseline poverty rate (6d)
 - c. The estimated net number of participants who crossed from below a given new-definition poverty line to above it since baseline is the negative of the estimated change (9a) expressed as a proportion, multiplied by the number of participants at baseline
10. Assuming that the "parallel lines" assumption holds,⁴² find the "grand" estimate of change that combines the hybrid and non-hybrid estimates:
- a. The "grand" estimate of change is the the hybrid estimate of change (4a) for a given old-definition (legacy) line plus the non-hybrid estimate of change for the corresponding new-definition line (9a)
 - b. The "grand" estimate of change relative to the share of participants who were below the given old-definition (legacy) line in the "past" baseline is the "grand" estimate of change (10a) divided by the share of participants who were below the given old-definition (legacy) line in the "past" baseline (2b). (There is no "grand" estimate of relative change for a given new-definition line because there is no estimate of the poverty rate for any new-definition line in the "past" baseline)

⁴² As discussed in the main text of this paper, it seems unlikely that the "parallel lines" assumption would hold well for Palestine.

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

The following hypothetical example illustrates the application of above steps using poverty likelihoods for the old-definition/legacy lines and the new-definition lines to which scores from the 2011 scorecard for Palestine have been calibrated:

1. *Select an old-definition (legacy) poverty line from among those supported in this paper:*

Select 100% of the old-definition national line.

2. *Estimate a baseline poverty rate for the given old-definition (legacy) line:*
 - a. *Retrieve (from a paper file, spreadsheet, or database) the scores and the poverty likelihoods for the given old-definition (legacy) line for each household in the representative sample of a given population to whom the old 2007 scorecard has already been applied. This likelihood is based on the look-up table for the given old-definition (legacy) line in Schreiner, 2010a (not the look-up tables in this paper)*

In this hypothetical example, the scores and likelihoods for the three⁴³ households in the sample are:

Score	Poverty likelihood (100% of the old-definition (legacy) national line)
15	81.7
20	82.2
25	63.1

The poverty likelihoods for 100% of the old-definition legacy national line come from p. 65 of Schreiner (2010a).⁴⁴

⁴³ Three households is an unrealistically small sample, but it is used in this hypothetical illustration to keep the arithmetic manageable.

⁴⁴ The page number here refers to “Figure 4: (National line): Estimated poverty likelihoods associated with scores)” as of 5 June 2014 at microfinance.com/English/Papers/Scoring_Poverty_Palestine_EN_2007.pdf.

- b. *Average the households' poverty likelihoods to get an estimate of their baseline poverty rate for the old-definition (legacy)*

$$(81.7 + 82.2 + 63.1) \div 3 = 75.7 \text{ percent.}$$

3. *Estimate a follow-up poverty rate for a given old-definition (legacy) line:*

- a. *Apply the new 2011 scorecard to a representative sample of the same population to which the old 2007 scorecard was originally applied*

Draw a new sample of three households
(or interview the same three households a second time)

- b. *Add up the score for each household from the new 2011 scorecard*

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

- c. *Convert each household's score to a poverty likelihood using the look-up tables for the given old-definition (legacy) line in this paper (not the look-up tables in Schreiner, 2010a)*

Look up the poverty likelihoods for 100% of the old-definition (legacy) national line on p. 175 in this paper.

Score	Poverty likelihood (100% of the old-definition (legacy) national line)
21	76.1
26	65.7
31	57.3

- d. *Average the households' poverty likelihoods to get an estimate of their follow-up poverty rate for the given old-definition (legacy) line*

$$(76.1 + 65.7 + 57.3) \div 3 = 66.4 \text{ percent.}$$

4. *Find hybrid estimates of change for the given old-definition (legacy) line:*

a. *The hybrid estimate of change is the estimated follow-up poverty rate (3d) minus the estimated baseline poverty rate (2b)*

$$66.4 \text{ percent} - 75.7 \text{ percent} = -9.3 \text{ percentage points.}$$

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

$$-9.3 \text{ percentage points} \div 75.7 \text{ percentage points} = -12.3 \text{ percent.}$$

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

Assuming for the sake of this hypothetical illustration
that there were 10,000 participants in the baseline population,
 $-(-0.093) \times 10,000 \text{ participants} = 930 \text{ participants.}$

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

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

For compatibility with the above, select 100% of the new-definition national line.

6. *Estimate a baseline poverty rate for the given new-definition line:*

a. *In addition to samples of households that are representative of the same population as that to which the new 2011 scorecard was applied in (3a), apply the new 2011 scorecard to samples of households that are representative of any additional populations of interest*

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

- b. Add up (or retrieve from 3b) the score for each household to which the new 2011 scorecard has been applied

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

- c. Convert each household's score to a poverty likelihood using the look-up tables for the given new-definition line in this paper (not the look-up tables in Schreiner, 2010a, none of which pertain to new-definition lines)

Look up the poverty likelihoods for the new-definition national line on p. 116 in this paper.

Score	Poverty likelihood (100% of the new-definition national line)
21	74.1
26	56.4
31	48.1

- d. Average the households' poverty likelihoods to get an estimate of their baseline poverty rate for the given new-definition line

$$(74.1 + 56.4 + 48.1) \div 3 = 59.5 \text{ percent.}$$

From this point on, all estimates of change are based solely on the new-definition lines:

7. Select a new-definition poverty line for which a baseline poverty rate has been estimated in 6d

For compatibility with the above, select 100% of the new-definition national line.

8. Estimate a follow-up poverty rate for the given new-definition line:

- a. Apply the new 2011 scorecard to a representative sample of the same population to which the new 2011 scorecard was originally applied in 3a, as well as any additional populations represented in 6a)

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

- b. Add up the score for each household to which the new 2011 scorecard has just been applied

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

- c. Convert each household's score to a poverty likelihood using the look-up tables for the given new-definition line in this paper (not the look-up tables in Schreiner, 2010a, none of which pertain to new-definition lines)

Look up the poverty likelihoods for the new-definition national line on p. 116 in this paper.

Score	Poverty likelihood (100% of the new-definition national line)
22	74.1
27	56.4
37	40.1

- d. For the sample representing a given population, average the households' poverty likelihoods to get an estimate of their follow-up poverty rate for the given new-definition line

$$(74.1 + 56.4 + 40.1) \div 3 = 56.9 \text{ percent.}$$

9. Find non-hybrid estimates of change for the given new-definition line:

- a. The estimated change is the estimated follow-up poverty rate (8d) minus the estimated baseline poverty rate (6d)

$$56.9 \text{ percent} - 59.5 \text{ percent} = -2.6 \text{ percentage points.}$$

- b. The estimated change relative to the share of participants who were under the given new-definition line at baseline is the estimated change (9a) divided by the estimated baseline poverty rate (6d)

$$-2.6 \text{ percentage points} \div 59.5 \text{ percentage points} = -4.4 \text{ percent.}$$

- c. *The estimated net number of participants who crossed from below the given new-definition poverty line to above it since baseline is the negative of the change (9a) expressed as a proportion, multiplied by the number of participants at baseline*

Assuming for the sake of this hypothetical illustration that there were 10,000 participants in the baseline population,
 $-(-0.026) \times 10,000$ participants = 260 participants.

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

- a. *The “grand” estimate of change is the hybrid estimate of change for the given old-definition (legacy) line (4a) plus the non-hybrid estimate of change the given new-definition line (9a)*

-9.3 percentage points + $(-2.6$ percentage points) = -11.9 percentage points.

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

$-11.9 \div 75.7 = -15.7$ percent.

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

Assuming for the sake of this hypothetical illustration that there were 10,000 participants in the baseline population,
 $-(-0.119) \times 10,000 = 1,190$.

This page sums up the process in the hypothetical illustration above. It focuses on estimates of changes in poverty rates.

Selected poverty line: 100% of the national line (old-definition/legacy and new-definition)

Scores and poverty likelihoods of sampled households for the selected poverty line

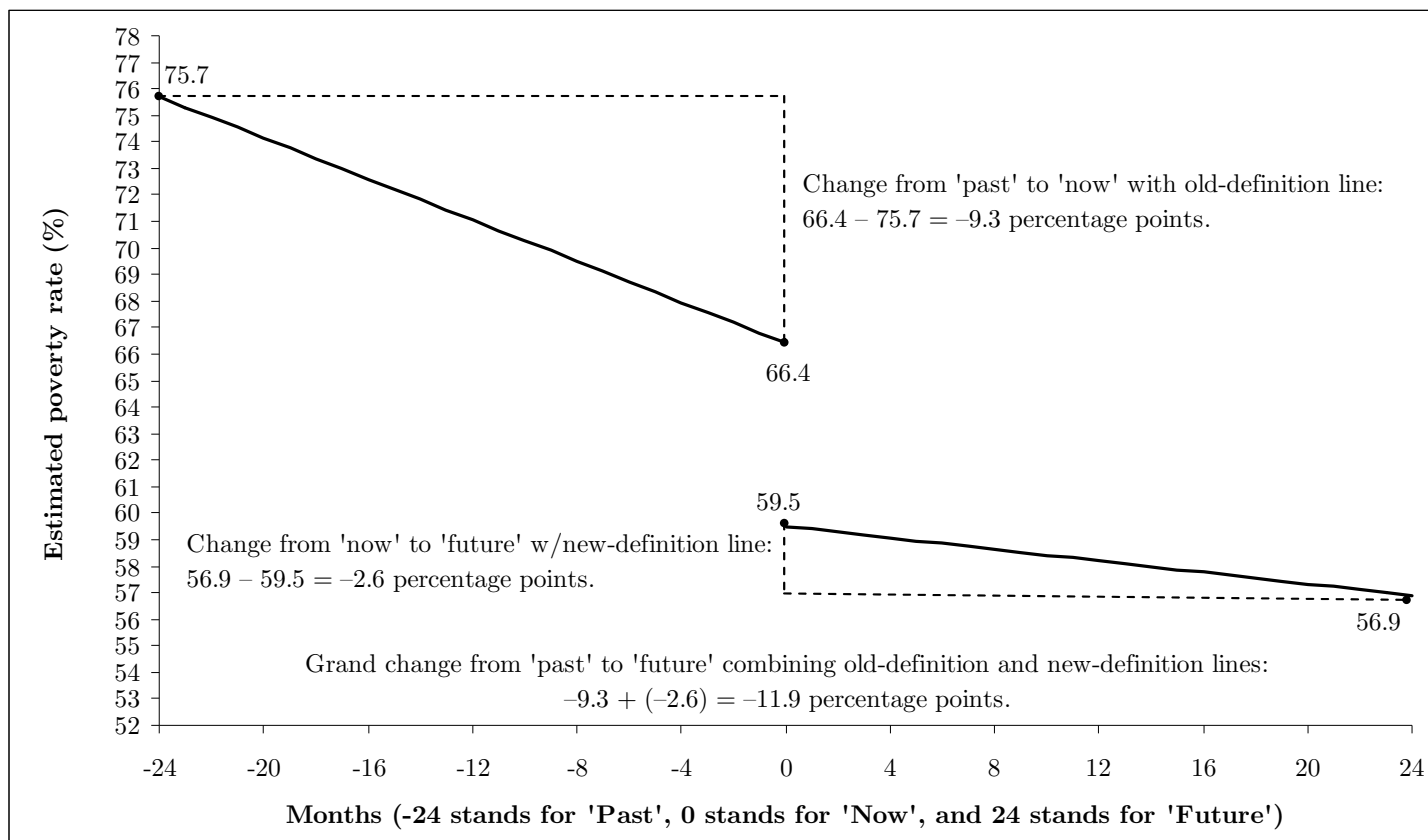
“Past”		“Now”			“Future”	
Score	Pov. like. (old def.) (%)	Score	Pov. like. (old def.) (%)	Pov. like. (new def.) (%)	Score	Pov. like. (new def.) (%)
15	81.7	21	76.1	74.1	22	74.1
20	82.2	26	65.7	56.4	27	56.4
25	63.1	31	57.3	48.1	37	40.1
Est. pov. rate (%)	75.7	—	66.4	59.5	—	56.9

Estimated change between:

“Past” and “now” (hybrid): $66.4 - 75.7 = -9.3$ percentage points

“Now” and “future” (non-hybrid): $56.9 - 59.5 = -2.6$ percentage points

“Past” and “future” (“grand”): $-9.3 + (-2.6) = -11.9$ percentage points



Guidelines for the Interpretation of Scorecard Indicators

The following is based on:

Palestinian Central Bureau of Statistics (2011) “User Manual: The Palestinian Expenditure and Consumption Survey 2011”, Ramallah, erfdataportal.com/index.php/catalog/64/download/529, retrieved 4 June 2014 [the *Manual*],

and

Palestinian Central Bureau of Statistics (2011) “Household Questionnaire: The Palestinian Expenditure and Consumption Survey 2011”, Ramallah, erfdataportal.com/index.php/catalog/64/download/527, retrieved 4 June 2014 [the *Questionnaire*],

1. How many household members are there?
 - A. Eight or more
 - B. Seven
 - C. Six
 - D. Five
 - E. Four
 - F. One, two, or three

According to p. 3 of the *Manual*, a *household* is “one person or a group of persons (with or without a family relationship) who live in the same dwelling unit, who share meals, and who make joint provisions for food and other essentials of living.”

2. How many household members are self-employed without employees, are unpaid employees, or work for an irregular wage?
 - A. Two or more
 - B. One
 - C. None

According to p. 3 of the *Manual*, a *household* is “one person or a group of persons (with or without a family relationship) who live in the same dwelling unit, who share meals, and who make joint provisions for food and other essentials of living.”

According to the PCBS (communication with Niveen Zein-Eldeen), a *self-employed* person “works in a business that completely or partly belongs to him/her and that does not hire any wage employees.” This includes the self-employed who work for themselves in their own businesses.

3. During the past week, was the male head/spouse employed for one hour or more?
 - A. No
 - B. No male head/spouse
 - C. Yes

According to p. 3 of the *Manual*, the *head of the household* is “the person who usually lives with the household and who is recognized as the head of the household by its other members. Often he/she is the main decision-maker and is responsible for the financial support and welfare of the household.”

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

- The household head, if the head is male
- The spouse/partner/companion of the household head, if the head is female
- Non-existent, if the head is female and if she does not have a spouse/partner/companion who is also a member of the household

4. During the past week, was the female head/spouse employed for one hour or more?
- A. No
 - B. Yes
 - C. No female head/spouse

According to p. 3 of the *Manual*, the *head of the household* is “the person who usually lives with the household and who is recognized as the head of the household by its other members. Often he/she is the main decision-maker and is responsible for the financial support and welfare of the household.”

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

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

5. How many rooms does the residence have (excluding kitchens and bathrooms)?
- A. One
 - B. Two
 - C. Three
 - D. Four
 - E. Five or more

There are no guidelines available for this indicator.

6. What is the main source of energy for heating the residence?
 - A. None, coal/firewood, solar, or other
 - B. Gas, kerosene, or electricity

There are no guidelines available for this indicator.

7. Does the household have a solar boiler?
 - A. No
 - B. Yes

According to the PCBS (communication with Niveen Zein-Eldeen), only solar boilers that are in working condition and that are used by the household should be counted.

8. Does the household have a filter?
- A. No
 - B. Yes

According to the PCBS (communication with Niveen Zein-Eldeen), *filter* refers to water filters. Only water filters that are in working condition and that are used by the household should be counted.

9. Does the household have a vacuum cleaner?
- A. No
 - B. Yes

According to the PCBS (communication with Niveen Zein-Eldeen), only vacuum cleaners that are in working condition and that are used by the household should be counted.

10. Does the household have a computer?

- A. No
- B. Yes

According to the PCBS (communication with Niveen Zein-Eldeen), only computers that are in working condition and that are used by the household should be counted.

Figure 1: New-definition poverty lines and poverty rates (for households and people) for all of Palestine, West Bank, and the Gaza Strip for 2005, 2007, and 2011 by poverty line and by sub-sample

Sample	Year	Line or rate	Level	n	Poverty rates (% with consumption less than a poverty line) and new-definition poverty lines (NIS per day per adult-equivalent or per person)									
					National poverty lines					Intl. 2005 PPP				
					Deep	100%	150%	200%	Median	\$1.25	\$2.00	\$2.50	\$3.75	\$5.00
<u>All Palestine</u>	2005	Line	People		—	—	—	—	—	—	—	—	—	—
		Rate	HHs	2,152	—	—	—	—	—	—	—	—	—	—
			People		—	—	—	—	—	—	—	—	—	—
<u>All Palestine</u>	2007	Line	People		—	—	—	—	—	—	—	—	—	—
		Rate	HHs	1,231	—	—	—	—	—	—	—	—	—	—
			People		—	—	—	—	—	—	—	—	—	—
<u>All Palestine</u>	2011	Line	People		19.85	24.84	37.26	49.69	11.87	3.68	5.89	7.36	11.04	14.72
		Rate	HHs	4,317	9.5	20.3	47.9	67.5	9.0	0.1	0.4	1.4	7.3	18.7
			People		12.4	25.4	54.5	73.6	12.7	0.2	0.7	2.3	10.6	25.0
<u>West Bank</u>	2011	Line	People		20.78	26.01	39.02	52.02	12.73	3.85	6.17	7.71	11.56	15.42
		Rate	HHs	2,909	5.9	14.5	39.7	59.9	6.2	0.1	0.5	0.8	4.2	12.3
			People		7.8	17.8	45.1	65.6	8.9	0.3	0.8	1.3	6.2	16.7
<u>Gaza Strip</u>	2011	Line	People		18.34	22.95	34.43	45.90	10.48	3.40	5.44	6.80	10.20	13.60
		Rate	HHs	1,408	16.3	31.5	63.7	82.2	14.1	0.0	0.3	2.6	13.2	31.0
			People		20.0	37.6	69.8	86.7	18.8	0.0	0.4	3.9	17.6	38.6
<u>Construction and calibration</u>														
Selecting indicators and points, and associating scores with likelihoods	2011	Rate	HHs	2,162	9.5	20.3	48.0	67.6	8.8	0.1	0.4	1.2	7.4	18.7
		Rate	People		12.4	25.5	54.7	73.7	12.5	0.1	0.5	1.8	10.7	25.2
<u>Validation</u>														
Measuring accuracy	2011	Rate	HHs	2,155	9.5	20.3	47.9	67.4	9.1	0.1	0.5	1.6	7.2	18.7
		Rate	People		12.5	25.2	54.4	73.5	12.9	0.3	0.8	2.7	10.4	24.9

Source: 2005, 2007, and 2011 PECS. Poverty lines in average calendar-year prices.

National lines are per-adult-equivalent. Median and international 2005 PPP lines are per-person.

Figure 1: Old-definition (legacy) poverty lines and poverty rates (for households and people) for all of Palestine, West Bank, and the Gaza Strip for 2005, 2007, and 2011 by poverty line and by sub-sample

Sample	Year	Line or rate	Level	n	Poverty rates (% with consumption < poverty line) and old-definition poverty lines (NIS/day per adult-equivalent or per-person)					
					National poverty lines				Intl. 2005 PPP	
					Deep	100%	150%	200%	\$3.75	\$5.00
<u>All Palestine</u>	2005	Line	People		7.18	10.51	15.77	21.02	8.66	11.55
		Rate	HHs	2,152	6.6	19.0	40.5	57.7	12.0	23.6
			People		9.1	24.7	49.4	66.6	16.0	30.2
<u>All Palestine</u>	2007	Line	People		10.33	12.92	19.39	25.85	9.23	12.31
		Rate	HHs	1,231	19.7	30.0	53.9	68.9	14.8	27.7
			People		25.9	37.4	62.3	76.4	20.2	34.9
<u>All Palestine</u>	2011	Line	People		12.46	15.58	23.37	31.15	11.12	14.83
		Rate	HHs	4,317	12.1	22.3	46.7	62.9	8.5	19.9
			People		16.8	29.5	56.7	72.3	12.2	26.6
West Bank	2011	Line	People		12.46	15.58	23.37	31.15	11.12	17.80
		Rate	HHs	2,909	5.6	12.5	34.5	52.4	3.7	11.2
			People		8.1	17.0	43.3	61.8	5.6	15.3
Gaza Strip	2011	Line	People		12.46	15.58	23.37	31.15	11.12	17.80
		Rate	HHs	1,408	24.4	41.0	69.8	82.8	17.6	36.5
			People		30.9	49.8	78.4	89.2	23.0	44.8
<u>Construction and calibration</u>										
Selecting indicators and points, and associating scores with likelihoods	2011	Rate	HHs	2,162	12.3	22.3	46.7	62.9	8.6	19.8
		Rate	People		17.5	29.8	56.9	72.2	12.4	26.9
<u>Validation</u>										
Measuring accuracy	2011	Rate	HHs	2,155	11.8	22.2	46.7	62.8	8.4	19.9
		Rate	People		16.2	29.3	56.5	72.3	12.1	26.4

Source: 2005, 2007, and 2011 PECS. Poverty lines in average calendar-year prices.

National lines are per-adult-equivalent. International 2005 PPP lines are per-person.

Figure 2: Poverty indicators by uncertainty coefficient

<u>Uncertainty coefficient</u>	<u>Indicator (Responses ordered starting with those most strongly linked with higher poverty likelihoods)</u>
865	Does the household have a vacuum cleaner? (No; Yes)
841	How many household members are there? (Eight or more; Seven; Six; Five; Four; One, two, or three)
646	What is the main source of energy for baking? (Wood; Olive cakes, coal, or other; Electricity; Gas; None or does not exist)
610	What is the occupation of the male head/spouse? (Elementary occupation, or other; Skilled agriculture or fishery worker; Does not work; Service, shop, or market worker; Craft and related trade worker; No male head/spouse; Plant and machine operator and assembler; Legislator, senior official, or manager; Professional, associate professional, or clerk)
608	What is the main source of energy for heating the residence? (None, coal/firewood, solar, or other; Gas, kerosene, or electricity)
558	Does the household have a private car? (No; Yes)
510	How many household members are 18-years-old or younger? (Six or more; Five; Four; Three; Two; One; None)
472	How many household members are self-employed without employees, are unpaid employees, or work for an irregular wage? (Two or more; One; None)
445	Does the household have a home library? (No; Yes)
438	Does the household have a land-line and/or a cellular telephone? (None; Only cellular; Land-line (regardless of cellular))
435	How many household members are 17-years-old or younger? (Six or more; Five; Four; Three; Two; One; None)
433	What is the highest educational level that the female head/spouse has completed? (None, illiterate, or can read and write; Elementary school; Preparatory school; Secondary school; Associate diploma; No female head/spouse; Bachelor's degree, high diploma, master's degree, or doctorate)

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

<u>Uncertainty coefficient</u>	<u>Indicator (Responses ordered starting with those most strongly linked with higher poverty likelihoods)</u>
432	Does the household have a land-line telephone? (No; Yes)
427	Does the household have a computer? (No; Yes)
422	What is the highest educational level that the male head/spouse has completed? (None, illiterate, or can read and write; Elementary school; Preparatory school; Secondary school; No male head/spouse; Associate diploma; Bachelor's degree, high diploma, master's degree, or doctorate)
410	How many household members are 16-years-old or younger? (Six or more; Five; Four; Three; Two; One; None)
409	Does the household have a VCR and/or DVD? (No; Yes)
409	Does the household have a T.V. and/or a VCR or DVD? (None; Only T.V.; VCR or DVD (regardless of T.V.))
399	What is the main building material used in the outside walls of the residence? (Cement cob, mud, or other; Cleaned stone, stone and cement, old stone, or concrete)
390	In their main occupation, do any household members work in elementary occupations? (Yes; No)
376	What is the main source of energy for heating water? (Gas, kerosene, wood, coal, solar, or other; Solar; Electricity)
358	What is the employment status of the male head/spouse? (Paid employee (irregular wage), or unpaid employee; Not employed; Self-employed; No male head/spouse; Paid employee (regular wage); Employer)
330	In their main occupation, how many household members work in elementary occupations or in agriculture, fishing, or forestry? (Two or more; One; None)
326	What is the occupation of the female head/spouse? (Craft and related trade worker, or plant and machine operators and assemblers; Skilled agriculture or fishery worker, elementary occupation, or other; Does not work; Legislator, senior official, manager, professional, associate professional, or clerk; Service, shop, or market worker; No female head/spouse)

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

<u>Uncertainty coefficient</u>	<u>Indicator (Responses ordered starting with those most strongly linked with higher poverty likelihoods)</u>
322	Does the household have a refrigerator? (No; Yes)
322	In their main occupation, do any household members work as a professional, technician, associate professional, clerk, legislator, senior official, or manager? (No; Yes)
319	How many household members ages 6 to 18 are currently attending school? (Not all; All; No members are ages 6 to 18)
314	Does the household have a radio/recorder?
314	Does the household have a filter?
314	How many household members are 15-years-old or younger? (Five or more; Four; Three; Two; One; None)
291	Does the household have a solar boiler?
288	How many household members are 14-years-old or younger? (Five or more; Four; Three; Two; One; None)
267	In his main occupation, in what sector does the male head/spouse work?
257	How many household members are 13-years-old or younger? (Five or more; Four; Three; Two; One; None)
254	How many household members are 12-years-old or younger? (Five or more; Four; Three; Two; One; None)
252	How many household members ages 6 to 17 are currently attending school? (Not all; All; No members are ages 6 to 17)
247	How many household members ages 6 to 16 are currently attending school? (Not all; All; No members are ages 6 to 16)
239	How many household members ages 6 to 15 are currently attending school? (Not all; All; No members are ages 6 to 15)
230	In their main occupation, how many household members work in the sector of agric
227	What is the employment status of the female head/spouse?
212	What type of water network is the residence connected to?
209	How many household members are 11-years-old or younger? (Four or more; Three; Two; One; None)
207	During the past week, was the male head/spouse employed for one hour or more?

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

<u>Uncertainty coefficient</u>	<u>Indicator (Responses ordered starting with those most strongly linked with higher poverty likelihoods)</u>
204	In her main occupation, in what sector does the female head/spouse work?
203	What is the main source of energy for air conditioning?
199	How many household members are employers with employees, or work for a regular w
197	How many household members ages 6 to 13 are currently attending school? (Not all; All; No members are ages 6 to 13)
185	How many household members ages 6 to 14 are currently attending school? (Not all; All; No members are ages 6 to 14)
182	How many household members ages 6 to 12 are currently attending school? (Not all; All; No members are ages 6 to 12)
174	Does the household have a clothes dryer/dehumidifier? (No; Yes)
159	How many rooms does the residence have (excluding kitchens and bathrooms)? (One; Two; Three; Four; Five or more)
153	How many household members ages 6 to 11 are currently attending school? (Not all; All; No members are ages 6 to 11)
137	Does the household have a washing machine? (No; Yes)
122	How many household members are 6-years-old or younger? (Three or more; Two; One; None)
86	Are any household members refugees (be they registered or unregistered)? (No; Yes)
85	In their main occupation, do any household members work in agriculture, fishing, or forestry? (Yes; No)
79	Does the household have any cattle, sheep, goats, poultry, horses, mules, or beehives? (Yes; No)
71	Does the household have a satellite dish? (No; Yes)
71	What type of residence does the household have? (House, or separate room, tent, marginal shelter, or other; Villa, or apartment)
66	Does the household have any agricultural land (be it owned, rented, or free)? (No; Yes)
54	How many bedrooms does the residence have? (One; Two; Three; Four or more)

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

<u>Uncertainty coefficient</u>	<u>Indicator (Responses ordered starting with those most strongly linked with higher poverty likelihoods)</u>
44	Does the household have a dishwasher? (No; Yes)
43	What is the structure of household headship? (Both male and female heads/spouses; Female head/spouse only; Male head/spouse only)
38	What is the marital status of the female head/spouse? (Married; Other (divorced, separated, or widowed); Single (never married or common-law married); No female head/spouse)
29	What is the marital status of the male head/spouse? (Married; No male head/spouse; Single (never married or common-law married), or other (divorced, separated, or widowed))
25	During the past week, was the female head/spouse employed for one hour or more? (No; Yes; No female head/spouse)
23	Do any household members work at more than one job? (No; Yes)
18	What type of sewage system is the residence connected to? (Hole absorption; No sewerage system; Cesspit; Public)
14	During the past week, were any household members employed for one hour or more? (No; Yes)
11	Does the household have a T.V.? (No; Yes)
8	Does the household have a cellular telephone? (No; Yes)
8	What is the main source of income for the household? (Wages and salaries from the private sector; International aid organizations; Social aid/affairs)
3	What the the tenancy status of the household in its residence? (Other (provided at subsidized rent, without payment, in exchange for work, etc.); Rented (with or without furniture); Owned)

Source: 2011 PECS and 100% of the new-definition national poverty line

**Tables for
the New-Definition Deep Poverty Line

(and Tables Pertaining
to All Ten New-Definition Poverty Lines
and All Six Legacy-Definition Poverty Lines)**

Figure 3 (New-definition deep 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	72.5
15-19	54.1
20-24	46.8
25-29	32.3
30-34	27.8
35-39	17.5
40-44	11.3
45-49	9.5
50-54	2.0
55-59	1.9
60-64	0.5
65-69	0.1
70-74	0.1
75-79	0.0
80-84	0.0
85-89	0.0
90-94	0.0
95-100	0.0

Figure 4 (New-definition deep line): Derivation of estimated poverty likelihoods associated with scores

Score	Households at score and < poverty line		All households at score		Poverty likelihood (%)
0-4	61	÷	61	=	100.0
5-9	33	÷	33	=	100.0
10-14	242	÷	334	=	72.5
15-19	577	÷	1,065	=	54.1
20-24	1,170	÷	2,501	=	46.8
25-29	1,165	÷	3,611	=	32.3
30-34	1,788	÷	6,426	=	27.8
35-39	1,564	÷	8,917	=	17.5
40-44	1,412	÷	12,490	=	11.3
45-49	1,152	÷	12,066	=	9.5
50-54	252	÷	12,415	=	2.0
55-59	193	÷	10,344	=	1.9
60-64	34	÷	7,134	=	0.5
65-69	11	÷	7,917	=	0.1
70-74	4	÷	7,273	=	0.1
75-79	0	÷	4,730	=	0.0
80-84	0	÷	1,495	=	0.0
85-89	0	÷	934	=	0.0
90-94	0	÷	187	=	0.0
95-100	0	÷	67	=	0.0

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

Figure 5 (Per-adult-equivalent new-definition poverty lines): Probability that a given household's consumption falls in a range demarcated by two adjacent new-definition per-adult-equivalent poverty lines

Score	Likelihood (%) of having daily per-adult-equivalent consumption in a range demarcated by adjacent new-definition poverty lines				
		\geq Deep	\geq 100% Natl.	\geq 150% Natl.	
	<Deep	and	and	and	\geq 200% Natl.
		<100% Natl.	<150% Natl.	<200% Natl.	
		\geq NIS19.85	\geq NIS24.84	\geq NIS37.26	
	<NIS19.85	and	and	and	\geq NIS49.69
		<NIS24.84	<NIS37.26	<NIS49.69	
0–4	100.0	0.0	0.0	0.0	0.0
5–9	100.0	0.0	0.0	0.0	0.0
10–14	72.5	21.7	2.9	2.0	0.9
15–19	54.1	29.7	10.7	3.8	1.6
20–24	46.8	27.3	20.5	3.9	1.6
25–29	32.3	24.1	32.1	9.9	1.6
30–34	27.8	20.2	34.2	14.2	3.5
35–39	17.5	22.6	34.6	16.6	8.7
40–44	11.3	20.7	38.2	15.8	14.0
45–49	9.5	14.1	36.4	24.1	15.8
50–54	2.0	6.4	34.3	28.5	28.9
55–59	1.9	4.5	29.2	29.4	35.0
60–64	0.5	2.5	21.6	22.4	53.0
65–69	0.1	1.0	20.9	21.6	56.4
70–74	0.1	1.1	13.1	20.0	65.8
75–79	0.0	0.7	7.1	9.4	82.8
80–84	0.0	0.0	4.2	4.5	91.3
85–89	0.0	0.0	2.7	5.4	91.9
90–94	0.0	0.0	0.0	6.0	94.0
95–100	0.0	0.0	0.0	0.0	100.0

Figure 5 (Per-capita new-definition poverty lines): Probability that a given household's consumption falls in a range demarcated by two adjacent new-definition per-capita poverty lines

Score	Likelihood (%) of having daily per-capita consumption in a range demarcated by adjacent new-definition poverty lines						
		$\geq \$1.25/\text{day}$	$\geq \$2.00/\text{day}$	$\geq \$2.50/\text{day}$	$\geq \$3.75/\text{day}$	$\geq \text{Median}$	$\geq \$5.00/\text{day}$
	$< \$1.25/\text{day}$	and	and	and	and	and	
		$< \$2.00/\text{day}$	$< \$2.50/\text{day}$	$< \$3.75/\text{day}$	$< \text{Median}$	$< \$5.00/\text{day}$	
	$\geq \text{NIS}3.68$	$\geq \text{NIS}5.89$	$\geq \text{NIS}7.36$	$\geq \text{NIS}11.04$	$\geq \text{NIS}11.87$		
	$< \text{NIS}3.68$	and	and	and	and	and	$\geq \text{NIS}14.72$
		$< \text{NIS}5.89$	$< \text{NIS}7.36$	$< \text{NIS}11.04$	$< \text{NIS}11.87$	$< \text{NIS}14.72$	
0-4	2.0	24.5	0.0	73.5	0.0	0.0	0.0
5-9	2.0	24.5	0.0	46.9	0.0	26.5	0.0
10-14	2.0	7.0	4.6	39.4	6.2	34.3	6.5
15-19	2.0	1.1	6.2	36.2	7.8	32.8	13.9
20-24	1.4	1.6	6.4	35.7	7.2	19.4	28.3
25-29	0.0	2.0	3.4	25.8	2.3	25.9	40.7
30-34	0.0	1.1	2.6	21.2	2.5	23.1	49.6
35-39	0.0	0.5	2.2	13.5	3.2	17.4	63.2
40-44	0.0	0.1	1.0	6.2	3.1	15.2	74.5
45-49	0.0	0.0	0.0	3.0	1.3	14.5	81.2
50-54	0.0	0.0	0.0	0.8	0.5	5.9	92.8
55-59	0.0	0.0	0.0	0.8	0.5	4.6	94.1
60-64	0.0	0.0	0.0	0.3	0.1	1.6	98.0
65-69	0.0	0.0	0.0	0.1	0.0	0.8	99.0
70-74	0.0	0.0	0.0	0.1	0.0	0.7	99.2
75-79	0.0	0.0	0.0	0.0	0.0	0.0	100.0
80-84	0.0	0.0	0.0	0.0	0.0	0.0	100.0
85-89	0.0	0.0	0.0	0.0	0.0	0.0	100.0
90-94	0.0	0.0	0.0	0.0	0.0	0.0	100.0
95-100	0.0	0.0	0.0	0.0	0.0	0.0	100.0

Figure 5 (Per-adult-equivalent old-definition poverty lines): Probability that a given household's consumption falls in a range demarcated by two adjacent old-definition per-adult-equivalent poverty lines

Score	Likelihood (%) of having daily per-adult-equivalent consumption in a range demarcated by adjacent old-definition poverty lines				
		\geq Deep	\geq 100% Natl.	\geq 150% Natl.	
	<Deep	and	and	and	\geq 200% Natl.
		<100% Natl.	<150% Natl.	<200% Natl.	
		\geq NIS12.46	\geq NIS15.58	\geq NIS23.37	
	<NIS12.46	and	and	and	\geq NIS31.15
		<NIS15.58	<NIS23.37	<NIS31.15	
0-4	100.0	0.0	0.0	0.0	0.0
5-9	100.0	0.0	0.0	0.0	0.0
10-14	92.2	1.7	4.4	0.2	1.5
15-19	70.6	15.5	10.8	0.4	2.8
20-24	58.0	18.0	20.2	1.0	2.8
25-29	42.5	23.2	23.2	6.9	4.1
30-34	37.7	19.6	25.0	13.0	4.8
35-39	24.3	21.5	30.1	15.4	8.7
40-44	15.7	18.5	35.0	14.5	16.3
45-49	10.7	12.2	38.4	16.8	21.9
50-54	2.2	7.0	30.9	23.3	36.7
55-59	2.2	6.1	24.7	24.1	42.9
60-64	0.5	2.2	16.4	21.8	59.1
65-69	0.1	1.0	15.6	20.8	62.5
70-74	0.1	0.9	13.7	12.5	72.9
75-79	0.0	0.0	8.0	5.4	86.6
80-84	0.0	0.0	2.4	3.7	93.9
85-89	0.0	0.0	1.5	4.5	93.9
90-94	0.0	0.0	0.0	5.1	94.9
95-100	0.0	0.0	0.0	0.0	100.0

Figure 5 (Per-capita old-definition poverty lines):
Probability that a given household's consumption
falls in a range demarcated by two adjacent old-
definition per-capita poverty lines

Score	Likelihood (%) of having daily per-capita consumption in a range demarcated by adjacent old-definition poverty lines		
	<\$3.75/day	≥\$3.75/day	≥\$5.00/day
		and	
	<NIS11.12	<\$5.00/day	≥NIS14.83
and			
	<NIS11.12	<NIS14.83	
0–4	100.0	0.0	0.0
5–9	100.0	0.0	0.0
10–14	77.5	16.4	6.0
15–19	53.7	32.3	13.9
20–24	50.2	24.8	25.0
25–29	34.5	26.7	38.8
30–34	24.7	29.0	46.3
35–39	18.2	22.6	59.2
40–44	9.2	18.2	72.6
45–49	6.1	14.4	79.5
50–54	0.8	6.3	92.9
55–59	0.7	5.0	94.3
60–64	0.2	1.9	97.9
65–69	0.1	1.0	98.9
70–74	0.1	0.9	99.1
75–79	0.0	0.0	100.0
80–84	0.0	0.0	100.0
85–89	0.0	0.0	100.0
90–94	0.0	0.0	100.0
95–100	0.0	0.0	100.0

Figure 6 (New-definition deep line): Average differences between estimated and true poverty likelihoods for households in 1,000 bootstraps of $n = 16,384$, with confidence intervals by score range, scorecard applied to the validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0–4	+0.0	0.0	0.0	0.0
5–9	+0.0	0.0	0.0	0.0
10–14	+32.6	11.5	13.4	18.5
15–19	+21.5	6.5	7.5	9.3
20–24	–24.6	14.2	14.5	15.3
25–29	–15.4	9.6	9.9	10.7
30–34	–1.9	2.6	3.0	3.8
35–39	+2.6	1.6	2.0	2.7
40–44	+4.6	0.9	1.0	1.4
45–49	+3.7	0.9	1.0	1.4
50–54	–2.0	1.4	1.5	1.7
55–59	+0.1	0.5	0.6	0.8
60–64	–0.7	0.6	0.7	0.7
65–69	+0.1	0.0	0.0	0.0
70–74	–0.5	0.5	0.5	0.6
75–79	+0.0	0.0	0.0	0.0
80–84	–7.0	5.0	5.2	5.8
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

Figure 7 (New-definition deep line): Average differences between estimated poverty rates and true values for a group at a point in time in 1,000 bootstraps of various sample sizes, with confidence intervals, scorecard applied to the validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+0.4	50.0	60.5	76.0
4	-0.8	23.6	30.0	41.5
8	-0.3	16.3	19.7	24.5
16	-0.1	11.1	13.2	17.1
32	-0.1	7.9	9.7	12.6
64	-0.1	5.8	7.1	8.7
128	-0.1	4.2	4.9	6.0
256	-0.1	2.9	3.5	4.4
512	-0.2	2.0	2.4	3.0
1,024	-0.2	1.4	1.7	2.1
2,048	-0.2	1.0	1.3	1.6
4,096	-0.2	0.7	0.9	1.2
8,192	-0.2	0.5	0.6	0.8
16,384	-0.2	0.4	0.4	0.6

Figure 8 (All new-definition poverty lines): Average differences between estimates and true values for poverty rates of a group of households at a point in time in 1,000 bootstraps of $n = 16,384$, precision, and the α factor for precision, scorecard applied to the validation sample

	Poverty line									
	National poverty lines				Median	Intl. 2005 PPP				
	Deep	100%	150%	200%		\$1.25	\$2.00	\$2.50	\$3.75	\$5.00
Estimate minus true value	-0.2	+1.3	+3.1	+3.0	+0.0	+0.0	-0.1	-0.4	+0.1	+0.5
Precision of difference	0.4	0.5	0.6	0.6	0.3	0.0	0.1	0.2	0.3	0.5
α factor for precision	0.98	0.93	0.98	1.01	0.93	0.83	1.13	1.19	0.96	0.93

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

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

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

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

Figure 8 (All old-definition poverty lines): Average differences between estimates and true values for poverty rates of a group of households at a point in time in 1,000 bootstraps of $n = 16,384$, precision, and the α factor for precision, scorecard applied to the validation sample

	Poverty line					
	Deep	National poverty lines			Intl. 2005 PPP	
		100%	150%	200%	\$3.75	\$5.00
Estimate minus true value	+0.6	+1.0	+2.9	+2.5	+0.1	+0.5
Precision of difference	0.4	0.5	0.6	0.6	0.3	0.5
α factor for precision	0.92	0.92	0.96	0.96	0.95	0.93

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

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

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

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

Figure 9 (All poverty lines): Possible targeting outcomes

		<u>Targeting segment</u>	
		<u>Targeted</u>	<u>Non-targeted</u>
<u>True poverty status</u>	<u>Below poverty line</u>	<u>Inclusion</u> Below poverty line Correctly Targeted	<u>Undercoverage</u> Below poverty line Mistakenly Non-targeted
	<u>Above poverty line</u>	<u>Leakage</u> Above poverty line Mistakenly Targeted	<u>Exclusion</u> Above poverty line Correctly Non-targeted

Figure 10 (New-definition deep line): Shares of households by cut-off score and targeting classification, along with the hit rate and BPAC, scorecard applied to the validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	9.5	0.0	90.5	90.5	-98.7
≤9	0.1	9.4	0.0	90.5	90.6	-98.0
≤14	0.2	9.3	0.2	90.3	90.5	-93.0
≤19	0.6	9.0	0.9	89.6	90.1	-78.3
≤24	2.1	7.4	1.9	88.6	90.7	-36.2
≤29	3.6	5.9	4.0	86.5	90.2	+18.0
≤34	5.4	4.1	8.6	81.9	87.3	+9.8
≤39	6.7	2.8	16.2	74.2	80.9	-70.4
≤44	7.7	1.8	27.7	62.7	70.4	-191.0
≤49	8.5	1.0	39.0	51.5	60.0	-308.9
≤54	9.1	0.5	50.8	39.6	48.7	-433.5
≤59	9.3	0.2	61.0	29.5	38.8	-539.9
≤64	9.4	0.1	68.0	22.5	31.9	-613.6
≤69	9.4	0.1	75.9	14.5	23.9	-696.7
≤74	9.4	0.1	83.1	7.3	16.8	-772.4
≤79	9.4	0.1	87.9	2.6	12.0	-822.1
≤84	9.5	0.0	89.3	1.2	10.7	-836.9
≤89	9.5	0.0	90.2	0.3	9.8	-846.7
≤94	9.5	0.0	90.4	0.1	9.6	-848.6
≤100	9.5	0.0	90.5	0.0	9.5	-849.3

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

Figure 11 (New-definition deep line): By score cut-off, the share of all households who are targeted (that is, score at or below the 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), 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	0.1	100.0	0.6	Only poor targeted
≤9	0.1	100.0	1.0	Only poor targeted
≤14	0.4	55.0	2.5	1.2:1
≤19	1.5	38.8	6.1	0.6:1
≤24	4.0	52.3	21.9	1.1:1
≤29	7.6	47.9	38.2	0.9:1
≤34	14.0	38.7	57.0	0.6:1
≤39	22.9	29.2	70.4	0.4:1
≤44	35.4	21.7	80.9	0.3:1
≤49	47.5	18.0	89.5	0.2:1
≤54	59.9	15.2	95.3	0.2:1
≤59	70.3	13.2	97.4	0.2:1
≤64	77.4	12.1	98.5	0.1:1
≤69	85.3	11.0	98.5	0.1:1
≤74	92.6	10.2	99.1	0.1:1
≤79	97.3	9.7	99.1	0.1:1
≤84	98.8	9.6	100.0	0.1:1
≤89	99.7	9.6	100.0	0.1:1
≤94	99.9	9.5	100.0	0.1:1
≤100	100.0	9.5	100.0	0.1:1

**Tables for
100% of the New-Definition National Poverty Line**

**Figure 3 (100% of the new-definition national line):
 Estimated poverty likelihoods associated with scores**

If a household's score is then the likelihood (%) of being below the poverty line is:
0-4	100.0
5-9	100.0
10-14	94.2
15-19	83.9
20-24	74.1
25-29	56.4
30-34	48.1
35-39	40.1
40-44	32.0
45-49	23.6
50-54	8.4
55-59	6.4
60-64	3.0
65-69	1.1
70-74	1.1
75-79	0.7
80-84	0.0
85-89	0.0
90-94	0.0
95-100	0.0

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

Score	Households at score and < poverty line		All households at score		Poverty likelihood (%)
0-4	61	÷	61	=	100.0
5-9	33	÷	33	=	100.0
10-14	314	÷	334	=	94.2
15-19	893	÷	1,065	=	83.9
20-24	1,853	÷	2,501	=	74.1
25-29	2,037	÷	3,611	=	56.4
30-34	3,088	÷	6,426	=	48.1
35-39	3,577	÷	8,917	=	40.1
40-44	3,993	÷	12,490	=	32.0
45-49	2,849	÷	12,066	=	23.6
50-54	1,040	÷	12,415	=	8.4
55-59	660	÷	10,344	=	6.4
60-64	215	÷	7,134	=	3.0
65-69	89	÷	7,917	=	1.1
70-74	82	÷	7,273	=	1.1
75-79	35	÷	4,730	=	0.7
80-84	0	÷	1,495	=	0.0
85-89	0	÷	934	=	0.0
90-94	0	÷	187	=	0.0
95-100	0	÷	67	=	0.0

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

**Figure 6 (100% of the new-definition national line):
Average differences between estimated and true
poverty likelihoods for households in 1,000
bootstraps of $n = 16,384$, with confidence intervals
by score range, scorecard applied to the validation
sample**

Score	Diff.	Difference between estimate and true value		
		Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	+0.0	0.0	0.0	0.0
5-9	+0.0	0.0	0.0	0.0
10-14	+8.4	6.7	8.1	10.9
15-19	+30.4	6.5	7.8	10.3
20-24	-11.5	7.0	7.2	7.7
25-29	-6.3	5.0	5.2	6.0
30-34	-5.2	4.0	4.2	4.5
35-39	+4.3	2.3	2.7	3.4
40-44	+10.8	1.6	1.9	2.4
45-49	+8.7	1.4	1.7	2.3
50-54	-3.4	2.3	2.5	2.6
55-59	-1.0	1.2	1.4	1.8
60-64	-1.4	1.2	1.3	1.5
65-69	-0.7	0.8	0.9	1.1
70-74	-1.2	0.9	1.0	1.1
75-79	-1.1	0.9	1.0	1.3
80-84	-7.0	5.0	5.2	5.8
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

**Figure 7 (100% of the new-definition national line):
Average differences between estimated poverty rates
and true values for a group at a point in time in
1,000 bootstraps of various sample sizes, with
confidence intervals, scorecard applied to the
validation sample**

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	-0.8	62.2	74.0	88.9
4	+0.2	32.6	39.1	54.3
8	+1.1	22.3	25.9	34.5
16	+1.5	15.2	18.8	25.3
32	+1.5	10.9	13.0	18.2
64	+1.4	7.6	9.1	11.8
128	+1.4	5.3	6.5	8.0
256	+1.4	3.8	4.5	6.0
512	+1.3	2.6	3.1	3.8
1,024	+1.3	1.9	2.2	3.0
2,048	+1.3	1.4	1.6	2.1
4,096	+1.3	1.0	1.1	1.6
8,192	+1.3	0.7	0.8	1.0
16,384	+1.3	0.5	0.5	0.7

Figure 10 (100% of the new-definition national line): Shares of households by cut-off score and targeting classification, along with the hit rate and BPAC, scorecard applied to the validation sample

Score	Inclusion:	Undercoverage:	Leakage:	Exclusion:	Hit rate	BPAC
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	20.2	0.0	79.7	79.8	-99.4
≤9	0.1	20.2	0.0	79.7	79.8	-99.1
≤14	0.4	19.9	0.1	79.6	80.0	-96.1
≤19	1.0	19.3	0.5	79.2	80.2	-87.9
≤24	2.9	17.4	1.1	78.7	81.6	-65.8
≤29	5.2	15.1	2.4	77.3	82.5	-37.0
≤34	8.6	11.7	5.5	74.3	82.8	+11.4
≤39	11.8	8.5	11.2	68.5	80.3	+44.9
≤44	14.7	5.6	20.7	59.0	73.7	-2.1
≤49	16.9	3.4	30.6	49.1	66.0	-50.9
≤54	18.6	1.7	41.3	38.4	56.9	-103.7
≤59	19.4	0.9	50.9	28.8	48.2	-150.7
≤64	19.8	0.5	57.6	22.1	41.8	-184.0
≤69	19.9	0.4	65.4	14.3	34.2	-222.4
≤74	20.1	0.2	72.5	7.3	27.4	-257.0
≤79	20.2	0.1	77.1	2.6	22.8	-279.9
≤84	20.3	0.0	78.5	1.2	21.5	-286.9
≤89	20.3	0.0	79.4	0.3	20.6	-291.5
≤94	20.3	0.0	79.6	0.1	20.4	-292.4
≤100	20.3	0.0	79.7	0.0	20.3	-292.7

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

Figure 11 (100% of the new-definition national line): By score cut-off, the share of all households who are targeted (that is, score at or below the 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), 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	0.1	100.0	0.3	Only poor targeted
≤9	0.1	100.0	0.5	Only poor targeted
≤14	0.4	83.4	1.8	5.0:1
≤19	1.5	65.1	4.8	1.9:1
≤24	4.0	73.6	14.5	2.8:1
≤29	7.6	68.3	25.6	2.2:1
≤34	14.0	61.1	42.3	1.6:1
≤39	22.9	51.3	58.0	1.1:1
≤44	35.4	41.5	72.5	0.7:1
≤49	47.5	35.5	83.2	0.6:1
≤54	59.9	31.0	91.5	0.4:1
≤59	70.3	27.6	95.5	0.4:1
≤64	77.4	25.5	97.4	0.3:1
≤69	85.3	23.3	98.0	0.3:1
≤74	92.6	21.7	99.2	0.3:1
≤79	97.3	20.8	99.6	0.3:1
≤84	98.8	20.5	100.0	0.3:1
≤89	99.7	20.3	100.0	0.3:1
≤94	99.9	20.3	100.0	0.3:1
≤100	100.0	20.3	100.0	0.3:1

**Tables for
150% of the New-Definition National Poverty Line**

**Figure 3 (150% of the new-definition national line):
 Estimated poverty likelihoods associated with scores**

If a household's score is then the likelihood (%) of being below the poverty line is:
0-4	100.0
5-9	100.0
10-14	97.1
15-19	94.5
20-24	94.5
25-29	88.5
30-34	82.3
35-39	74.7
40-44	70.1
45-49	60.0
50-54	42.7
55-59	35.6
60-64	24.6
65-69	22.0
70-74	14.2
75-79	7.9
80-84	4.2
85-89	2.7
90-94	0.0
95-100	0.0

**Figure 6 (150% of the new-definition national line):
Average differences between estimated and true
poverty likelihoods for households in 1,000
bootstraps of $n = 16,384$, with confidence intervals
by score range, scorecard applied to the validation
sample**

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	+0.0	0.0	0.0	0.0
5-9	+0.0	0.0	0.0	0.0
10-14	-2.9	1.4	1.4	1.4
15-19	+8.6	4.6	5.5	6.9
20-24	-1.0	1.3	1.6	2.2
25-29	-0.3	2.3	2.6	3.5
30-34	+4.7	2.5	2.9	3.9
35-39	-0.8	2.0	2.4	3.1
40-44	+11.2	2.0	2.3	3.1
45-49	+12.8	2.0	2.4	3.4
50-54	-2.6	2.3	2.5	3.1
55-59	+1.0	2.0	2.5	3.3
60-64	-8.8	5.7	6.0	6.4
65-69	+5.5	1.8	2.1	2.9
70-74	+1.5	1.6	1.8	2.4
75-79	+5.2	0.9	1.1	1.4
80-84	-4.7	3.9	4.2	4.9
85-89	-4.5	4.0	4.5	5.2
90-94	+0.0	0.0	0.0	0.0
95-100	+0.0	0.0	0.0	0.0

**Figure 7 (150% of the new-definition national line):
Average differences between estimated poverty rates
and true values for a group at a point in time in
1,000 bootstraps of various sample sizes, with
confidence intervals, scorecard applied to the
validation sample**

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	-0.9	72.8	76.4	87.1
4	+1.9	37.4	44.9	56.1
8	+2.8	27.9	32.8	40.9
16	+2.9	19.7	23.9	29.0
32	+3.2	14.0	16.1	21.1
64	+3.0	10.1	12.0	15.5
128	+3.0	6.8	8.2	10.3
256	+2.9	4.8	5.7	7.6
512	+2.9	3.4	4.3	5.8
1,024	+3.0	2.5	3.0	4.0
2,048	+3.1	1.8	2.1	2.9
4,096	+3.1	1.3	1.5	2.0
8,192	+3.1	0.9	1.1	1.3
16,384	+3.1	0.6	0.8	1.0

Figure 10 (150% of the new-definition national line): Shares of households by cut-off score and targeting classification, along with the hit rate and BPAC, scorecard applied to the validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	47.8	0.0	52.1	52.2	-99.7
≤9	0.1	47.8	0.0	52.1	52.2	-99.6
≤14	0.4	47.5	0.0	52.1	52.5	-98.2
≤19	1.4	46.5	0.1	52.0	53.3	-94.1
≤24	3.7	44.2	0.3	51.8	55.4	-84.0
≤29	6.8	41.1	0.8	51.3	58.1	-69.9
≤34	12.0	35.9	2.0	50.1	62.1	-45.6
≤39	18.8	29.1	4.1	48.0	66.8	-12.8
≤44	26.7	21.3	8.8	43.3	70.0	+29.6
≤49	33.0	14.9	14.5	37.6	70.7	+68.1
≤54	38.9	9.0	21.0	31.1	69.9	+56.1
≤59	42.7	5.2	27.6	24.5	67.2	+42.4
≤64	45.0	2.9	32.4	19.7	64.7	+32.3
≤69	46.4	1.5	38.9	13.2	59.6	+18.7
≤74	47.6	0.3	45.0	7.1	54.6	+6.0
≤79	47.7	0.2	49.6	2.5	50.2	-3.5
≤84	47.8	0.1	51.0	1.1	49.0	-6.4
≤89	47.9	0.0	51.8	0.3	48.2	-8.2
≤94	47.9	0.0	52.0	0.1	48.0	-8.6
≤100	47.9	0.0	52.1	0.0	47.9	-8.7

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

Figure 11 (150% of the new-definition national line): By score cut-off, the share of all households who are targeted (that is, score at or below the 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), 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	0.1	100.0	0.1	Only poor targeted
≤9	0.1	100.0	0.2	Only poor targeted
≤14	0.4	100.0	0.9	Only poor targeted
≤19	1.5	90.9	2.8	9.9:1
≤24	4.0	91.7	7.6	11.1:1
≤29	7.6	89.6	14.2	8.6:1
≤34	14.0	85.8	25.1	6.1:1
≤39	22.9	81.9	39.2	4.5:1
≤44	35.4	75.2	55.6	3.0:1
≤49	47.5	69.6	69.0	2.3:1
≤54	59.9	64.9	81.2	1.8:1
≤59	70.3	60.8	89.1	1.5:1
≤64	77.4	58.1	93.9	1.4:1
≤69	85.3	54.4	96.8	1.2:1
≤74	92.6	51.4	99.3	1.1:1
≤79	97.3	49.0	99.6	1.0:1
≤84	98.8	48.4	99.9	0.9:1
≤89	99.7	48.0	100.0	0.9:1
≤94	99.9	47.9	100.0	0.9:1
≤100	100.0	47.9	100.0	0.9:1

**Tables for
200% of the New-Definition National Poverty Line**

**Figure 3 (200% of the new-definition national line):
 Estimated poverty likelihoods associated with scores**

If a household's score is then the likelihood (%) of being below the poverty line is:
0-4	100.0
5-9	100.0
10-14	99.1
15-19	98.4
20-24	98.4
25-29	98.4
30-34	96.5
35-39	91.3
40-44	86.0
45-49	84.2
50-54	71.1
55-59	65.0
60-64	47.0
65-69	43.6
70-74	34.2
75-79	17.2
80-84	8.7
85-89	8.1
90-94	6.0
95-100	0.0

Figure 6 (200% of the new-definition national line):
Average differences between estimated and true
poverty likelihoods for households in 1,000
bootstraps of $n = 16,384$, with confidence intervals
by score range, scorecard applied to the validation
sample

Score	Diff.	Difference between estimate and true value		
		Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	+0.0	0.0	0.0	0.0
5-9	+0.0	0.0	0.0	0.0
10-14	-0.9	0.4	0.4	0.4
15-19	-1.6	0.8	0.8	0.8
20-24	+2.9	1.3	1.6	2.2
25-29	+0.2	0.8	1.0	1.2
30-34	-1.3	1.0	1.1	1.1
35-39	+3.4	1.6	2.0	2.4
40-44	+5.9	1.8	2.1	3.0
45-49	+15.5	1.9	2.4	3.2
50-54	-3.8	2.7	2.8	3.1
55-59	+5.0	2.3	2.7	3.5
60-64	-9.3	6.0	6.2	6.6
65-69	+6.2	2.4	2.8	3.8
70-74	+1.7	2.5	2.8	3.8
75-79	+5.3	2.0	2.3	2.9
80-84	-0.3	3.1	3.7	5.0
85-89	-13.3	9.4	9.9	10.8
90-94	+6.0	0.0	0.0	0.0
95-100	+0.0	0.0	0.0	0.0

Figure 7 (200% of the new-definition national line):
Average differences between estimated poverty rates and true values for a group at a point in time in 1,000 bootstraps of various sample sizes, with confidence intervals, scorecard applied to the validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+0.5	70.3	75.9	89.7
4	+1.5	35.6	41.8	52.9
8	+2.5	26.5	30.0	39.5
16	+2.6	18.5	21.9	29.2
32	+2.8	13.1	15.3	20.0
64	+3.0	9.5	11.6	14.7
128	+3.1	6.8	7.9	11.0
256	+3.0	4.7	5.5	7.5
512	+3.0	3.4	4.1	5.4
1,024	+3.0	2.4	2.9	3.8
2,048	+3.1	1.7	2.1	2.7
4,096	+3.1	1.2	1.4	2.0
8,192	+3.0	0.9	1.0	1.4
16,384	+3.0	0.6	0.7	1.0

Figure 10 (200% of the new-definition national line): Shares of households by cut-off score and targeting classification, along with the hit rate and BPAC, scorecard applied to the validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	67.4	0.0	32.6	32.6	-99.8
≤9	0.1	67.4	0.0	32.6	32.6	-99.7
≤14	0.4	67.0	0.0	32.6	33.0	-98.7
≤19	1.5	66.0	0.0	32.6	34.0	-95.6
≤24	3.8	63.6	0.2	32.4	36.2	-88.4
≤29	7.3	60.1	0.3	32.3	39.6	-77.9
≤34	13.6	53.9	0.5	32.1	45.6	-59.1
≤39	21.5	45.9	1.4	31.1	52.6	-34.1
≤44	32.0	35.5	3.5	29.1	61.1	-0.0
≤49	40.8	26.6	6.7	25.9	66.7	+30.9
≤54	50.1	17.4	9.8	22.7	72.8	+63.1
≤59	56.7	10.8	13.6	19.0	75.7	+79.9
≤64	60.6	6.8	16.8	15.8	76.4	+75.1
≤69	63.9	3.5	21.4	11.1	75.0	+68.3
≤74	66.4	1.0	26.2	6.4	72.8	+61.2
≤79	67.1	0.3	30.2	2.3	69.4	+55.2
≤84	67.2	0.2	31.6	1.0	68.2	+53.2
≤89	67.4	0.0	32.3	0.3	67.7	+52.1
≤94	67.4	0.0	32.5	0.1	67.5	+51.8
≤100	67.4	0.0	32.6	0.0	67.4	+51.7

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

Figure 11 (200% of the new-definition national line): By score cut-off, the share of all households who are targeted (that is, score at or below the 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), 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	0.1	100.0	0.1	Only poor targeted
≤9	0.1	100.0	0.1	Only poor targeted
≤14	0.4	100.0	0.6	Only poor targeted
≤19	1.5	100.0	2.2	5,028.1:1
≤24	4.0	95.1	5.6	19.5:1
≤29	7.6	96.1	10.8	24.9:1
≤34	14.0	96.7	20.1	29.0:1
≤39	22.9	93.7	31.9	14.9:1
≤44	35.4	90.3	47.4	9.3:1
≤49	47.5	85.9	60.5	6.1:1
≤54	59.9	83.6	74.2	5.1:1
≤59	70.3	80.7	84.1	4.2:1
≤64	77.4	78.3	89.9	3.6:1
≤69	85.3	74.9	94.7	3.0:1
≤74	92.6	71.7	98.5	2.5:1
≤79	97.3	69.0	99.5	2.2:1
≤84	98.8	68.0	99.7	2.1:1
≤89	99.7	67.6	100.0	2.1:1
≤94	99.9	67.5	100.0	2.1:1
≤100	100.0	67.4	100.0	2.1:1

**Tables for
the New-Definition Median Poverty Line**

Figure 3 (New-definition median 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	73.5
10-14	59.3
15-19	53.3
20-24	52.3
25-29	33.5
30-34	27.4
35-39	19.4
40-44	10.3
45-49	4.3
50-54	1.3
55-59	1.3
60-64	0.4
65-69	0.1
70-74	0.1
75-79	0.0
80-84	0.0
85-89	0.0
90-94	0.0
95-100	0.0

Figure 6 (New-definition median line): Average differences between estimated and true poverty likelihoods for households in 1,000 bootstraps of $n = 16,384$, with confidence intervals by score range, scorecard applied to the validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	+0.0	0.0	0.0	0.0
5-9	-26.5	13.3	13.3	13.3
10-14	-26.6	16.5	17.0	17.9
15-19	+11.9	6.5	7.7	10.1
20-24	-17.5	10.7	10.9	11.8
25-29	-6.0	4.8	5.2	6.1
30-34	-4.2	3.4	3.6	4.1
35-39	+4.1	1.7	2.1	2.7
40-44	+7.3	0.6	0.7	1.0
45-49	+0.5	0.7	0.8	1.1
50-54	-3.2	2.0	2.1	2.3
55-59	+0.1	0.4	0.5	0.7
60-64	-0.8	0.7	0.7	0.8
65-69	+0.1	0.0	0.0	0.1
70-74	+0.1	0.0	0.0	0.0
75-79	+0.0	0.0	0.0	0.0
80-84	+0.0	0.0	0.0	0.0
85-89	+0.0	0.0	0.0	0.0
90-94	+0.0	0.0	0.0	0.0
95-100	+0.0	0.0	0.0	0.0

Figure 7 (New-definition median line): Average differences between estimated poverty rates and true values for a group at a point in time in 1,000 bootstraps of various sample sizes, with confidence intervals, scorecard applied to the validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+0.2	50.0	57.0	76.0
4	-0.3	21.9	26.8	39.4
8	+0.0	15.3	19.1	26.1
16	+0.3	10.6	12.5	16.7
32	+0.2	7.7	8.9	11.1
64	+0.1	5.4	6.6	8.3
128	+0.1	3.7	4.2	5.5
256	+0.1	2.7	3.1	3.9
512	+0.1	1.9	2.2	2.7
1,024	+0.0	1.3	1.6	2.0
2,048	+0.0	1.0	1.1	1.5
4,096	+0.0	0.7	0.8	1.1
8,192	+0.0	0.5	0.6	0.8
16,384	+0.0	0.3	0.4	0.5

Figure 10 (New-definition median line): Shares of households by cut-off score and targeting classification, along with the hit rate and BPAC, scorecard applied to the validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	9.0	0.0	90.9	91.0	-98.7
≤9	0.1	9.0	0.0	90.9	91.0	-97.9
≤14	0.4	8.7	0.1	90.8	91.2	-91.3
≤19	0.9	8.2	0.6	90.3	91.1	-74.1
≤24	2.4	6.7	1.6	89.3	91.7	-29.5
≤29	3.7	5.3	3.9	87.0	90.7	+25.1
≤34	5.7	3.4	8.3	82.6	88.2	+7.8
≤39	7.0	2.0	15.9	75.0	82.1	-75.3
≤44	7.6	1.5	27.8	63.1	70.6	-207.5
≤49	8.2	0.9	39.3	51.6	59.8	-333.9
≤54	8.8	0.3	51.1	39.8	48.6	-464.6
≤59	8.9	0.1	61.3	29.6	38.5	-577.1
≤64	9.0	0.0	68.3	22.6	31.6	-654.6
≤69	9.1	0.0	76.2	14.7	23.7	-741.9
≤74	9.1	0.0	83.5	7.4	16.5	-822.2
≤79	9.1	0.0	88.2	2.7	11.7	-874.5
≤84	9.1	0.0	89.7	1.2	10.2	-891.0
≤89	9.1	0.0	90.7	0.3	9.3	-901.3
≤94	9.1	0.0	90.8	0.1	9.1	-903.4
≤100	9.1	0.0	90.9	0.0	9.1	-904.1

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

Figure 11 (New-definition median line): By score cut-off, the share of all households who are targeted (that is, score at or below the 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), 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	0.1	100.0	0.7	Only poor targeted
≤9	0.1	100.0	1.0	Only poor targeted
≤14	0.4	83.4	3.9	5.0:1
≤19	1.5	57.0	9.4	1.3:1
≤24	4.0	59.9	26.4	1.5:1
≤29	7.6	48.9	41.1	1.0:1
≤34	14.0	40.5	62.8	0.7:1
≤39	22.9	30.7	77.8	0.4:1
≤44	35.4	21.3	83.5	0.3:1
≤49	47.5	17.2	90.4	0.2:1
≤54	59.9	14.6	96.8	0.2:1
≤59	70.3	12.7	98.6	0.1:1
≤64	77.4	11.7	99.9	0.1:1
≤69	85.3	10.6	100.0	0.1:1
≤74	92.6	9.8	100.0	0.1:1
≤79	97.3	9.3	100.0	0.1:1
≤84	98.8	9.2	100.0	0.1:1
≤89	99.7	9.1	100.0	0.1:1
≤94	99.9	9.1	100.0	0.1:1
≤100	100.0	9.1	100.0	0.1:1

**Tables for
The New-Definition \$1.25/day 2005 PPP Poverty Line**

Figure 3 (New-definition \$1.25/day line): Estimated poverty likelihoods associated with scores

If a household's score is then the likelihood (%) of being below the poverty line is:
0-4	2.0
5-9	2.0
10-14	2.0
15-19	2.0
20-24	1.4
25-29	0.0
30-34	0.0
35-39	0.0
40-44	0.0
45-49	0.0
50-54	0.0
55-59	0.0
60-64	0.0
65-69	0.0
70-74	0.0
75-79	0.0
80-84	0.0
85-89	0.0
90-94	0.0
95-100	0.0

Figure 6 (New-definition \$1.25/day line): Average differences between estimated and true poverty likelihoods for households in 1,000 bootstraps of $n = 16,384$, with confidence intervals by score range, scorecard applied to the validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0–4	+0.0	0.0	0.0	0.0
5–9	+2.0	0.0	0.0	0.0
10–14	+2.0	0.0	0.0	0.0
15–19	+2.0	0.0	0.0	0.0
20–24	+1.4	0.0	0.0	0.0
25–29	+0.0	0.0	0.0	0.0
30–34	–0.9	0.7	0.8	0.9
35–39	+0.0	0.0	0.0	0.0
40–44	+0.0	0.0	0.0	0.0
45–49	+0.0	0.0	0.0	0.0
50–54	+0.0	0.0	0.0	0.0
55–59	+0.0	0.0	0.0	0.0
60–64	+0.0	0.0	0.0	0.0
65–69	+0.0	0.0	0.0	0.0
70–74	+0.0	0.0	0.0	0.0
75–79	+0.0	0.0	0.0	0.0
80–84	+0.0	0.0	0.0	0.0
85–89	+0.0	0.0	0.0	0.0
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 7 (New-definition \$1.25/day line): Average differences between estimated poverty rates and true values for a group at a point in time in 1,000 bootstraps of various sample sizes, with confidence intervals, scorecard applied to the validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+0.1	0.0	0.7	1.0
4	+0.0	0.3	0.3	0.4
8	+0.0	0.2	0.2	0.3
16	+0.0	0.1	0.1	2.9
32	-0.0	0.1	0.1	1.6
64	+0.0	0.1	0.8	0.8
128	+0.0	0.4	0.4	0.4
256	+0.0	0.2	0.2	0.4
512	+0.0	0.1	0.2	0.2
1,024	+0.0	0.1	0.1	0.2
2,048	+0.0	0.1	0.1	0.1
4,096	+0.0	0.1	0.1	0.1
8,192	+0.0	0.0	0.0	0.1
16,384	+0.0	0.0	0.0	0.0

Figure 10 (New-definition \$1.25/day line): Shares of households by cut-off score and targeting classification, along with the hit rate and BPAC, scorecard applied to the validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	0.1	0.0	99.9	99.9	+0.0
≤9	0.1	0.1	0.0	99.8	99.9	+27.2
≤14	0.1	0.1	0.4	99.5	99.6	-200.5
≤19	0.1	0.1	1.4	98.4	98.5	-1,072.8
≤24	0.1	0.1	3.9	95.9	96.0	-3,121.0
≤29	0.1	0.1	7.5	92.3	92.4	-6,078.0
≤34	0.1	0.0	13.9	86.0	86.1	-11,290.6
≤39	0.1	0.0	22.8	77.1	77.2	-18,592.9
≤44	0.1	0.0	35.3	64.6	64.7	-28,821.9
≤49	0.1	0.0	47.4	52.5	52.6	-38,703.3
≤54	0.1	0.0	59.8	40.1	40.2	-48,870.3
≤59	0.1	0.0	70.1	29.7	29.9	-57,341.8
≤64	0.1	0.0	77.3	22.6	22.7	-63,183.9
≤69	0.1	0.0	85.2	14.7	14.8	-69,668.0
≤74	0.1	0.0	92.5	7.4	7.5	-75,624.2
≤79	0.1	0.0	97.2	2.7	2.8	-79,498.2
≤84	0.1	0.0	98.7	1.2	1.3	-80,722.5
≤89	0.1	0.0	99.6	0.3	0.4	-81,487.5
≤94	0.1	0.0	99.8	0.1	0.2	-81,641.0
≤100	0.1	0.0	99.9	0.0	0.1	-81,696.1

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

Figure 11 (New-definition \$1.25/day line): By score cut-off, the share of all households who are targeted (that is, score at or below the 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), 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	0.1	100.0	50.0	Only poor targeted
≤9	0.1	64.7	50.0	1.8:1
≤14	0.4	14.3	50.0	0.2:1
≤19	1.5	4.1	50.0	0.0:1
≤24	4.0	1.5	50.0	0.0:1
≤29	7.6	0.8	50.0	0.0:1
≤34	14.0	0.9	100.0	0.0:1
≤39	22.9	0.5	100.0	0.0:1
≤44	35.4	0.3	100.0	0.0:1
≤49	47.5	0.3	100.0	0.0:1
≤54	59.9	0.2	100.0	0.0:1
≤59	70.3	0.2	100.0	0.0:1
≤64	77.4	0.2	100.0	0.0:1
≤69	85.3	0.1	100.0	0.0:1
≤74	92.6	0.1	100.0	0.0:1
≤79	97.3	0.1	100.0	0.0:1
≤84	98.8	0.1	100.0	0.0:1
≤89	99.7	0.1	100.0	0.0:1
≤94	99.9	0.1	100.0	0.0:1
≤100	100.0	0.1	100.0	0.0:1

**Tables for
the New-Definition \$2.00/day 2005 PPP Poverty Line**

Figure 3 (New-definition \$2.00/day line): Estimated poverty likelihoods associated with scores

If a household's score is then the likelihood (%) of being below the poverty line is:
0-4	26.5
5-9	26.5
10-14	9.0
15-19	3.1
20-24	2.9
25-29	2.0
30-34	1.1
35-39	0.5
40-44	0.1
45-49	0.0
50-54	0.0
55-59	0.0
60-64	0.0
65-69	0.0
70-74	0.0
75-79	0.0
80-84	0.0
85-89	0.0
90-94	0.0
95-100	0.0

Figure 6 (New-definition \$2.00/day line): Average differences between estimated and true poverty likelihoods for households in 1,000 bootstraps of $n = 16,384$, with confidence intervals by score range, scorecard applied to the validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	+0.0	0.0	0.0	0.0
5-9	-73.5	36.7	36.7	36.7
10-14	+8.9	0.2	0.2	0.3
15-19	+3.1	0.0	0.0	0.0
20-24	+1.8	0.7	0.8	1.1
25-29	+0.5	0.8	1.0	1.2
30-34	-1.4	1.1	1.1	1.2
35-39	-1.2	0.9	1.0	1.1
40-44	+0.1	0.0	0.0	0.0
45-49	+0.0	0.0	0.0	0.0
50-54	-0.2	0.1	0.2	0.2
55-59	+0.0	0.0	0.0	0.0
60-64	+0.0	0.0	0.0	0.0
65-69	+0.0	0.0	0.0	0.0
70-74	+0.0	0.0	0.0	0.0
75-79	+0.0	0.0	0.0	0.0
80-84	+0.0	0.0	0.0	0.0
85-89	+0.0	0.0	0.0	0.0
90-94	+0.0	0.0	0.0	0.0
95-100	+0.0	0.0	0.0	0.0

Figure 7 (New-definition \$2.00/day line): Average differences between estimated poverty rates and true values for a group at a point in time in 1,000 bootstraps of various sample sizes, with confidence intervals, scorecard applied to the validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+0.1	1.0	1.5	1.6
4	-0.0	0.6	0.7	15.7
8	-0.1	0.4	5.2	9.7
16	-0.0	1.9	3.0	4.7
32	-0.1	1.7	2.3	3.1
64	-0.1	1.2	1.5	2.5
128	-0.1	0.9	1.2	1.6
256	-0.1	0.7	0.8	1.0
512	-0.1	0.5	0.6	0.8
1,024	-0.1	0.3	0.4	0.5
2,048	-0.1	0.2	0.3	0.4
4,096	-0.1	0.2	0.2	0.3
8,192	-0.1	0.1	0.1	0.2
16,384	-0.1	0.1	0.1	0.1

Figure 10 (New-definition \$2.00/day line): Shares of households by cut-off score and targeting classification, along with the hit rate and BPAC, scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.1	0.5	0.0	99.5	99.5	-77.4
≤9	0.1	0.4	0.0	99.5	99.6	-65.1
≤14	0.1	0.4	0.3	99.1	99.2	-2.7
≤19	0.1	0.4	1.4	98.1	98.2	-157.6
≤24	0.1	0.4	3.9	95.6	95.8	-611.4
≤29	0.2	0.3	7.4	92.1	92.3	-1,266.4
≤34	0.4	0.2	13.6	85.8	86.2	-2,421.7
≤39	0.5	0.0	22.4	77.0	77.5	-4,046.4
≤44	0.5	0.0	34.9	64.5	65.0	-6,353.9
≤49	0.5	0.0	47.0	52.5	53.0	-8,583.1
≤54	0.5	0.0	59.4	40.1	40.6	-10,869.6
≤59	0.5	0.0	69.7	29.7	30.3	-12,780.7
≤64	0.5	0.0	76.9	22.6	23.1	-14,098.6
≤69	0.5	0.0	84.8	14.7	15.2	-15,561.3
≤74	0.5	0.0	92.0	7.4	8.0	-16,905.0
≤79	0.5	0.0	96.8	2.7	3.2	-17,778.9
≤84	0.5	0.0	98.3	1.2	1.7	-18,055.1
≤89	0.5	0.0	99.2	0.3	0.8	-18,227.7
≤94	0.5	0.0	99.4	0.1	0.6	-18,262.3
≤100	0.5	0.0	99.5	0.0	0.5	-18,274.7

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

Figure 11 (New-definition \$2.00/day line): By score cut-off, the share of all households who are targeted (that is, score at or below the 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), 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	0.1	100.0	11.3	Only poor targeted
≤9	0.1	100.0	17.4	Only poor targeted
≤14	0.4	23.0	18.2	0.3:1
≤19	1.5	6.6	18.2	0.1:1
≤24	4.0	3.6	26.5	0.0:1
≤29	7.6	2.7	38.6	0.0:1
≤34	14.0	2.7	70.4	0.0:1
≤39	22.9	2.2	93.0	0.0:1
≤44	35.4	1.4	93.0	0.0:1
≤49	47.5	1.1	93.0	0.0:1
≤54	59.9	0.9	100.0	0.0:1
≤59	70.3	0.8	100.0	0.0:1
≤64	77.4	0.7	100.0	0.0:1
≤69	85.3	0.6	100.0	0.0:1
≤74	92.6	0.6	100.0	0.0:1
≤79	97.3	0.6	100.0	0.0:1
≤84	98.8	0.5	100.0	0.0:1
≤89	99.7	0.5	100.0	0.0:1
≤94	99.9	0.5	100.0	0.0:1
≤100	100.0	0.5	100.0	0.0:1

**Tables for
the New-Definition \$2.50/day 2005 PPP Poverty Line**

Figure 3 (New-definition \$2.50/day line): Estimated poverty likelihoods associated with scores

If a household's score is then the likelihood (%) of being below the poverty line is:
0-4	26.5
5-9	26.5
10-14	13.7
15-19	9.3
20-24	9.3
25-29	5.4
30-34	3.7
35-39	2.7
40-44	1.1
45-49	0.0
50-54	0.0
55-59	0.0
60-64	0.0
65-69	0.0
70-74	0.0
75-79	0.0
80-84	0.0
85-89	0.0
90-94	0.0
95-100	0.0

Figure 6 (New-definition \$2.50/day line): Average differences between estimated and true poverty likelihoods for households in 1,000 bootstraps of $n = 16,384$, with confidence intervals by score range, scorecard applied to the validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	+0.0	0.0	0.0	0.0
5-9	-73.5	36.7	36.7	36.7
10-14	+13.6	0.2	0.2	0.3
15-19	-1.9	4.4	5.3	7.3
20-24	-1.2	2.5	2.9	4.0
25-29	-4.6	3.4	3.6	4.1
30-34	-1.2	1.1	1.3	1.8
35-39	-1.2	1.1	1.2	1.5
40-44	+0.6	0.3	0.3	0.4
45-49	-0.7	0.5	0.6	0.6
50-54	-0.2	0.1	0.2	0.2
55-59	+0.0	0.0	0.0	0.0
60-64	+0.0	0.0	0.0	0.0
65-69	+0.0	0.0	0.0	0.0
70-74	+0.0	0.0	0.0	0.0
75-79	+0.0	0.0	0.0	0.0
80-84	+0.0	0.0	0.0	0.0
85-89	+0.0	0.0	0.0	0.0
90-94	+0.0	0.0	0.0	0.0
95-100	+0.0	0.0	0.0	0.0

Figure 7 (New-definition \$2.50/day line): Average differences between estimated poverty rates and true values for a group at a point in time in 1,000 bootstraps of various sample sizes, with confidence intervals, scorecard applied to the validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	-0.2	2.7	4.7	52.4
4	-0.6	10.6	15.5	24.5
8	-0.5	7.7	9.7	14.6
16	-0.3	4.7	5.8	9.1
32	-0.3	3.3	4.2	5.9
64	-0.3	2.6	3.1	4.0
128	-0.4	1.9	2.2	2.7
256	-0.4	1.3	1.6	2.1
512	-0.4	0.9	1.2	1.5
1,024	-0.4	0.7	0.8	1.1
2,048	-0.4	0.5	0.6	0.8
4,096	-0.4	0.3	0.4	0.5
8,192	-0.4	0.2	0.3	0.4
16,384	-0.4	0.2	0.2	0.3

Figure 10 (New-definition \$2.50/day line): Shares of households by cut-off score and targeting classification, along with the hit rate and BPAC, scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.1	1.6	0.0	98.4	98.4	-92.6
≤9	0.1	1.6	0.0	98.4	98.4	-88.5
≤14	0.1	1.5	0.3	98.0	98.1	-68.0
≤19	0.2	1.4	1.3	97.1	97.3	+3.4
≤24	0.5	1.1	3.5	94.9	95.4	-110.7
≤29	0.9	0.8	6.7	91.6	92.5	-310.2
≤34	1.2	0.5	12.8	85.5	86.7	-680.8
≤39	1.5	0.2	21.5	76.9	78.3	-1,205.7
≤44	1.5	0.1	33.9	64.5	66.0	-1,960.9
≤49	1.6	0.0	45.9	52.5	54.1	-2,689.8
≤54	1.6	0.0	58.3	40.1	41.7	-3,442.2
≤59	1.6	0.0	68.6	29.7	31.4	-4,070.9
≤64	1.6	0.0	75.8	22.6	24.2	-4,504.6
≤69	1.6	0.0	83.7	14.7	16.3	-4,985.8
≤74	1.6	0.0	90.9	7.4	9.1	-5,427.9
≤79	1.6	0.0	95.7	2.7	4.3	-5,715.5
≤84	1.6	0.0	97.2	1.2	2.8	-5,806.3
≤89	1.6	0.0	98.1	0.3	1.9	-5,863.1
≤94	1.6	0.0	98.3	0.1	1.7	-5,874.5
≤100	1.6	0.0	98.4	0.0	1.6	-5,878.6

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

Figure 11 (New-definition \$2.50/day line): By score cut-off, the share of all households who are targeted (that is, score at or below the 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), 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	0.1	100.0	3.7	Only poor targeted
≤9	0.1	100.0	5.7	Only poor targeted
≤14	0.4	23.0	6.0	0.3:1
≤19	1.5	13.9	12.6	0.2:1
≤24	4.0	13.2	32.1	0.2:1
≤29	7.6	11.3	52.1	0.1:1
≤34	14.0	8.5	72.1	0.1:1
≤39	22.9	6.4	89.2	0.1:1
≤44	35.4	4.3	93.2	0.0:1
≤49	47.5	3.4	97.7	0.0:1
≤54	59.9	2.7	100.0	0.0:1
≤59	70.3	2.3	100.0	0.0:1
≤64	77.4	2.1	100.0	0.0:1
≤69	85.3	1.9	100.0	0.0:1
≤74	92.6	1.8	100.0	0.0:1
≤79	97.3	1.7	100.0	0.0:1
≤84	98.8	1.7	100.0	0.0:1
≤89	99.7	1.6	100.0	0.0:1
≤94	99.9	1.6	100.0	0.0:1
≤100	100.0	1.6	100.0	0.0:1

**Tables for
the New-Definition \$3.75/day 2005 PPP Poverty Line**

Figure 3 (New-definition \$3.75/day line): Estimated poverty likelihoods associated with scores

If a household's score is then the likelihood (%) of being below the poverty line is:
0-4	100.0
5-9	73.5
10-14	53.1
15-19	45.6
20-24	45.0
25-29	31.2
30-34	24.9
35-39	16.2
40-44	7.3
45-49	3.0
50-54	0.8
55-59	0.8
60-64	0.3
65-69	0.1
70-74	0.1
75-79	0.0
80-84	0.0
85-89	0.0
90-94	0.0
95-100	0.0

Figure 6 (New-definition \$3.75/day line): Average differences between estimated and true poverty likelihoods for households in 1,000 bootstraps of $n = 16,384$, with confidence intervals by score range, scorecard applied to the validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0–4	+0.0	0.0	0.0	0.0
5–9	–26.5	13.3	13.3	13.3
10–14	–21.7	15.5	16.2	17.6
15–19	+6.5	6.4	7.5	10.2
20–24	–20.2	12.1	12.4	13.3
25–29	–8.2	5.9	6.3	7.2
30–34	+1.2	2.2	2.7	3.6
35–39	+4.1	1.6	2.0	2.6
40–44	+4.8	0.5	0.7	0.9
45–49	+0.9	0.5	0.6	0.8
50–54	–1.9	1.3	1.4	1.5
55–59	+0.7	0.1	0.1	0.1
60–64	–0.3	0.3	0.4	0.5
65–69	+0.1	0.0	0.0	0.0
70–74	+0.1	0.0	0.0	0.0
75–79	+0.0	0.0	0.0	0.0
80–84	+0.0	0.0	0.0	0.0
85–89	+0.0	0.0	0.0	0.0
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 7 (New-definition \$3.75/day line): Average differences between estimated poverty rates and true values for a group at a point in time in 1,000 bootstraps of various sample sizes, with confidence intervals, scorecard applied to the validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+0.7	39.9	53.1	71.3
4	-0.3	20.8	24.7	34.4
8	+0.2	13.8	16.9	23.5
16	+0.3	9.6	11.8	15.9
32	+0.3	6.9	8.0	10.9
64	+0.2	4.9	5.7	7.8
128	+0.1	3.5	4.0	5.2
256	+0.2	2.5	2.9	3.7
512	+0.1	1.7	2.1	2.6
1,024	+0.1	1.3	1.5	1.9
2,048	+0.1	0.9	1.1	1.4
4,096	+0.1	0.7	0.8	1.0
8,192	+0.1	0.5	0.6	0.8
16,384	+0.1	0.3	0.4	0.5

Figure 10 (New-definition \$3.75/day line): Shares of households by cut-off score and targeting classification, along with the hit rate and BPAC, scorecard applied to the validation sample

Score	< poverty line	< poverty line	≥ poverty line	≥ poverty line	Inclusion	
	correctly targeted	mistakenly non-targeted	mistakenly targeted	correctly non-targeted	+ Exclusion	See text
≤4	0.1	7.1	0.0	92.8	92.9	-98.3
≤9	0.1	7.1	0.0	92.8	92.9	-97.4
≤14	0.3	6.8	0.1	92.7	93.0	-89.7
≤19	0.7	6.4	0.8	92.1	92.8	-68.8
≤24	2.1	5.0	1.9	91.0	93.1	-14.4
≤29	3.4	3.7	4.2	88.7	92.1	+41.8
≤34	5.0	2.2	9.1	83.8	88.7	-26.8
≤39	6.0	1.2	17.0	75.9	81.8	-137.3
≤44	6.4	0.8	29.1	63.8	70.2	-305.9
≤49	6.7	0.4	40.8	52.1	58.8	-469.5
≤54	7.1	0.1	52.9	40.0	47.0	-638.5
≤59	7.1	0.1	63.2	29.7	36.8	-782.4
≤64	7.2	0.0	70.2	22.6	29.8	-881.4
≤69	7.2	0.0	78.2	14.7	21.8	-992.0
≤74	7.2	0.0	85.4	7.4	14.6	-1,093.6
≤79	7.2	0.0	90.2	2.7	9.8	-1,159.7
≤84	7.2	0.0	91.7	1.2	8.3	-1,180.6
≤89	7.2	0.0	92.6	0.3	7.4	-1,193.6
≤94	7.2	0.0	92.8	0.1	7.2	-1,196.3
≤100	7.2	0.0	92.8	0.0	7.2	-1,197.2

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

Figure 11 (New-definition \$3.75/day line): By score cut-off, the share of all households who are targeted (that is, score at or below the 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), 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	0.1	100.0	0.9	Only poor targeted
≤9	0.1	100.0	1.3	Only poor targeted
≤14	0.4	73.1	4.4	2.7:1
≤19	1.5	49.7	10.4	1.0:1
≤24	4.0	53.4	29.8	1.1:1
≤29	7.6	45.3	48.1	0.8:1
≤34	14.0	35.3	69.2	0.5:1
≤39	22.9	26.0	83.4	0.4:1
≤44	35.4	18.0	89.2	0.2:1
≤49	47.5	14.2	94.2	0.2:1
≤54	59.9	11.8	98.6	0.1:1
≤59	70.3	10.1	99.3	0.1:1
≤64	77.4	9.2	100.0	0.1:1
≤69	85.3	8.4	100.0	0.1:1
≤74	92.6	7.7	100.0	0.1:1
≤79	97.3	7.4	100.0	0.1:1
≤84	98.8	7.2	100.0	0.1:1
≤89	99.7	7.2	100.0	0.1:1
≤94	99.9	7.2	100.0	0.1:1
≤100	100.0	7.2	100.0	0.1:1

**Tables for
the New-Definition \$5.00/day 2005 PPP Poverty Line**

Figure 3 (New-definition \$5.00/day line): Estimated poverty likelihoods associated with scores

If a household's score is then the likelihood (%) of being below the poverty line is:
0-4	100.0
5-9	100.0
10-14	93.5
15-19	86.1
20-24	71.7
25-29	59.3
30-34	50.4
35-39	36.8
40-44	25.5
45-49	18.8
50-54	7.2
55-59	5.9
60-64	2.0
65-69	1.0
70-74	0.8
75-79	0.0
80-84	0.0
85-89	0.0
90-94	0.0
95-100	0.0

Figure 6 (New-definition \$5.00/day line): Average differences between estimated and true poverty likelihoods for households in 1,000 bootstraps of $n = 16,384$, with confidence intervals by score range, scorecard applied to the validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0–4	+0.0	0.0	0.0	0.0
5–9	+0.0	0.0	0.0	0.0
10–14	+7.7	6.7	8.1	10.9
15–19	+30.4	6.6	8.0	10.0
20–24	–15.6	9.0	9.2	9.7
25–29	–5.6	4.6	4.9	5.7
30–34	+2.0	2.8	3.3	4.4
35–39	–1.7	2.2	2.7	3.7
40–44	+7.0	1.5	1.8	2.2
45–49	+4.4	1.4	1.6	2.2
50–54	–3.5	2.4	2.5	2.8
55–59	+0.6	1.0	1.1	1.4
60–64	–0.6	0.6	0.8	1.0
65–69	+0.9	0.1	0.1	0.1
70–74	–0.3	0.4	0.5	0.7
75–79	–1.8	1.3	1.4	1.6
80–84	–7.0	5.0	5.2	5.8
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

Figure 7 (New-definition \$5.00/day line): Average differences between estimated poverty rates and true values for a group at a point in time in 1,000 bootstraps of various sample sizes, with confidence intervals, scorecard applied to the validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+0.6	62.4	65.8	86.5
4	+0.1	30.1	36.2	48.2
8	+0.6	20.0	24.0	33.8
16	+0.7	14.8	17.6	24.0
32	+0.9	9.7	11.6	15.5
64	+0.6	6.9	8.4	11.8
128	+0.6	5.1	6.3	7.8
256	+0.6	3.7	4.5	5.6
512	+0.5	2.7	3.2	3.9
1,024	+0.5	1.8	2.2	2.8
2,048	+0.5	1.3	1.6	2.0
4,096	+0.5	0.9	1.1	1.5
8,192	+0.5	0.7	0.8	1.0
16,384	+0.5	0.5	0.6	0.7

Figure 10 (New-definition \$5.00/day line): Shares of households by cut-off score and targeting classification, along with the hit rate and BPAC, scorecard applied to the validation sample

Score	<u>Inclusion:</u>	<u>Undercoverage:</u>	<u>Leakage:</u>	<u>Exclusion:</u>	<u>Hit rate</u>	<u>BPAC</u>
	< poverty line correctly targeted	< poverty line mistakenly non-targeted	≥ poverty line mistakenly targeted	≥ poverty line correctly non-targeted	Inclusion + Exclusion	See text
≤4	0.1	18.7	0.0	81.3	81.3	-99.3
≤9	0.1	18.7	0.0	81.3	81.3	-99.0
≤14	0.4	18.4	0.1	81.2	81.5	-95.8
≤19	1.0	17.7	0.5	80.8	81.8	-86.6
≤24	3.1	15.7	0.9	80.3	83.4	-62.3
≤29	5.3	13.4	2.3	79.0	84.3	-30.9
≤34	8.4	10.3	5.6	75.7	84.1	+19.9
≤39	11.9	6.9	11.1	70.2	82.0	+40.9
≤44	14.4	4.3	21.0	60.2	74.7	-12.1
≤49	16.3	2.4	31.2	50.1	66.4	-66.4
≤54	17.6	1.1	42.3	39.0	56.6	-125.5
≤59	18.2	0.6	52.1	29.2	47.4	-177.8
≤64	18.4	0.3	59.0	22.3	40.7	-214.5
≤69	18.5	0.3	66.8	14.4	32.9	-256.5
≤74	18.6	0.2	74.0	7.3	25.8	-294.7
≤79	18.7	0.1	78.7	2.6	21.3	-319.6
≤84	18.7	0.0	80.1	1.2	19.9	-327.1
≤89	18.7	0.0	81.0	0.3	19.0	-332.1
≤94	18.7	0.0	81.2	0.1	18.8	-333.1
≤100	18.7	0.0	81.3	0.0	18.7	-333.4

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

Figure 11 (New-definition \$5.00/day line): By score cut-off, the share of all households who are targeted (that is, score at or below the 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), 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	0.1	100.0	0.3	Only poor targeted
≤9	0.1	100.0	0.5	Only poor targeted
≤14	0.4	83.4	1.9	5.0:1
≤19	1.5	68.1	5.4	2.1:1
≤24	4.0	76.9	16.4	3.3:1
≤29	7.6	70.2	28.5	2.4:1
≤34	14.0	60.2	45.0	1.5:1
≤39	22.9	51.7	63.3	1.1:1
≤44	35.4	40.7	76.9	0.7:1
≤49	47.5	34.3	87.0	0.5:1
≤54	59.9	29.4	94.1	0.4:1
≤59	70.3	25.9	97.0	0.3:1
≤64	77.4	23.8	98.4	0.3:1
≤69	85.3	21.7	98.5	0.3:1
≤74	92.6	20.1	99.1	0.3:1
≤79	97.3	19.2	99.5	0.2:1
≤84	98.8	19.0	100.0	0.2:1
≤89	99.7	18.8	100.0	0.2:1
≤94	99.9	18.8	100.0	0.2:1
≤100	100.0	18.7	100.0	0.2:1

**Tables for
the Old-Definition Deep Poverty Line**

Figure 3 (Old-definition deep 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	92.2
15-19	70.6
20-24	58.0
25-29	42.5
30-34	37.7
35-39	24.3
40-44	15.7
45-49	10.7
50-54	2.2
55-59	2.2
60-64	0.5
65-69	0.1
70-74	0.1
75-79	0.0
80-84	0.0
85-89	0.0
90-94	0.0
95-100	0.0

Figure 6 (Old-definition deep line): Average differences between estimated and true poverty likelihoods for households in 1,000 bootstraps of $n = 16,384$, with confidence intervals by score range, scorecard applied to the validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	+0.0	0.0	0.0	0.0
5-9	+0.0	0.0	0.0	0.0
10-14	+17.5	9.5	11.2	15.0
15-19	+26.9	6.7	7.9	9.9
20-24	-18.7	11.0	11.3	12.1
25-29	-10.4	7.1	7.4	7.9
30-34	+0.5	2.6	3.1	4.1
35-39	-0.8	2.1	2.4	3.2
40-44	+5.8	1.1	1.3	1.8
45-49	+6.4	0.8	0.9	1.2
50-54	-3.2	2.0	2.1	2.3
55-59	+0.9	0.4	0.5	0.7
60-64	-0.7	0.6	0.6	0.7
65-69	+0.1	0.0	0.0	0.0
70-74	+0.1	0.0	0.0	0.0
75-79	+0.0	0.0	0.0	0.0
80-84	+0.0	0.0	0.0	0.0
85-89	+0.0	0.0	0.0	0.0
90-94	+0.0	0.0	0.0	0.0
95-100	+0.0	0.0	0.0	0.0

Figure 7 (Old-definition deep line): Average differences between estimated poverty rates and true values for a group at a point in time in 1,000 bootstraps of various sample sizes, with confidence intervals, scorecard applied to the validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+0.7	56.7	63.4	84.2
4	-0.0	25.5	31.0	40.1
8	+0.7	17.7	20.4	28.8
16	+0.7	12.0	14.3	19.5
32	+0.9	8.6	10.1	12.8
64	+0.7	5.8	6.7	8.4
128	+0.6	4.3	4.9	6.4
256	+0.6	3.0	3.5	4.4
512	+0.6	2.2	2.5	3.3
1,024	+0.5	1.5	1.8	2.4
2,048	+0.6	1.1	1.3	1.7
4,096	+0.6	0.8	1.0	1.3
8,192	+0.6	0.6	0.7	0.9
16,384	+0.6	0.4	0.5	0.6

**Tables for
100% of the Old-Definition National Poverty Line**

Figure 3 (100% of the old-definition national line):
Estimated poverty likelihoods associated with scores

If a household's score is then the likelihood (%) of being below the poverty line is:
0-4	100.0
5-9	100.0
10-14	94.0
15-19	86.1
20-24	76.1
25-29	65.7
30-34	57.3
35-39	45.7
40-44	34.2
45-49	22.9
50-54	9.2
55-59	8.3
60-64	2.7
65-69	1.1
70-74	0.9
75-79	0.0
80-84	0.0
85-89	0.0
90-94	0.0
95-100	0.0

**Figure 6 (100% of the old-definition national line):
Average differences between estimated and true
poverty likelihoods for households in 1,000
bootstraps of $n = 16,384$, with confidence intervals
by score range, scorecard applied to the validation
sample**

Score	Diff.	Difference between estimate and true value		
		Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	+0.0	0.0	0.0	0.0
5-9	+0.0	0.0	0.0	0.0
10-14	+1.7	5.2	6.3	7.9
15-19	+25.4	6.6	7.9	10.6
20-24	-17.9	9.7	9.9	10.1
25-29	-8.5	5.8	6.1	6.8
30-34	+0.0	2.7	3.3	4.2
35-39	+2.7	2.3	2.8	3.6
40-44	+10.6	1.6	1.9	2.4
45-49	+5.6	1.5	1.8	2.3
50-54	-3.8	2.6	2.8	3.1
55-59	-0.5	1.2	1.5	2.1
60-64	-0.7	0.7	0.9	1.2
65-69	+0.8	0.2	0.2	0.3
70-74	-0.8	0.7	0.7	0.8
75-79	-1.8	1.3	1.4	1.6
80-84	-7.0	5.0	5.2	5.8
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

**Figure 7 (100% of the old-definition national line):
Average differences between estimated poverty rates
and true values for a group at a point in time in
1,000 bootstraps of various sample sizes, with
confidence intervals, scorecard applied to the
validation sample**

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+0.7	61.4	74.1	85.3
4	+0.6	31.1	38.3	50.4
8	+1.2	21.7	26.1	36.6
16	+1.2	15.2	18.9	25.4
32	+1.4	10.5	12.5	16.7
64	+1.0	7.5	9.0	13.0
128	+1.0	5.4	6.4	8.7
256	+1.0	3.9	4.6	5.8
512	+0.9	2.7	3.3	4.2
1,024	+0.9	1.9	2.2	3.0
2,048	+1.0	1.4	1.7	2.2
4,096	+1.0	1.0	1.2	1.5
8,192	+1.0	0.7	0.8	1.1
16,384	+1.0	0.5	0.6	0.8

**Tables for
150% of the Old-Definition National Poverty Line**

Figure 3 (150% of the old-definition national line):
Estimated poverty likelihoods associated with scores

If a household's score is then the likelihood (%) of being below the poverty line is:
0-4	100.0
5-9	100.0
10-14	98.4
15-19	96.9
20-24	96.2
25-29	88.9
30-34	82.3
35-39	75.9
40-44	69.1
45-49	61.3
50-54	40.0
55-59	33.0
60-64	19.1
65-69	16.7
70-74	14.6
75-79	8.0
80-84	2.4
85-89	1.5
90-94	0.0
95-100	0.0

Figure 6 (150% of the old-definition national line):
Average differences between estimated and true
poverty likelihoods for households in 1,000
bootstraps of $n = 16,384$, with confidence intervals
by score range, scorecard applied to the validation
sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	+0.0	0.0	0.0	0.0
5-9	+0.0	0.0	0.0	0.0
10-14	-1.6	0.8	0.8	0.8
15-19	+11.6	4.8	5.8	7.2
20-24	+0.7	1.3	1.6	2.2
25-29	-3.1	2.5	2.7	2.9
30-34	-4.2	3.1	3.5	3.9
35-39	-1.3	1.9	2.4	3.0
40-44	+15.3	2.0	2.4	3.2
45-49	+11.4	2.1	2.4	3.4
50-54	-5.0	3.5	3.7	4.1
55-59	+7.5	1.9	2.2	2.8
60-64	-7.8	5.0	5.3	5.6
65-69	+4.0	1.6	1.9	2.4
70-74	-0.6	1.9	2.2	2.9
75-79	+2.2	1.3	1.6	2.2
80-84	-4.6	3.8	4.1	4.6
85-89	+1.5	0.0	0.0	0.0
90-94	+0.0	0.0	0.0	0.0
95-100	+0.0	0.0	0.0	0.0

Figure 7 (150% of the old-definition national line):
Average differences between estimated poverty rates and true values for a group at a point in time in 1,000 bootstraps of various sample sizes, with confidence intervals, scorecard applied to the validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+0.7	68.1	79.6	88.8
4	+1.9	36.5	44.0	54.0
8	+2.8	26.0	32.7	43.8
16	+2.9	19.0	23.7	30.3
32	+2.8	14.3	16.2	21.4
64	+2.7	9.4	11.9	14.9
128	+2.9	6.9	8.0	10.5
256	+2.8	4.9	5.8	7.1
512	+2.7	3.4	3.9	5.2
1,024	+2.8	2.4	3.0	3.8
2,048	+2.9	1.7	2.0	2.6
4,096	+2.9	1.3	1.5	1.9
8,192	+2.8	0.9	1.0	1.3
16,384	+2.9	0.6	0.7	1.0

**Tables for
200% of the Old-Definition National Poverty Line**

Figure 3 (200% of the old-definition national line):
Estimated poverty likelihoods associated with scores

If a household's score is then the likelihood (%) of being below the poverty line is:
0-4	100.0
5-9	100.0
10-14	98.5
15-19	97.2
20-24	97.2
25-29	95.9
30-34	95.2
35-39	91.3
40-44	83.7
45-49	78.1
50-54	63.3
55-59	57.1
60-64	40.9
65-69	37.5
70-74	27.1
75-79	13.4
80-84	6.1
85-89	6.1
90-94	5.1
95-100	0.0

Figure 6 (200% of the old-definition national line):
Average differences between estimated and true
poverty likelihoods for households in 1,000
bootstraps of $n = 16,384$, with confidence intervals
by score range, scorecard applied to the validation
sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	+0.0	0.0	0.0	0.0
5-9	+0.0	0.0	0.0	0.0
10-14	-1.5	0.7	0.7	0.7
15-19	-2.8	1.4	1.4	1.4
20-24	+1.7	1.3	1.6	2.2
25-29	-1.7	1.3	1.4	1.5
30-34	-1.4	1.2	1.2	1.5
35-39	+3.3	1.5	1.9	2.4
40-44	+5.6	1.7	2.1	2.7
45-49	+11.3	2.0	2.4	3.1
50-54	-4.8	3.3	3.4	3.7
55-59	+8.7	2.1	2.6	3.5
60-64	-2.8	2.8	3.2	4.1
65-69	+7.5	2.4	2.8	3.5
70-74	-5.4	3.9	4.1	4.5
75-79	+2.8	1.9	2.2	2.8
80-84	-2.9	3.1	3.7	5.0
85-89	-7.9	6.3	6.6	7.7
90-94	+5.1	0.0	0.0	0.0
95-100	+0.0	0.0	0.0	0.0

Figure 7 (200% of the old-definition national line):
Average differences between estimated poverty rates and true values for a group at a point in time in 1,000 bootstraps of various sample sizes, with confidence intervals, scorecard applied to the validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+0.0	70.3	78.3	90.9
4	+1.0	37.3	42.9	57.2
8	+2.2	27.7	32.2	39.9
16	+2.1	18.1	22.6	30.4
32	+2.5	13.3	16.6	21.3
64	+2.5	9.7	11.2	14.8
128	+2.7	6.8	8.2	10.4
256	+2.5	4.8	5.6	7.2
512	+2.5	3.3	3.9	5.6
1,024	+2.4	2.3	2.7	3.7
2,048	+2.5	1.7	1.9	2.7
4,096	+2.5	1.2	1.4	1.9
8,192	+2.5	0.8	1.0	1.2
16,384	+2.5	0.6	0.7	0.9

**Tables for
the Old-Definition \$3.75/day 2005 PPP Poverty Line**

Figure 3 (Old-definition \$3.75/day line): Estimated poverty likelihoods associated with scores

If a household's score is then the likelihood (%) of being below the poverty line is:
0-4	100.0
5-9	100.0
10-14	77.5
15-19	53.7
20-24	50.2
25-29	34.5
30-34	24.7
35-39	18.2
40-44	9.2
45-49	6.1
50-54	0.8
55-59	0.7
60-64	0.2
65-69	0.1
70-74	0.1
75-79	0.0
80-84	0.0
85-89	0.0
90-94	0.0
95-100	0.0

Figure 6 (Old-definition \$3.75/day line): Average differences between estimated and true poverty likelihoods for households in 1,000 bootstraps of $n = 16,384$, with confidence intervals by score range, scorecard applied to the validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	+0.0	0.0	0.0	0.0
5-9	+0.0	0.0	0.0	0.0
10-14	+2.7	9.5	11.2	15.0
15-19	+21.5	6.3	7.4	9.8
20-24	-21.6	12.6	13.0	13.7
25-29	-9.9	6.8	7.2	8.1
30-34	-5.1	3.7	4.0	4.5
35-39	+4.2	1.7	2.0	2.6
40-44	+3.9	0.8	1.0	1.2
45-49	+3.3	0.6	0.7	0.9
50-54	-1.8	1.3	1.3	1.4
55-59	+0.6	0.1	0.1	0.1
60-64	-0.3	0.3	0.4	0.5
65-69	+0.1	0.0	0.0	0.0
70-74	+0.1	0.0	0.0	0.0
75-79	+0.0	0.0	0.0	0.0
80-84	+0.0	0.0	0.0	0.0
85-89	+0.0	0.0	0.0	0.0
90-94	+0.0	0.0	0.0	0.0
95-100	+0.0	0.0	0.0	0.0

Figure 7 (Old-definition \$3.75/day line): Average differences between estimated poverty rates and true values for a group at a point in time in 1,000 bootstraps of various sample sizes, with confidence intervals, scorecard applied to the validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	+0.8	45.1	54.9	73.8
4	-0.2	22.4	28.1	38.0
8	+0.2	15.1	18.1	25.1
16	+0.3	10.2	12.7	16.4
32	+0.3	7.3	8.9	11.3
64	+0.2	5.4	6.4	8.1
128	+0.1	3.7	4.3	5.6
256	+0.1	2.6	3.1	3.9
512	+0.0	1.9	2.2	2.8
1,024	+0.0	1.4	1.6	2.0
2,048	+0.0	1.0	1.1	1.5
4,096	+0.0	0.7	0.8	1.1
8,192	+0.1	0.5	0.6	0.8
16,384	+0.1	0.3	0.4	0.5

**Tables for
the Old-Definition \$5.00/day 2005 PPP Poverty Line**

Figure 3 (Old-definition \$5.00/day line): Estimated poverty likelihoods associated with scores

If a household's score is then the likelihood (%) of being below the poverty line is:
0-4	100.0
5-9	100.0
10-14	94.0
15-19	86.1
20-24	75.0
25-29	61.2
30-34	53.7
35-39	40.8
40-44	27.4
45-49	20.5
50-54	7.1
55-59	5.7
60-64	2.1
65-69	1.1
70-74	0.9
75-79	0.0
80-84	0.0
85-89	0.0
90-94	0.0
95-100	0.0

Figure 6 (Old-definition \$5.00/day line): Average differences between estimated and true poverty likelihoods for households in 1,000 bootstraps of $n = 16,384$, with confidence intervals by score range, scorecard applied to the validation sample

Score	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
0-4	+0.0	0.0	0.0	0.0
5-9	+0.0	0.0	0.0	0.0
10-14	+8.1	6.7	8.1	10.9
15-19	+27.6	6.7	7.7	10.6
20-24	-14.0	8.2	8.4	8.9
25-29	-10.0	6.7	7.0	7.5
30-34	-2.4	2.8	3.4	4.2
35-39	+3.7	2.3	2.7	3.6
40-44	+7.9	1.5	1.8	2.3
45-49	+5.8	1.4	1.7	2.1
50-54	-4.2	2.8	2.9	3.2
55-59	-3.0	2.1	2.3	2.7
60-64	-0.4	0.6	0.8	1.0
65-69	+0.8	0.2	0.2	0.3
70-74	-0.2	0.4	0.5	0.7
75-79	+0.0	0.0	0.0	0.0
80-84	-7.0	5.0	5.2	5.8
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

Figure 7 (Old-definition \$5.00/day line): Average differences between estimated poverty rates and true values for a group at a point in time in 1,000 bootstraps of various sample sizes, with confidence intervals, scorecard applied to the validation sample

Sample Size <i>n</i>	Difference between estimate and true value			
	Diff.	Confidence interval (\pm percentage points)		
		90-percent	95-percent	99-percent
1	-0.3	63.1	69.9	85.5
4	-0.3	30.4	37.1	49.4
8	+0.5	21.7	25.5	37.1
16	+0.6	15.4	18.0	24.0
32	+0.8	10.4	12.6	16.0
64	+0.5	7.2	8.6	11.9
128	+0.6	5.2	6.4	8.2
256	+0.5	3.8	4.5	5.8
512	+0.4	2.7	3.1	3.9
1,024	+0.4	1.8	2.1	3.0
2,048	+0.5	1.3	1.6	2.1
4,096	+0.5	1.0	1.1	1.5
8,192	+0.5	0.7	0.8	1.1
16,384	+0.5	0.5	0.5	0.7