Simple Poverty Scorecard® Poverty-Assessment Tool Egypt

Mark Schreiner

27 April 2010

This document is at `SimplePovertyScorecard.com`

Abstract

The Simple Poverty Scorecard-brand poverty-assessment tool uses 10 low-cost indicators from Egypt's 2004/5 Household Income, 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 Egypt to measure poverty rates, to track changes in poverty rates over time, and to segment clients for differentiated treatment.

Acknowledgements

This paper was funded by Grameen-Jameel Pan-Arab Microfinance Ltd. The data were provided by Egypt's Central Agency for Public Mobilisation and Statistics, with thanks to Mohamed Morsey, Abdel Hamid Sahrf Eldin, Khaled Mohamed Maher, and Madiha Saleh. Thanks go also to Julia Assaad, Nigel Bigger, Sharlene Brown, Reem Nejdawi, Sarah Sabry, Hoda Salman, Brian Slocum, and Jeff Toohig. Special thanks go to Heba El-Laithy for the computation of poverty lines. This scorecard was re-branded by Grameen Foundation (GF) as a Progress out of Poverty Index® tool. The PPI® is a performance-management tool that GF promotes to help organizations achieve their social objectives more effectively. "Progress out of Poverty Index" and "PPI" are Registered Trademarks of Innovations for Poverty Action. "Simple Poverty Scorecard" is a Registered Trademark of Microfinance Risk Management, L.L.C. for its brand of poverty-assessment tools.

Simple Poverty Scorecard® Poverty-Assessment Tool

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

Indicator	Response	Points	Score
1. How many members does the household have?	A. Seven or more	0	
	B. Six	5	
	C. Five	11	
	D. Four	18	
	E. Three	19	
	F. One or two	36	
2. Do all children ages 6 to 18 attend school?	A. No	0	
	B. Yes	2	
	C. No children 6 to 18	4	
3. Can the female head/spouse read and write?	A. No	0	
	B. No female head/spouse	4	
	C. Yes	7	
4. In their main line of work, do any household members have non-permanent (temporary, seasonal, or irregular) wage jobs?	A. Yes	0	
	B. No	7	
5. What is the material of the walls of the residence?	A. Stones, mud, wood, tin, asbestos, or other	0	
	B. Bricks with mortar	4	
	C. Concrete	6	
6. How many rooms does the residence of the household have (including parlor/reception hall)?	A. One	0	
	B. Two	1	
	C. Three	2	
	D. Four or more	8	
7. What is the source of water for the household?	A. Well, pump, public network with no connection, public network with tap outside building, or other	0	
	B. Public network with tap inside building	4	
8. What toilet arrangement does the household have?	A. No toilet available, or shared toilet	0	
	B. Private non-flush toilet	2	
	C. Private flush toilet	7	
9. Does the household own any gas or electric water heaters?	A. No	0	
	B. Yes	6	
10. Does the household own any clothes-washing machines?	A. No	0	
	B. Yes, only non-automatic	4	
	C. Yes, automatic	15	

Score:

Simple Poverty Scorecard® Poverty-Assessment Tool Egypt

1. Introduction

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

The direct approach to poverty measurement via surveys is difficult and costly. For example, Egypt's 2004/5 Household Income, Expenditure, and Consumption Survey (HIECS) runs more than 50 pages, with consumption module with hundreds of questions, such as: "How much market-price *baladi* bread did the household consume in the past month? How much was this worth? How much *baladi* bread did the household receive as an in-kind transfer? How much was this worth? Now then, how much subsidized *baladi* bread did the household consume in the past month . . .".

In contrast, the indirect approach via the scorecard is simple, quick, and inexpensive. It uses ten verifiable indicators (such as "What is the material of the walls of the residence?" or "What toilet arrangement does the household have?") to get a score that is highly correlated with poverty status as measured by consumption from the lengthy survey.

The scorecard differs from "proxy means tests" (Coady, Grosh, and Hoddinott, 2002) in that it is tailored to the capabilities and purposes not of national governments but rather of local, pro-poor organizations. The feasible poverty-measurement options for these local organizations are typically subjective and relative (such as participatory wealth ranking by skilled field workers) or blunt (such as rules based on land-ownership or housing quality). These approaches may be costly, their results are not comparable across organizations nor across countries, and their accuracy and precision are unknown.

Suppose an organization wants to know what share of its participants are below a poverty line, perhaps because it wants to relate their poverty status to the Millennium Development Goals' $1.25/day poverty line at 2005 purchase-power parity (PPP). Or an organization might want to report (as required of USAID microenterprise partners) how many of its participants are among the poorest half of people below the national poverty line. Or an organization might want to measure movement across a poverty line (for example, Daley-Harris, 2009). In these cases, what is needed is an consumption-based, objective tool with known accuracy that can serve for monitoring, management, and/or targeting. While consumption surveys are costly even for governments, many small, local organizations can afford to implement a simple, inexpensive scorecard.

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, 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 these tools do not work, but because they are presented (when they are presented at all) as tables of regression coefficients incomprehensible to non-specialists (with indicator names such as "`LGHHSZ_2`", negative points, and points with many decimal places). Thanks to the predictive-modeling phenomenon known as the "flat maximum", simple, trasparent scorecards are often about as accurate as complex, opaque ones.

The technical approach here is also 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 standard 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 based on the 2004/5 HIECS conducted by Egypt's Central Agency for Public Mobilization and Statistics (CAPMAS). Indicators for the scorecard are selected to be:

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

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

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

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

Third, the scorecard can estimate changes in the poverty rate for a given group of households (or for two independent representative samples of households from the same population) between two points in time. This estimate is the change in the average poverty likelihood of the group(s) of households over time.

The scorecard can also be used for targeting services to poorer households. To help managers select a targeting cut-off, 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 and Egypt's upper national poverty line. Scores from this scorecard are calibrated to poverty likelihoods for seven poverty lines.

CAPMAS provided this project with a 25-percent random sample from the full 2004/5 HIECS. The scorecard is then constructed and calibrated using a sub-sample of this data, and its accuracy is then validated on a different sub-sample. While all three scoring estimators are unbiased when applied to the population from which they are derived (that is, they match the true value on average in repeated samples from the

same population from which the scorecard is built), they are—like all predictive models—biased to some extent when applied to a different population.[1]

Thus, while the indirect scoring approach is less costly than the direct survey approach, it is also biased in practice. (The direct survey approach is unbiased by definition.) There is bias because scoring must assume that the future relationships between indicators and poverty will be the same as in the data used to build the scorecard. It must also assume that these relationships will be the same in all sub-groups as in the population as a whole.[2] Of course, these assumptions—ubiquitous and inevitable in predictive modeling—hold only partly.

When applied to the 2004/5 validation sample for Egypt with the upper national poverty line and $n = 16{,}384$, the difference between scorecard estimates of groups' poverty rates and the true rates at a point in time is +0.4 percentage points. Across all seven lines, the average absolute difference is +0.4 percentage points, and the maximum difference is +1.2 percentage points. Because the 2004/5 validation sample is representative of the same population as is the data used to construct the scorecard and because all the data come from the same time frame, the scorecard estimators are unbiased and these observed differences are due to sampling variation; the average

[1] Important examples of "different populations" are nationally representative samples at another point in time or non-representative sub-groups (Tarozzi and Deaton, 2007).
[2] Bias may also result from changes over time in the quality of data collection, from changes in the real value of poverty lines, from imperfect adjustment of poverty lines to account for differences in cost-of-living across time or geographic regions, or from sampling variation.

difference would be zero if the 2004/5 HIECS were to be repeatedly redrawn and divided into sub-samples before repeating the entire scorecard-building and accuracy-testing process.

For $n = 16{,}384$, the 90-percent confidence intervals for these estimates are ± 0.5 percentage points or less. For $n = 1{,}024$, these intervals are ± 1.9 percentage points or less.

Section 2 below documents data, poverty rates, and poverty lines for Egypt. Sections 3 and 4 describe scorecard construction and offer practical guidelines for use. 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, and Section 8 covers targeting. Section 9 places the new scorecard here in the context of similar existing exercises for Egypt. The final 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 lines to which scores are calibrated.

2.1 Data

The scorecard is based on data from 11,774 households randomly sampled from the 2004/5 HIECS. This is the best, most recent national consumption survey available for Egypt. For scoring, these 11,774 households are further divided into three sub-samples (Figure 2):

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

2.2 Poverty rates and poverty lines

2.2.1 Rates

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

Beyond this general definition, there two special cases, *household-level poverty rates* and *person-level poverty rates*. With household-level rates, each household is counted as if it had only one person, regardless of true household size, so all households

are counted equally. With person-level rates (the "head-count index"), each household is weighted by the number of people in it, so larger households have greater weight.

For example, consider a group of two households, the first with one member and the second with two members. Suppose further that the first household has per-capita consumption above a poverty line (it is "non-poor") and that the second household has per-capita consumption below a poverty line (it is "poor"). The household-level rate counts both households as if they had only one person and so gives a poverty rate for the group of $1 \div (1 + 1) = 50$ percent. In contrast, the person-level rate weighs each household by the number of people in it and so gives a poverty rate for the group of $2 \div (1 + 2) = 67$ percent.

Which rate is more relevant depends on the situation. If an organization's "participants" include all the people in a household, then the person-level rate is relevant. Governments, for example, are concerned with the well-being of their people, regardless of how those people are arranged in households, so governments typically report person-level poverty rates.

If an organization has only one "participant" per household, however, then the household-level rate may be relevant. For example, if a microlender has only one borrower in a household, then it might want to report household-level poverty rates.

The scorecard here is constructed using Egypt's 2004/5 HIECS and household-level lines, scores are calibrated to household-level poverty likelihoods, and accuracy is

measured for household-level rates. This use of household-level rates reflects the belief that they are the most relevant for most pro-poor organizations.

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

2.2.2 Poverty lines

Figure 3 reports average poverty lines for Egypt as a whole and for its seven regions. It also reports poverty rates at both the household-level and the person-level.

Egypt's food poverty line is defined as the cost of a food basket that satisfies a minimum caloric requirement as determined by prices in a given region and by the age, sex, and activity level (proxied by urban/rural location) of people in a given household. Heba El-Laithy used the 2004/5 HIECS to derive the cost of this basket and the poverty-line formulas for Egypt's seven regions (World Bank, 2007).[3] For Egypt as a whole, the average food line was EGP2.73 per person per day (Figure 3), giving a household-level poverty rate of 2.5 percent and a person-level rate of 3.9 percent.

Using the approach in Ravallion (1994), Egypt's lower national line is then defined as the food line plus the non-food consumption observed for households in the 2004/5 HIECS whose total consumption is at the food line. Like the food line, this lower

[3] El-Laithy was instrumental in enabling this paper to use the same algorithm to generate household-specific poverty lines as in World Bank (2008).

national line is household-specific and accounts for differences in household composition, differences in the cost-of-living across regions, and economies of scale within households. The average lower line for Egypt is EGP3.90, giving a household-level poverty rate of 13.9 percent and a person-level rate of 19.4 percent (Figure 3).

Egypt's upper national poverty line is defined as the food line plus the non-food consumption for households whose observed food consumption is at the food line. The average upper line for Egypt is EGP5.08, giving a household-level poverty rate of 31.4 percent and a person-level rate of 40.3 percent (Figure 3).

The upper national poverty line is used here to construct the scorecard. Because local pro-poor organizations in Egypt may want to use different or various poverty lines, this paper calibrates scores from its single scorecard to poverty likelihoods for seven lines:

- Upper national
- Lower national
- Food
- USAID "extreme"
- $1.25/day 2005 PPP
- $2.50/day 2005 PPP
- $3.75/day 2005 PPP

The USAID "extreme" line is defined as the median consumption of people (not households) below the upper national line (U.S. Congress, 2002). It is calculated by region and averages EGP3.98 overall.

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

- 2005 PPP exchange rate for "individual consumption expenditure by households" (World Bank, 2008): EGP2.02 per $1.00
- Consumer price indices from the Central Bank of Egypt:[4] 133.1 on average for July 2004 to June 2005 when the HIECS was in the field, and 135.0 for all of 2005

The $1.25/day 2005 PPP line for Egypt as a whole in July 2004 through June 2005 is then (Sillers, 2006):

$$(\text{2005 PPP exchange rate}) \cdot \$1.25 \cdot \frac{\text{CPI}_{\text{July 2004-June 2005}}}{\text{CPI}_{\text{Ave. 2005}}} =$$

$$\left(\frac{\text{EGP2.02}}{\$1.00}\right) \cdot \$1.25 \cdot \frac{133.1}{135.0} = \text{EGP2.4895}.$$

This $1.25/day 2005 PPP line applies to Egypt as a whole. It is adjusted for differences in cost-of-living across regions and households using the upper national poverty line as a deflator. That is, each household-specific $1.25/day 2005 PPP line is defined as the national-level $1.25/day 2005 PPP line, multiplied by that household's upper national line, and divided by the average all-Egypt upper line.

The $2.50/day line and the $3.75/day line are multiples of the $1.25/day line.

[4] `http://www.cbe.org.eg/TimeSeries/prices/Consumer%20Price%20Index%20(Urban).xls` and `http://www.cbe.org.eg/TimeSeries/prices/Consumer%20Price%20Index%20(Urban)2.xls`, both accessed 21 December 2009.

3. Scorecard construction

For the Egypt scorecard, about 100 potential indicators are initially prepared in the areas of:

- Household composition (such as number of members)
- Education (such as school attendance by children)
- Employment (such as whether any household members have non-permanent wage jobs)
- Housing (such as the number of rooms)
- Ownership of durable goods (such as water heaters or washing machines)

Figure 4 lists all the candidate indicators, ranked by the entropy-based "uncertainty coefficient" that is a measure of how well the indicator predicts poverty on its own (Goodman and Kruskal, 1979). Responses for each indicator are ordered starting with those most strongly linked with higher poverty likelihoods.

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

The scorecard itself is built using the upper national poverty line and Logit regression on the construction sub-sample. Indicator selection uses both judgment and statistics (forward stepwise, based on "c"). The first step is to use Logit to build one scorecard for each candidate indicator. Each scorecard's accuracy is taken as "c", a measure of ability to rank by poverty status (SAS Institute Inc., 2004).

One of these one-indicator scorecards is then selected based on several factors (Schreiner *et al.*, 2004; Zeller, 2004), including improvement in accuracy, likelihood of acceptance by users (determined by simplicity, cost of collection, and "face validity" in terms of experience, theory, and common sense), sensitivity to changes in poverty status, variety among indicators, and verifiability.

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

This algorithm is the Logit analogue to the familiar R^2-based stepwise least-squares regression. It differs from naïve stepwise in that the criteria for selecting indicators include not only statistical accuracy but also judgment and non-statistical factors. The use of non-statistical criteria can improve robustness through time and helps ensure that indicators are simple and make sense to users.

The 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).

The single scorecard here applies to all of Egypt. Tests for Mexico and India (Schreiner, 2006a and 2006b), Sri Lanka (Narayan and Yoshida, 2005), and Jamaica (Grosh and Baker, 1995) suggests that segmenting poverty-assessment tools by

urban/rural does not improve targeting much, although such segmentation may improve the accuracy of estimated poverty rates (Tarozzi and Deaton, 2007).

4. Practical guidelines for scorecard use

The main challenge of scorecard design is not to squeeze out the last drops of accuracy but rather to improve the chances that scoring is actually used (Schreiner, 2005). When scoring projects fail, the reason is not usually technical inaccuracy but rather the failure of an organization to decide to do what is needed to integrate scoring in its processes and to learn to use it properly (Schreiner, 2002). After all, most reasonable scorecards predict tolerably well, thanks to the empirical phenomenon known as the "flat maximum" (Falkenstein, 2008; Hand, 2006; Baesens *et al.*, 2003; Lovie and Lovie, 1986; Kolesar and Showers, 1985; Stillwell, Barron, and Edwards, 1983; Dawes, 1979; Wainer, 1976; Myers and Forgy, 1963). The bottleneck is less technical and more human, not statistics but organizational-change management. Accuracy is easier to achieve than adoption.

The scorecard here is designed to encourage understanding and trust so that users will want to adopt it and use it properly. Of course, accuracy is important, but so are simplicity, ease-of-use, and "face validity". Programs are more likely to collect data, compute scores, and pay attention to the results if, in their view, scoring does not make a lot of "extra" work and if the whole process generally seems to make sense.

To this end, the scorecard here fits on a single page. The construction process, indicators, and points are simple and transparent. "Extra" work is minimized; non-specialists can compute scores by hand in the field because the scorecard has:

- Only 10 indicators
- Only categorical indicators
- Simple points (non-negative integers, and no arithmetic beyond addition)

The scorecard is ready to be photocopied. A field worker using the paper scorecard would:

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

4.1 Quality control

Of course, field workers must be trained and monitored. The quality of outputs depends on the quality of inputs. If organizations or field workers gather their own data and if they 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).[5] IRIS Center (2007a) and Toohig (2008) are useful nuts-and-bolts guides for planning,

[5] If an organization does not want field workers to know the points associated with indicators, then they can use the version of Figure 1 without points and apply the points later in a spreadsheet or database at the central office.

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 most alternatives, it is still absolutely difficult. Training and explicit definitions of the terms and concepts in the scorecard is essential. For the example of Nigeria, one study finds distressingly low inter-rater and test-retest correlations for indicators as seemingly simple and obvious as whether the household owns an automobile (Onwujekwe, Hanson, and Fox-Rushby, 2006).

As another example, Martinelli and Parker (2007) find that in the first stage of targeting in a Mexican conditional cash-transfer program, "underreporting [of asset ownership] is widespread but not overwhelming, except for a few goods . . . [and] overreporting is common for a few goods, which implies that self-reporting may lead to the exclusion of deserving households" (pp. 24–25). Still—as done in the second stage of Mexico's targeting process—field agents can verify responses with a home visit and correct false reports, and this same procedure is suggested for the scorecard as well.

4.2 Implementation and sampling

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

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

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

- Employees of the organization
- Third-party contractors

Responses, scores, and poverty likelihoods can be recorded:

- On paper in the field and then filed at an office
- On paper in the field and then keyed into a database or spreadsheet at an office
- On portable electronic devices in the field and then downloaded to a database

The subjects to be scored can be:

- All participants
- A representative sample of all participants
- All participants in a representative sample of branches
- A representative sample of all participants in a representative sample of branches
- A representative sample of participants in a particular group of interest

If not determined by other factors, the number of participants to be scored can be derived from sample-size formulas (presented later) for a desired confidence level and a desired confidence interval.

Frequency of application can be:

- At in-take of new clients only (precluding measuring change in poverty rates)
- As a once-off project for current participants (precluding measuring change)
- Once a year or at some other fixed time interval (allowing measuring change)
- Each time a field worker visits a participant at home (allowing measuring change)

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

- With different sets of participants, with each set representative of all participants
- With a single set of participants

An example set of choices for implementation and design is provided by BRAC and ASA, two microlenders in Bangladesh (each with more than 7 million participants) who are applying the scorecard (Schreiner, 2013). Their design is that loan officers in a random sample of branches apply the scorecard to their clients each time they visit a homestead (about once a year) as part of their standard due diligence prior to loan disbursement. Responses in the field are recorded on paper before being sent to a central office to be entered into a database. The sampling plans of ASA and BRAC cover 50,000–100,000 participants each (far more than would be required to inform most relevant decisions at a typical pro-poor organization).

5. Estimates of household poverty likelihoods

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

To get absolute units, scores must be converted to *poverty likelihoods*, that is, probabilities of being below a poverty line. This is done via simple look-up tables. For the example of the upper national line with the 2004/5 HIECS, scores of 40–44 have a poverty likelihood of 44.6 percent, and scores of 45–49 have a poverty likelihood of 37.1 percent (Figure 5).

The poverty likelihood associated with a score varies by poverty line. For example, scores of 40–44 are associated with a poverty likelihood of 44.6 percent for the upper national line but 14.3 percent for the lower national line.[6]

[6] Starting with Figure 5, many figures have seven versions, one for each of the seven poverty lines. The tables are grouped by poverty line. Single tables that pertain to all poverty lines are placed with the tables for the upper national line.

5.1 Calibrating scores with poverty likelihoods

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

For the example of the upper national line (Figure 6), there are 8,870 (normalized) households in the calibration sub-sample with a score of 40–44, of whom 3,955 (normalized) are below the poverty line. The estimated poverty likelihood associated with a score of 40–44 is then 44.6 percent, as 3,955 ÷ 8,870 = 44.6 percent.

To illustrate further with the upper national line and a score of 45–49, there are 9,452 (normalized) households in the calibration sample, of whom 3,505 (normalized) are below the line (Figure 6). Thus, the poverty likelihood for this score is 3,505 ÷ 9,452 = 37.1 percent.

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

Figure 7 shows, for all scores, the likelihood that consumption falls in a range demarcated by two adjacent poverty lines. For example, the daily consumption of someone with a score of 40–44 falls in the following ranges with probability:

- 0.3 percent below the $1.25/day 2005 PPP line
- 1.2 percent between the $1.25/day 2005 PPP and the food lines
- 12.9 percent between the food and the lower national lines
- 1.8 percent between the lower national and the USAID "extreme" lines
- 27.1 percent between the USAID "extreme" and the $2.50/day 2005 PPP lines
- 1.4 percent between the $2.50/day 2005 PPP and the upper national lines
- 42.7 percent between the upper national and $3.75/day 2005 PPP lines
- 12.7 percent above the $3.75/day 2005 PPP line

Even though the scorecard is constructed partly based on judgment, this calibration process produces poverty likelihoods that are objective, that is, derived from quantitative poverty lines and survey data on consumption. The poverty likelihoods would be objective even if indicators and/or points were selected without any data at all. In fact, scorecards with objective poverty likelihoods of proven accuracy are often constructed using only judgment (Fuller, 2006; Caire, 2004; Schreiner *et al.*, 2004). Of course, the scorecard here is constructed with both data and judgment. The fact that this paper acknowledges that some choices in scorecard construction—as in any statistical analysis—are informed by judgment in no way impugns the objectivity of the poverty likelihoods, as this depends on using data in score calibration, not on using data (and nothing else) in scorecard construction.

Although the points in Egypt's scorecard are transformed coefficients from a Logit regression, scores are not converted to poverty likelihoods via the Logit formula of $2.718281828^{\text{score}} \times (1 + 2.718281828^{\text{score}})^{-1}$. This is because the Logit formula is esoteric and

difficult to compute by hand. Non-specialists find it more intuitive to define the poverty likelihood as the share of households with a given score in the calibration sample who are below a poverty line. In the field, converting scores to poverty likelihoods requires no arithmetic at all, just a look-up table. This non-parametric calibration can also improve accuracy, especially with large calibration samples.

5.2 Accuracy of estimates of households' poverty likelihoods

As long as the relationship between indicators and poverty does not change and as long as the scorecard is applied to households who are representative of the same population from which the scorecard was constructed, this calibration process produces unbiased estimates of poverty likelihoods. *Unbiased* means that in repeated samples from the same population, the average estimate matches the true poverty likelihood. The scorecard also produces unbiased estimates of poverty rates at a point in time, as well as unbiased estimates of changes in poverty rates between two points in time.[7]

But the relationship between indicators and poverty does change with time and also across sub-groups in Egypt's population, so the scorecard will generally be biased when applied after the end date of fieldwork for the 2004/5 HIECS (as it must necessarily be applied in practice) or when applied with non-nationally representative groups (as it probably would be applied by local, pro-poor organizations).

[7] 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 representativeness? To check, the scorecard is applied to 1,000 bootstrap samples of size n = 16,384 from the 2004/5 validation sub-sample. Bootstrapping entails (Efron and Tibshirani, 1993):

- Score each household in the validation sample
- Draw a new bootstrap sample *with replacement* from the validation sample
- For each score, compute the true poverty likelihood in the bootstrap sample, that is, the share of households with the score who have consumption below a poverty line
- For each score, record the difference between the estimated poverty likelihood (Figure 5) and the true poverty likelihood in the bootstrap sample
- Repeat the previous three steps 1,000 times
- For each score, report the average difference between estimated and true poverty likelihoods across the 1,000 bootstrap samples
- For each score, report the two-sided interval containing the central 900, 950, or 990 differences between estimated and true poverty likelihoods

For each score range and for n = 16,384, Figure 8 shows the average difference between estimated and true poverty likelihoods as well as confidence intervals for the differences.

For the upper national line in the 2004/5 validation sample, the average poverty likelihood across bootstrap samples for scores of 40–44 is too low by 2.9 percentage points (Figure 8). For scores of 45–49, the estimate is too high by 2.8 percentage points.[8]

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

The 90-percent confidence interval for the differences for scores of 40–44 is ±2.6 percentage points (Figure 8). This means that in 900 of 1,000 bootstraps, the difference between the estimate and the true value is between –5.5 and –0.3 percentage points (because –2.9 – 2.6 = –5.5, and –2.9 + 2.6 = –0.3). In 950 of 1,000 bootstraps (95 percent), the difference is –2.9 ±2.9 percentage points, and in 990 of 1,000 bootstraps (99 percent), the difference is –2.9 ±3.6 percentage points.

For many scores, Figure 8 shows differences—some of them large—between estimated poverty likelihoods and true values. This is because the validation sub-sample is a single sample that—thanks to sampling variation—differs in distribution from the construction/calibration sub-samples and from Egypt's population. For targeting, however, what matters is less the differences across all score ranges and more the differences in score ranges just above and below the targeting cut-off. This mitigates the effects of bias and sampling variation on targeting (Friedman, 1997). Section 8 below looks at targeting accuracy in detail.

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

Another possible source of bias is overfitting. By construction, the scorecard here is unbiased, but it may still be *overfit* when applied after the June 2005 end of field work for the 2004/5 HIECS. That is, the scorecard may fit the data from the 2004/5 HIECS so closely that it captures not only real patterns but also some random patterns

that, due to sampling variation, show up only in the 2004/5 HIECS. Or the scorecard may be overfit in the sense that it is not robust to changes through time in the relationships between indicators and poverty. Finally, the scorecard could also be overfit when it is applied to samples from non-nationally representative sub-groups.

Overfitting can be mitigated by simplifying the scorecard and by not relying only on the 2004/5 HIECS data but rather also considering experience, judgment, and theory. Of course, the scorecard here does this. Bootstrapping scorecard construction—which is not done here—can also mitigate overfitting by reducing (but not eliminating) dependence on a single sampling instance. Combining scorecards can also help, at the cost of complexity.

In any case, most errors in individual households' likelihoods balance out in the estimates of groups' poverty rates (see later sections). Furthermore, much of the differences between scorecard estimates and true values may come from non-scorecard sources such as changes in the relationship between indicators and poverty, sampling variation, changes in poverty lines, inconsistencies in data quality across time, and inconsistencies/imperfections in cost-of-living adjustments across geography and time. These factors can be addressed only by improving data quantity and quality (which is beyond the scope of the scorecard), by updating data, or by reducing overfitting (which likely has limited returns, given the scorecard's parsimony).

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

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

To illustrate, suppose a program samples three households on Jan. 1, 2010 and that they have scores of 20, 30, and 40, corresponding to poverty likelihoods of 86.5, 65.2, and 44.6 percent (upper national line, Figure 5). The group's estimated poverty rate is the households' average poverty likelihood of $(86.5 + 65.2 + 44.6) \div 3 = 65.4$ percent.[9]

6.1 Accuracy of estimated poverty rates at a point in time

How accurate is this estimate? For a range of sample sizes, Figure 10 reports average differences between estimated and true poverty rates as well as precision (confidence intervals for the differences) for the Egypt scorecard applied to 1,000 bootstrap samples from the 2004/5 validation sample.

Summarizing Figure 10 across poverty lines and years for $n = 16{,}384$, Figure 9 shows that the absolute differences between the estimated poverty rate and the true rate for the 2004/5 scorecard applied to the 2004/5 validation sample are 1.2 percentage

[9] The group's poverty rate is *not* the poverty likelihood associated with the average score. Here, the average score is $(20 + 30 + 40) \div 3 = 30$, and the poverty likelihood associated with the average score is 65.2 percent. This is not the 65.4 percent found as the average of the three poverty likelihoods associated with each of the three scores.

points or less. The average absolute difference across the seven poverty lines is +0.4 percentage points.

In terms of precision, the 90-percent confidence interval for a group's estimated poverty rate at a point in time in 2004/5 with n = 16,384 and for all poverty lines is ±0.5 percentage points or less (Figure 9). This means that in 900 of 1,000 bootstraps of this size, the absolute difference between the estimate and the average estimate is 0.5 percentage points or less.

In the specific case of the upper national line, 90 percent of all samples of n = 16,384 produce estimates that differ from the true value in the range of +0.4 + 0.5 = +0.9 to +0.4 – 0.5 = –0.1 percentage points. This is because +0.4 is the average difference and ±0.5 is its 90-percent confidence interval. The average difference is +0.4 because the average scorecard estimate is too high by 0.4 percentage points; the scorecard tends to estimate a poverty rate of 31.9 percent for the 2004/5 validation sample, but the true value is 31.5 percent (Figure 2). Future accuracy will depend on how closely the period of application resembles 2004/5.

6.2 Standard-error formula for estimates of poverty rates at a point in time

How precise are the point-in-time estimates? Because they are averages, the estimates have a Normal distribution and can be characterized by their average difference vis-à-vis true values, along 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 for indirect measurement via poverty-assessment tools (Schreiner, 2008a), note that the textbook formula (Cochran, 1977) that relates confidence intervals with standard errors in the case of direct measurement of poverty rates is $c = +/- z \cdot \sigma$, where:

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

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

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

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

n is the sample size.

For example, with a sample $n = 16{,}384$, 90-percent confidence ($z = 1.64$), and a poverty rate p of 31.5 percent (the true rate in the 2004/5 validation sample for the upper national line in Figure 2), the confidence interval c is

$$+/- z \cdot \sqrt{\frac{p \cdot (1-p)}{n}} = +/- 1.64 \cdot \sqrt{\frac{0.315 \cdot (1-0.315)}{16{,}384}} = \pm 0.595 \text{ percentage points.}$$

The scorecard, however, does not measure poverty directly, so this formula is not applicable. To derive a formula for the Egypt scorecard, consider Figure 10, which reports empirical confidence intervals c for the differences for the scorecard applied to 1,000 bootstrap samples of various sample sizes from the validation sample. For $n =$

16,384, the upper national line, and the 2004/5 validation sub-sample, the 90-percent confidence interval is ±0.455 percentage points.[10] Thus, the ratio of confidence intervals with the scorecard and with direct measurement is 0.455 ÷ 0.595 = 0.76.

Now consider the same case, but with n = 8,192. The confidence interval under direct measurement is $+/-1.64 \cdot \sqrt{\frac{0.315 \cdot (1-0.315)}{8{,}192}} = \pm 0.842$ percentage points. The empirical confidence interval with the Egypt scorecard for the upper national line (Figure 10) is ±0.685 percentage points. Thus for n = 8,192, the ratio is 0.685 ÷ 0.842 = 0.81.

This ratio of 0.81 for n = 8,192 is close to the ratio of 0.76 for n = 16,384. Indeed, across all sample sizes of 256 or more in Figure 10, the average ratio turns out to be 0.81, implying that confidence intervals for indirect estimates of poverty rates via the Egypt scorecard and the upper national poverty line are about 19 percent narrower than those for direct estimates. This 0.81 appears in Figure 9 as the "α factor" because if $\alpha = 0.81$, then the formula relating confidence intervals c and standard errors σ for the Egypt scorecard is $c = +/- z \cdot \alpha \cdot \sigma$. The standard error σ for point-in-time estimates of poverty rates via scoring is $\alpha \cdot \sqrt{\frac{p \cdot (1-p)}{n}}$.

In general, α could 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 in all seven cases in Figure 9.

[10] Due to rounding, Figure 10 displays 0.5, not 0.455.

The formula relating confidence intervals to standard errors for the scorecard can be rearranged to give a formula for determining sample size n before measurement.[11] If $\hat{p}$ is the expected poverty rate before measurement, then the formula for n based on the desired confidence level that corresponds to z and the desired confidence interval $\pm c$ under the scorecard is $n = \left(\frac{\alpha \cdot z}{c}\right)^2 \cdot \hat{p} \cdot (1 - \hat{p})$.

To illustrate how to use this, suppose $c = 0.04005$ and $z = 1.64$ (90-percent confidence), and $\hat{p} = 0.314$ (the average poverty rate for the upper national line in the 2004/5 construction and calibration sub-samples, Figure 2). Then the formula gives $n = \left(\frac{0.81 \cdot 1.64}{0.04005}\right)^2 \cdot 0.314 \cdot (1 - 0.314) = 237$, close to the sample size of 256 observed for these parameters in Figure 10.

Of course, the α factors in Figure 9 are specific to Egypt, its poverty lines, its poverty rates, and this scorecard. The method for deriving the formulas, however, is valid for any poverty-assessment tool following the approach in this paper.

In practice after the end of the HIECS field work in June 2005, an organization would select a poverty line (say, the upper national line), select a desired confidence level (say, 90 percent, or $z = 1.64$), select a desired confidence interval (say, ±2.0

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

percentage points, or $c = 0.02$), make an assumption about $\hat{p}$ (perhaps based on a previous measurement such as the 31.4 percent average for the upper national line in the 2004/5 HIECS in Figure 2), look up α (here, 0.81), assume that the scorecard is still valid in the future and/or for non-nationally representative sub-groups,[12] and then compute the required sample size. In this illustration,

$$n = \left(\frac{0.81 \cdot 1.64}{0.02}\right)^2 \cdot 0.341 \cdot (1 - 0.341) = 992.$$

[12] This paper reports accuracy for the scorecard applied to the 2004/5 validation sample, but it cannot test accuracy for later years or other groups. Performance will deteriorate with time as the relationship between indicators and poverty changes.

7. Estimates of changes in group poverty rates over time

The change in a group's poverty rate between two points in time is estimated as the change in the average poverty likelihood of the households in the group. With data only from the 2004/5 HIECS, this paper cannot test estimates of change over time for Egypt, and it can only suggest approximate formulas for standard errors. Nevertheless, the relevant concepts are presented here because, in practice, pro-poor organizations can apply the scorecard to collect their own data and measure change through time.

7.1 Warning: Change is not impact

Scoring can estimate change. Of course, change could be for the better or for the worse, and scoring does not indicate what caused change. This point is often forgotten, confused, or ignored, 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 program participation on poverty status 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 program impact only if there is some way to know what would have happened in the absence of the program. And that information must come from somewhere beyond the scorecard.

7.2 Calculating estimated changes in poverty rates over time

Consider the illustration begun in the previous section. On Jan. 1, 2010, a program samples three households who score 20, 30, and 40 and so have poverty likelihoods of 86.5, 65.2, and 44.6 percent (upper national line, Figure 5). The group's baseline estimated poverty rate is the households' average poverty likelihood of (86.5 + 65.2 + 44.6) ÷ 3 = 65.4 percent.

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

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

By way of illustration, suppose that a year later on Jan. 1, 2011, the program samples three additional households who are in the same cohort as the three households originally sampled (or suppose that the program scores the same three original households a second time) and finds that their scores are now 25, 35, and 45 (poverty likelihoods of 76.8, 50.9, and 37.1 percent, upper national line, Figure 5). Their average poverty likelihood at follow-up is (76.8 + 50.9 + 37.1) ÷ 3 = 54.9 percent, an improvement of 65.4 – 54.9 = 10.5 percentage points.[13]

This suggests that about one of ten participants crossed the poverty line in 2010. (This is a net figure; some people start above the line and end below it, and vice versa.) Among those who started below the line, about one in six (10.5 ÷ 65.4 = 16.1 percent)

[13] Of course, such a huge reduction in poverty is unlikely in a year's time, but this is just an example to show how the scorecard can be used to estimate change.

ended up above the line. Of course, the scorecard does not reveal the reasons for this change.

7.3 Estimated changes in poverty rates in Egypt

With only the 2004/5 HIECS, it is not possible to measure the accuracy of scorecard estimates of changes in groups' poverty rates over time. In practice, of course, local pro-poor organizations can still apply the Egypt scorecard to estimate change. The rest of this section suggests approximate formulas for standard errors and sample sizes that may be used until there is additional data.

7.4 Accuracy for estimated change in two independent samples

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

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

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

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

intervals from the textbook formula for direct measurement for two equal-sized independent samples.

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

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

For countries for which this α has been measured (Schreiner, 2009a, 2009b, 2009c, 2009d, 2009e, and 2008b and Chen and Schreiner, 2009), the simple average of α across poverty lines, years, and countries is 1.11. This is as reasonable a figure as any to use for Egypt.

To illustrate the use of the formula above to determine sample size for estimating changes in poverty rates across two independent samples, suppose the desired confidence level is 90 percent ($z = 1.64$), the desired confidence interval is 2 percentage points ($c = 0.02$), the poverty line is the upper national line, $\alpha = 1.11$, and $\hat{p} = 0.341$ (from Figure 2). Then the baseline sample size is $n = 2 \cdot \left(\frac{1.11 \cdot 1.64}{0.02} \right)^2 \cdot 0.341 \cdot (1 - 0.341)$ = 3,724, and the follow-up sample size is also 3,724.

7.5 Accuracy for estimated change for one sample, scored twice

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

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

z, c, and α are defined as before, p_{12} is the share of all sampled households that move from below the poverty line to above it, and p_{21} is the share of all sampled households that move from above the line to below it.

As usual, the formula for σ can be rearranged to give a formula for sample size n before measurement. This requires an estimate (based on information available before measurement) of the expected shares of all households who cross the poverty line $\hat{p}_{12}$ and $\hat{p}_{21}$. Before measurement, it is reasonable to assume that the overall change in the poverty rate will be zero, which implies $\hat{p}_{12} = \hat{p}_{21} = \hat{p}_*$, giving:

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

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

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

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

Given this, a sample-size formula for a group of households to whom the Egypt scorecard is applied twice (once after the end of field work for the 2004/5 HIECS and then again later) is:

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

In Peru (the only other country for which there is a data-based estimate, Schreiner 2009a), the average α across years and poverty lines is about 1.30.

To illustrate the use of this formula, suppose the desired confidence level is 90 percent ($z = 1.64$), the desired confidence interval is 2.0 percentage points ($c = 0.02$), the poverty line is the upper national line, and the sample will be scored first in 2010 and then again in 2013 ($y = 3$). The before-baseline poverty rate is 34.1 percent ($p_{2004/5} = 0.341$, Figure 2), and suppose $\alpha = 1.30$. Then the baseline sample size is $n = 2 \cdot \left(\frac{1.30 \cdot 1.64}{0.02} \right)^2 \cdot \{ -0.02 + 0.016 \cdot 3 + 0.47 \cdot [0.341 \cdot (1 - 0.341)] \} = 3{,}037$. The same group of 3,037 households is scored at follow-up as well.

8. Targeting

When a program 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 depends on whether consumption is below a poverty line as directly measured by a survey. In contrast, targeting status is a program's policy choice that depends on a cut-off and on an indirect estimate from a scorecard.

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

A program 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 12 shows the distribution of households by targeting outcome. For an example cut-off of 44 or less and the 2004/5 scorecard applied to the 2004/5 validation sample, outcomes for the upper national line are:

- Inclusion: 23.1 percent are below the line and correctly targeted
- Undercoverage: 8.4 percent are below the line and mistakenly not targeted
- Leakage: 13.0 percent are above the line and mistakenly targeted
- Exclusion: 55.6 percent are above the line and correctly not targeted

Increasing the cut-off to 49 or less improves inclusion and undercoverage but worsens leakage and exclusion:

- Inclusion: 26.3 percent are below the line and correctly targeted
- Undercoverage: 5.2 percent are below the line and mistakenly not targeted
- Leakage: 19.2 percent are above the line and mistakenly targeted
- Exclusion: 49.4 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 12 for a given poverty line
- Select the cut-off with the highest total net benefit

The most difficult step is assigning benefits and costs to targeting outcomes. Any program that uses targeting—with or without scoring—should thoughtfully consider

how it values successful inclusion or exclusion versus errors of undercoverage and leakage. It is healthy to go through a process of thinking explicitly and intentionally about how possible targeting outcomes are valued.

A common choice of benefits and costs is "Total Accuracy" (IRIS Center, 2005; Grootaert and Braithwaite, 1998). With "Total Accuracy", total net benefit is the number of households correctly included or correctly excluded:

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

Figure 12 shows "Total Accuracy" for all cut-offs for Egypt's scorecard. For the upper national line in the 2004/5 validation sample, total net benefit is greatest (79.2) for a cut-off of 44 or less, with about four in five Egyptian households correctly classified.

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

[16] Figure 12 also reports the Balanced Poverty Accuracy Criteria used by USAID as its criterion for certifying poverty-assessment tools. IRIS Center (2005) says that BPAC considers accuracy in terms of the estimated poverty rate and in terms of inclusion. BPAC is (Inclusion – |Undercoverage – Leakage|) x [100 ÷ (Inclusion+Undercoverage)].

As an alternative to assigning benefits and costs to targeting outcomes and then choosing a cut-off to maximize total net benefit, a program could set a cut-off to achieve a desired poverty rate among targeted households. The third column of Figure 13 ("% targeted who are poor") shows the expected poverty rate among Egypt households who score at or below a given cut-off. For the example of the upper national line and the 2004/5 validation sample, targeting households who score 44 or less would target 36.0 percent of all households (second column) and lead to a poverty rate among those targeted of 64.0 percent (third column).

Figure 13 also reports two other measures of targeting accuracy. The first is a version of inclusion ("% of poor who are targeted"). For the example of the upper national line and the 2004/5 validation sample with a cut-off of 44 or less, 73.3 percent of all poor households are covered.

The final targeting measure in Figure 13 is the number of successfully targeted poor households for each non-poor household mistakenly targeted (right-most column). For the upper national line, the 2004/5 validation sample, and a cut-off of 44 or less, covering 1.8 poor households means leaking to 1 non-poor household.

9. Context of poverty-assessment tools for Egypt

This section discusses four existing poverty-assessment tools for Egypt in terms of their goals, data, methods, poverty lines, definition of poverty, cost, indicators, accuracy, and precision. The relative strengths of the scorecard are that it uses the latest nationally representative data, it uses simpler and fewer indicators, it costs less to implement, it reports accuracy and precision out-of-sample, and it provides formulas for standard errors.

9.1 Gwatkin *et al.*

Gwatkin *et al.* (2007) apply to Egypt an approach used in 56 countries with Demographic and Health Surveys (Rutstein and Johnson, 2004). Principal Components Analysis is used to make an asset index from simple, low-cost indicators available for the 16,957 households in Egypt's 2000 DHS.[17] The PCA index is like the scorecard here except that, because the DHS does not collect data on income nor consumption, it defines poverty based on the indicators in the index.[18] Well-known examples of the PCA

[17] El-Kogali and El-Daw (2001) make a similar index for Egypt using the 1995 DHS.

[18] Still, because the indicators are similar and because the "flat maximum" is important, carefully built PCA indices and consumption-based poverty-assessment tools seem to pick up the same underlying construct (perhaps "permanent income", see Bollen, Glanville, and Stecklov, 2007), and they rank households much the same. Tests of how well rankings by PCA indices correspond with rankings by consumption are Filmer and Scott (2008), Lindelow (2006), Wagstaff and Watanabe (2003), and Montgomery *et al.* (2000).

asset-index approach include Stifel and Christiaensen (2007), Zeller *et al.* (2006), Sahn and Stifle (2000 and 2003), and Filmer and Pritchett (2001).

The 23 indicators in Gwatkin *et al.* are similar to those in the new scorecard here in terms of their simplicity, low cost, and verifiability:

- Characteristics of the residence:
  - — Type of residence
  - — Tenancy status
  - — Type of floor
  - — Electrical connection
  - — Source of drinking water
  - — Type of toilet arrangement
  - — Type of cooking fuel
- Ownership of consumer durables:
  - — Radios
  - — Televisions
  - — Video players
  - — Fans
  - — Refrigerators
  - — Water heaters
  - — Telephones
  - — Bicycles
  - — Motorcycles
  - — Cars or trucks
  - — Sewing machines
  - — Automatic clothes-washing machines
  - — Non-automatic clothes-washing machines
- Ownership of agricultural assets:
  - — Farm or other land
  - — Livestock
- Number of people per sleeping room

Gwatkin *et al.* discuss three basic uses for their index:

- Segmenting households by quintiles to see how health, population, and nutrition vary with socio-economic status
- Monitoring (via exit surveys) how well local health-service agencies reach the poor
- Measuring coverage of health services via local, small-scale surveys

The first goal is akin to targeting, and the last two goals resemble the monitoring goals here, so the uses of the index are about the same as those of the scorecard here.

Still, the Gwatkin *et al.* index is more difficult and costly: it has 23 indicators, it cannot easily fit on a single page, and it cannot be computed by hand in the field, as it has 123 point values, half of them negative, and all with five decimal places.

Finally, the scorecard here—unlike the PCA index—is linked directly to an absolute, consumption-based poverty line. Thus, while both approaches can rank households, only the scorecard can estimate consumption-based poverty status.

9.2 Datt and Jolliffe

Datt and Jolliffe (2005) construct a poverty-assessment tool to use in simulating effects of policy changes on poverty. Given this government-level policy focus, they include only indicators whose values are determined before the time period covered by the survey. They do not evaluate their tool in terms of targeting nor estimating poverty rates.

Datt and Jolliffe use nationally representative data on 2,500 households in the 1997 Egypt Integrated Household Survey (EIHS) by the International Food Policy Research Institute and Egypt's Ministry of Trade and Supply. As here, they construct regional, household-specific poverty lines. They build two tools (urban and rural) using least-squares regression on the logarithm of per-capita consumption. The tools use an average of eleven of the following indicators:

- Household demographics:
  - — Number of members (and its square)
  - — Age of the head (and its square)
  - — Whether the head is female
- Education:
  - — Average years of schooling among household members
  - — Whether the parent of the head completed primary school
  - — Whether the parent of the head's spouse completed primary school
- Area of cultivated land owned
- Location:
  - — Whether in Upper Egypt
  - — Governorate
  - — Distance (measured in time on foot) to nearest:
    - Secondary school
    - Hospital or health post

In addition, the tools include an average of six combination terms (such as whether the parent of the head's spouse completed primary school, combined with the distance to the nearest hospital or health post).

While Datt and Jolliffe's poverty-assessment tool is appropriate for its purpose of simulating the effects of counterfactual policies—and while it is an improvement on bivariate "poverty profiles"—its complexity and restricted set of indicators make it infeasible for local, pro-poor organizations who want a tool for targeting or estimating their participants' poverty rates.

9.3 Ahmed and Bouis

Like Datt and Jolliffe, Ahmed and Bouis (2002) use the 1997 EIHS. Their goal, however, is closer to the goal of this paper: they seek to provide a feasible tool to improve the targeting of ration cards for sugar and cooking oil. To this end—and again like the approach in this paper—tool construction is informed by feedback from the government agency for whom the tool is intended. In particular, indicators are selected based on their ease of collection, verifiability, and political acceptability.

Ahmed and Bouis' initial poverty-assessment tool (constructed from a regression on the logarithm of per-capita consumption) has 56 indicators. Users, however, asked them to drop all indicators requiring staff calculations (such as the dependency ratio) or judgment (such as whether a household is urban or rural), lest the tool be too complex for existing personnel to apply to all households in Egypt. Furthermore, the use of the governorate of residence was dropped due to its political sensitivity. The final tool has nine indicators:

- Number of household members
- Education:
  - — Highest level completed by any employed household member
  - — Whether any children attend a private school
  - — Whether all household member aged 15 or older have attended school
- Type of toilet arrangement
- Amount paid for utilities in EGP:[19]
  - — Electricity
  - — Telephone

[19] Ahmed and Bouis say that these amounts are simple to verify using bill stubs.

- Ownership of consumer durables:
  - — Motor vehicle
  - — Refrigerator

To evaluate targeting accuracy, Ahmed and Bouis apply the tool with the same households in the 1997 EIHS that are used to construct the tool. Such *in-sample* testing is known to overstate accuracy (Johannsen, 2006; Copestake *et al.*, 2005). A more accurate approach to measuring accuracy is *out-of-sample* testing that uses different data than in tool construction. This paper tests out-of-sample. For accuracy comparisons with Ahmed and Bouis, this puts the scorecard here at a disadvantage, as does the fact that it is built using household-level weights but—to be compared with Ahmed and Bouis—must be applied with person-level weights.

For a cut-off that targets 36.5 percent of people, Ahmed and Bouis report inclusion of 26.2 percent and exclusion of 10.4 percent. For the new scorecard here, this same cut-off leads to inclusion of 23.7 percent and exclusion of 10.2 percent. Given sampling variation and Ahmed and Bouis' use of an in-sample test, it seems fair to say that the two tools target about equally well.

Still, the new scorecard here is simpler and costs less to apply. For example, Ahmed and Bouis still require field agents to perform multiplication. While weights are

rounded off to whole numbers, their tool still requires subtraction because it has negative points. Finally, their score is not normalized to a range of 0 to 100.[20]

[20] The approach to the scorecard in this paper is nevertheless similar in some ways to that in Ahmed and Bouis, although the approach here was developed before the author saw their paper. What is not new in the approach here was learned from Dean Caire.

9.4 World Bank

World Bank (2007) uses the 2004/5 HIECS and the 2005 Population Census to construct region-specific poverty-assessment tools that feed into a "poverty map" (Elbers, Lanjouw, and Lanjouw, 2003; Hentschel *et al.*, 2000) that estimates poverty rates at the sub-district level (rural villages and urban *shayakha*). The poverty-mapping estimates are more precise for these small areas than would be direct estimates based on the HIECS. World Bank also creates a poverty map using 1996 data and uses the two maps to track changes over time.

Poverty-assessment tools are built using stepwise generalized least-squares on the logarithm of per-capita consumption, using only indicators found both in the 2004/5 HIECS and the 2005 Census. The tools then are applied to households in the 2005 Census and used to make poverty maps that quickly show—in a way that is clear for non-specialists—how poverty rates vary across sub-districts. World Bank notes that the poverty map is preliminary, as some household-level data from the census had not yet been released.

Poverty mapping in World Bank and the scorecard in this paper are similar in that they both:

- Build poverty-assessment tools with nationally representative survey data and then apply them to other data on sub-groups that may not be nationally representative
- Use simple, verifiable indicators that are quick and inexpensive to collect
- Provide unbiased estimates when their assumptions hold
- Are used to estimate poverty rates for groups
- Seek to be useful in practice and so aim to be understood by non-specialists

Strengths of poverty mapping include that it:

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

Strengths of the scorecard include that it:

- Is simpler in terms of both construction and application
- Tests accuracy empirically
- Associates poverty likelihoods with scores non-parametrically
- Uses judgment and theory in scorecard construction to reduce overfitting
- Estimates poverty likelihoods for individual households
- Reports straightforward formulas for standard errors

The most basic difference between the two approaches is that poverty mapping seeks to help governments design and target pro-poor policies, while the scorecard seeks to help small, local pro-poor organizations to manage their outreach when implementing policies.[21]

[21] Another apparent difference is that the developers of the poverty-mapping approach (Elbers, Lanjouw, and Lanjouw, 2003; Demombynes *et al.*, 2002) say that it is too inaccurate to be used for targeting individual households, while Schreiner (2008c) supports such targeting as a legitimate, potentially useful application of the scorecard. Recently, the developers of poverty mapping seem to have taken a small step away from their original position (Elbers *et al.*, 2007).

World Bank uses the following indicators for Egypt:

- Demographics of the household:
  - Number of members (and its square, logarithm, and square of logarithm)
  - Age of head
  - Sex of head
- Education of head
- Employment of head:
  - Status
  - Sector
  - Activity
- Characteristics of the residence:
  - Electrical connection
  - Type of kitchen
  - Source of water
  - Type of bathroom
- Characteristics of the sub-district (share of population):
  - Female adults
  - Male adults
  - Children
  - Children ages 6 and younger
  - Employed persons
  - Employers in agriculture
  - Employers in non-agriculture
  - Self-employed in non-agriculture
  - Self-employed in agriculture
  - Unpaid in agriculture
  - Wage workers in agriculture
  - Wage workers in non-agriculture
  - Employed in private sector
  - Employed by government
  - Out of labor force
  - Unemployed
  - College graduates
  - Illiterates
- Other characteristics of the sub-district:
  - Presence of tap water
  - Presence of sewer system
  - Average years of education
- Geopolitical identity

The average regional tool uses about 18 of these 33 indicators and is based on about 2,100 households. Although a central feature of the poverty-mapping approach is the ability to generate standard errors, World Bank does not report standard errors, so the precision of its estimates cannot be compared with those in this paper.

10. Conclusion

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

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

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

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

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

When the scorecard is applied to the 2004/5 validation sample with $n = 16{,}384$, the absolute difference between estimates and true poverty rates at a point in time is 1.2 percentage points or less and averages—across the seven poverty lines—0.4 percentage points. With 90-percent confidence, the precision of these differences is ± 0.5 percentage points or less. The scorecard is more precise than direct measurement.

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

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

In sum, the scorecard is a practical, objective way for pro-poor programs in Egypt to measure poverty rates, track changes in poverty rates over time, and target services, provided that the scorecard is applied during a period similar to that of 2004/5, the time period when the data used to construct the scorecard was collected. The same approach can be applied to any country with similar data from a national income or consumption survey.

References

Adams, Niall M.; and David J. Hand. (2000) "Improving the Practice of Classifier Performance Assessment", *Neural Computation*, Vol. 12, pp. 305–311.

Ahmed, Akhter U.; and Howarth E. Bouis. (2002) "Weighing What Is Practical: Proxy Means Tests for Targeting Food Subsidies in Egypt", *Food Policy*, Vol. 27, No. 5–6, pp. 519–540.

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

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

Caire, Dean. (2004) "Building Credit Scorecards for Small Business Lending in Developing Markets", `microfinance.com/English/Papers/Scoring_SMEs_Hybrid.pdf`, retrieved 27 April 2010.

Chen, Shiyuan; and Mark Schreiner. (2009b) "Simple Poverty Scorecard Poverty-Assessment Tool: Vietnam", `SimplePovertyScorecard.com/VNM_2006_ENG.pdf`, retrieved 8 July 2016.

Coady, David; Grosh, Margaret; and John Hoddinott. (2002) "The Targeting of Transfers in Developing Countries", `hdl.handle.net/10986/14902`, retrieved 3 November 2015.

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

Copestake, James G.; Dawson, Peter; Fanning, John-Paul; McKay, Andrew; and Katie Wright-Revolledo. (2005) "Monitoring the Diversity of the Poverty Outreach and Impact of Microfinance: A Comparison of Methods Using Data from Peru", *Development Policy Review*, Vol. 23, No. 6, pp. 703–723.

Daley-Harris, Sam. (2009) *State of the Microcredit Summit Campaign Report 2009*, `microcreditsummit.org/state_of_the_campaign_report/`, retrieved 27 April 2010.

Datt, Gaurav; and Dean Jolliffe. (2005) "Poverty in Egypt: Modeling and Policy Simulations", *Economic Development and Cultural Change*, Vol. 53, No. 2, pp. 327–346.

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. (2002) "Producing an Improved Geographic Profile of Poverty: Methodology and Evidence from Three Developing Countries", World Institute for Development Economics Research Discussion Paper No. 2002/39, `go.worldbank.org/UMQCZ1BW00`, retrieved 27 April 2010.

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

El-Kogali, Safaa; and Suliman El-Daw. (2001) "Poverty, Human Capital, and Gender: A Comparative Study of Yemen and Egypt", Economic Research Forum Working Paper No. 0123, `www.erf.org.eg/CMS/getFile.php?id=86`, retrieved 27 April 2010.

Elbers, Chris; Lanjouw, Jean O.; and Peter Lanjouw. (2003) "Micro-Level Estimation of Poverty and Inequality", *Econometrica*, Vol. 71, No. 1, pp. 355–364.

_____.; 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.

Falkenstein, Eric. (2008) "DefProb™: A Corporate Probability of Default Model", `papers.ssrn.com/sol3/papers.cfm?abstract_id=1103404`, retrieved 27 April 2010.

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

_____.; and Kinnon Scott. (2008) "Assessing Asset Indices", World Bank Policy Research Working Paper No. 4605, `papers.ssrn.com/sol3/papers.cfm?abstract_id=1149108`, retrieved 27 April 2010.

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 27 April 2010.

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

Grootaert, Christiaan; and Jeanine Braithwaite. (1998) "Poverty Correlates and Indicator-Based Targeting in Eastern Europe and the Former Soviet Union", World Bank Policy Research Working Paper No. 1942, go.worldbank.org/VPMWVLU8E0, retrieved 27 April 2010.

Grosh, Margaret; and Judy L. Baker. (1995) "Proxy Means Tests for Targeting Social Programs: Simulations and Speculation", World Bank LSMS Working Paper No. 118, go.worldbank.org/W9OWN57PD0, retrieved 27 April 2010.

Gwatkin, Davidson R.; Rutstein, Shea; Johnson, Kiersten; Suliman, Eldaw; Wagstaff, Adam; and Agbessi Amouzou. (2007) "Socio-Economic Differences in Health, Nutrition, and Population: Egypt", World Bank Country Reports on HNP and Poverty, go.worldbank.org/T6LCN5A340, retrieved 27 April 2010.

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

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

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

IRIS Center. (2007a) "Manual for the Implementation of USAID Poverty Assessment Tools", povertytools.org/training_documents/Manuals/USAID_PAT_Manual_Eng.pdf, retrieved 27 April 2010.

_____. (2007b) "Introduction to Sampling for the Implementation of PATs", povertytools.org/training_documents/Sampling/Introduction_Sampling.ppt, retrieved 27 April 2010.

_____. (2005) "Notes on Assessment and Improvement of Tool Accuracy", povertytools.org/other_documents/AssessingImproving_Accuracy.pdf, retrieved 27 April 2010.

Johannsen, Julia. (2006) "Operational Poverty Targeting in Peru—Proxy Means Testing with Non-Income Indicators", IPC Working Paper No. 30, undp-povertycentre.org/pub/IPCWorkingPaper30.pdf, retrieved 27 April 2010.

Johnson, Glenn. (2007) "Lesson 3: Two-Way Tables—Dependent Samples", www.stat.psu.edu/online/development/stat504/03_2way/53_2way_compare.htm, retrieved 27 April 2010.

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

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

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

Martinelli, César; and Susan W. Parker. (2007) "Deception and Misreporting in a Social Program", ciep.itam.mx/~martinel/lies4.pdf, retrieved 27 April 2010.

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 27 April 2010.

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

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

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, siteresources.worldbank.org/EXTSAREGTOPPOVRED/Resources/493440-1102216396155/572861-1102221461685/Proxy+Means+Test+for+Targeting+Welfare+Benefits.pdf, retrieved 27 April 2010.

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

Ravallion, Martin. (1994) *Poverty Comparisons*.

Rutstein, Shea Oscar; and Kiersten Johnson. (2004) "The DHS Wealth Index", DHS Comparative Reports No. 6, www.measuredhs.com/pubs/pdf/CR6/CR6.pdf, retrieved 27 April 2010.

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

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

SAS Institute Inc. (2004) "The LOGISTIC Procedure: Rank Correlation of Observed Responses and Predicted Probabilities", in *SAS/STAT User's Guide, Version 9*, support.sas.com/documentation/cdl/en/statug/63033/HTML/default/statug_logistic_sect035.htm, retrieved 27 April 2010.

Schreiner, Mark. (2013) "Simple Poverty Scorecard Poverty-Assessment Tool: Bangladesh", SimplePovertyScorecard.com/BGD_2010_ENG.pdf, retrieved 8 July 2016.

Schreiner, Mark. (2009a) "Simple Poverty Scorecard Poverty-Assessment Tool: Peru", SimplePovertyScorecard.com/PER_2007_ENG.pdf, retrieved 8 July 2016.

_____. (2009b) "Simple Poverty Scorecard Poverty-Assessment Tool: Philippines", SimplePovertyScorecard.com/PHL_2002_ENG.pdf, retrieved 8 July 2016.

_____. (2009c) "Simple Poverty Scorecard Poverty-Assessment Tool: Pakistan", SimplePovertyScorecard.com/PAK_2005_ENG.pdf, retrieved 8 July 2016.

_____. (2009d) "Simple Poverty Scorecard Poverty-Assessment Tool: Bolivia", SimplePovertyScorecard.com/BOL_2007_ENG.pdf, retrieved 8 July 2016.

_____. (2009e) "Simple Poverty Scorecard Poverty-Assessment Tool: Mexico", SimplePovertyScorecard.com/MEX_2008_ENG.pdf, retrieved 8 July 2016.

_____. (2008a) "Simple Poverty Scorecard Poverty-Assessment Tool: Peru", SimplePovertyScorecard.com/PER_2003_ENG.pdf, retrieved 8 July 2016.

_____. (2008b) "Simple Poverty Scorecard Poverty-Assessment Tool: India", SimplePovertyScorecard.com/IND_2005_ENG.pdf, retrieved 8 July 2016.

_____. (2008c) "Simple Poverty Scorecard Poverty-Assessment Tool: Ecuador", SimplePovertyScorecard.com/ECU_2005_ENG.pdf, retrieved 8 July 2016.

_____. (2006a) "La Herramienta del Índice de Calificación de la Pobreza™: México", SimplePovertyScorecard.com/MEX_2002_SPA.pdf, retrieved 27 April 2010.

_____. (2006b) "Is One Simple Poverty Scorecard Poverty-Assessment Tool Enough for India?", microfinance.com/English/Papers/Scoring_Poverty_India_Segments.pdf, retrieved 27 April 2010.

_____. (2005) "IRIS Questions on the Simple Poverty Scorecard Poverty-Assessment Tool", microfinance.com/English/Papers/Scoring_Poverty_Response_to_IRIS.pdf, retrieved 27 April 2010.

_____. (2002) *Scoring: The Next Breakthrough in Microfinance?* CGAP Occasional Paper No. 7, pdf.usaid.gov/pdf_docs/PNACQ633.pdf, retrieved 27 April 2010.

_____; 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 27 April 2010.

Sillers, Don. (2006) "National and International Poverty Lines: An Overview", pdf.usaid.gov/pdf_docs/Pnadh069.pdf, retrieved 31 May 2012.

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

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

Tarozzi, Alessandro; and Angus Deaton. (2007) "Using Census and Survey Data to Estimate Poverty and Inequality for Small Areas", princeton.edu/~deaton/downloads/20080301SmallAreas_FINAL.pdf, retrieved 27 April 2010.

Toohig, Jeff. (2008) "PPI Pilot Training Guide", progressoutofpoverty.org/toolkit, retrieved 27 April 2010.

United States Congress. (2004) "Microenterprise Results and Accountability Act of 2004 (HR 3818 RDS)", November 20, smith4nj.com/laws/108-484.pdf, retrieved 11 January 2017.

Wagstaff, Adam; and Naoko Watanabe. (2003) "What Difference Does the Choice of SES Make in Health Inequality Measurement?", *Health Economics*, Vol. 12, No. 10, pp. 885–890.

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

World Bank. (2008) "Tables of Results", siteresources.worldbank.org/ICPINT/Resources/icp-final-tables.pdf, retrieved 27 April 2010.

_____. (2007) *Arab Republic of Egypt Poverty Assessment Update, Volume I: Main Report*, and *Volume II: Annexes*, Report No. 39885–EGT, www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2008/02/04/000020953_20080204085105/Original/398850v101OFFI10Update10Main0Report.doc, retrieved 27 April 2010.

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

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

Figure 2: Sample sizes and household poverty rates by sub-sample and poverty line

		% with expenditure below a poverty line						
		National			USAID	International 2005 PPP		
Sub-sample	**Households**	**Upper**	**Lower**	**Food**	**'Extreme'**	**$1.25/day**	**$2.50/day**	**$3.75/day**
All Egypt	11,774	31.4	13.9	2.5	14.7	1.4	29.9	62.4
Construction								
Selecting indicators and points	3,931	31.4	14.0	2.6	14.6	1.5	29.8	62.7
Calibration								
Associating scores with likelihoods	3,857	31.4	13.9	2.7	14.6	1.5	29.9	62.6
Validation								
Measuring accuracy	3,986	31.5	13.9	2.4	14.9	1.3	30.1	62.0
Change in poverty rate (percentage points)								
From construction/calibration to validation		–0.1	+0.1	+0.2	–0.3	+0.3	–0.2	+0.7

Source: 2004/5 HIECS

Figure 3: Poverty lines and poverty rates, by region and for all Egypt, at household- and person-level

		Average poverty line (EGP/person/day) and poverty rate (%)						
		National				International 2005 PPP		
Region	Item	Upper	Lower	Food	'Extreme'	$1.25/day	$2.50/day	$3.75/day
All Egypt	Poverty line	5.08	3.90	2.73	3.98	2.49	4.98	7.47
	Household-level poverty rate	31.4	13.9	2.5	14.7	1.4	29.9	62.4
	Person-level poverty rate	40.3	19.4	3.9	20.4	2.3	38.7	71.8
Urban Egypt	Poverty line	5.21	3.92	2.74	4.16	2.55	5.10	7.65
	Household-level poverty rate	19.8	7.2	1.0	9.5	0.7	18.9	45.6
	Person-level poverty rate	26.5	10.4	1.5	13.4	1.1	25.5	55.3
Rural Egypt	Poverty line	4.99	3.89	2.72	3.84	2.44	4.88	7.33
	Household-level poverty rate	41.6	19.8	3.9	19.2	2.1	39.6	77.2
	Person-level poverty rate	51.0	26.4	5.7	25.9	3.2	48.8	84.5
Metropolitan	Poverty line	5.27	3.98	2.81	4.33	2.58	5.16	7.74
	Household-level poverty rate	13.5	4.5	0.6	6.9	0.3	12.9	34.9
	Person-level poverty rate	18.6	6.4	0.7	9.5	0.4	17.9	43.9
Lower Urban	Poverty line	5.07	3.85	2.67	4.22	2.49	4.97	7.46
	Household-level poverty rate	21.7	6.4	1.0	10.5	0.6	20.6	53.4
	Person-level poverty rate	27.4	8.7	1.2	13.8	0.6	26.1	62.0
Lower Rural	Poverty line	5.26	3.89	2.70	3.78	2.57	5.15	7.72
	Household-level poverty rate	29.4	13.2	1.9	13.6	1.5	28.5	56.5
	Person-level poverty rate	39.2	19.1	3.2	19.5	2.6	38.0	67.1
Upper Urban	Poverty line	5.02	3.93	2.72	4.10	2.46	4.91	7.37
	Household-level poverty rate	32.0	11.8	0.9	15.3	0.3	29.9	72.7
	Person-level poverty rate	40.1	16.1	1.4	20.6	0.5	37.6	80.1
Lower Urban	Poverty line	4.95	3.85	2.72	3.54	2.42	4.85	7.27
	Household-level poverty rate	54.2	30.3	7.7	24.4	4.2	52.3	83.4
	Person-level poverty rate	63.9	38.5	10.8	32.2	6.2	62.0	89.9
Border urban	Poverty line	5.21	3.84	2.68	4.56	2.55	5.10	7.65
	Household-level poverty rate	7.9	0.0	0.0	4.8	0.0	6.8	34.3
	Person-level poverty rate	10.0	0.0	0.0	6.5	0.0	9.2	42.8
Border rural	Poverty line	5.14	3.73	2.70	3.25	2.52	5.03	7.55
	Household-level poverty rate	43.0	24.3	8.3	16.5	8.3	40.9	63.7
	Person-level poverty rate	55.0	33.3	12.0	26.6	13.0	52.4	75.7

Figure 4: Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly indicative of poverty)
1501	How many members does the household have? (Seven or more; Six; Five; Four; Three; One or two)
1086	What is the highest educational level that the female head/spouse has completed? (Illiterate; No educational certificate, but can read and write; Less than intermediate certificate; No female head/spouse; Intermediate certificate; More than intermediate but less than university; College or graduate degree (masters or Ph.D.))
1022	Does the household own any gas or electric water heaters? (No; Yes)
939	What is the highest educational level that the male head/spouse has completed? (Illiterate; No educational certificate, but can read and write; No male head/spouse; Less than intermediate certificate; Intermediate certificate; More than intermediate but less than university; College or graduate degree (masters or Ph.D.))
905	What type of residence does the household have? (Countryside house or non-traditional house; Separate room inside a residence with a shared entrance; Separate room with a private entrance; One or more apartments/flats, or a villa/detached house)
901	Can the female head/spouse read and write? (No; No female head/spouse; Yes)
879	What toilet arrangement does the household have? (No toilet available, or shared toilet; Private non-flush toilet; Private flush toilet)
861	Do all children ages 6 to 18 attend school? (No; Yes; No children 6 to 18)
830	How many members ages 18 or younger does the household have? (Four or more; Three; Two; One; None)
792	Does the household own any land-line and/or cellular telephones? (None; Cellular but not land-line, or land-line but not cellular; Both land-line and cellular;)
759	Do all children ages 6 to 17 attend school? (No; Yes; No children 6 to 17)
731	Does the household own any clothes-washing machines? (No; Yes, only non-automatic; Yes, automatic)
708	How many members ages 17 or younger does the household have? (Four or more; Three; Two; One; None)
687	Do all children ages 6 to 16 attend school? (No; Yes; No children 6 to 16)

Figure 4 (cont.): Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly associated with poverty)
649	Does the household own any automatic clothes-washing machines? (No; Yes)
643	How does the household dispose of its rubbish? (Other; In the street; Public garbage bins; Sweeper; Cleaning company)
640	How many members ages 16 or younger does the household have? (Four or more; Three; Two; One; None)
638	How many household members were employed in the past week? (Four or more; Three; Two; One; None)
604	Can the male head/spouse read and write? (No; Yes; No male head/spouse)
580	What is the occupation of the male head/spouse in his main job? (Skilled agricultural and fishery workers; Service workers and shop/market sales workers, and elementary occupations; Craft and related trade workers; Plant and machine operators and assemblers; Does not work; No male head/spouse, legislators, senior officials, managers, technicians and associate professionals; Clerks and professionals)
576	Does the household own any vacuum cleaners? (No; Yes)
572	Do all children ages 6 to 15 attend school? (No; Yes; No children 6 to 15)
569	In their main employment activity, how many household members are in agricultural, forestry, or fishing? (Two or more; One; None)
561	In their main occupation, are many household members skilled agricultural or fishery workers? (Yes; No)
548	How many members ages 15 or younger does the household have? (Four or more; Three; Two; One; None)
535	Does the household own any cellular telephones? (No; Yes)
518	What is the occupation of the female head/spouse in her main job? (Skilled agricultural and fishery workers; Service workers and shop/market sales workers, craft and related trade workers, or elementary occupations; Plant and machine operators and assemblers, or does not work; No female head/spouse; Legislators, senior officials, managers, professionals, technicians and associate professionals, and clerks)
499	Does the residence have a sewer connection? (No; Yes)

Figure 4 (cont.): Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly associated with poverty)
492	What is the source of water for the household? (Well, pump, public network with no connection, public network with tap outside building, or other; Public network with tap inside building)
486	What is the employment status of the female head/spouse in her main job? (Non-permanent (temporary, seasonal, or irregular) wage job, or unpaid family labor; Self-employed; Employer; Does not work; No female head/spouse; Permanent wage job;)
477	Does the household own any telephones? (None; Cellular but not land-line, or land-line but not cellular; Both land-line and cellular)
472	What type of kitchen does the household use? (None; Shared; Private)
464	Does the household own any refrigerators and/or deep freezers? (None; Refrigerator, but no deep freezer; Deep freezer, but no refrigerator, or both refrigerator and deep freezer)
462	Do all children ages 6 to 14 attend school? (No; Yes; No children 6 to 14)
454	Does the household own any televisions? (No; Yes, only black-and-white; Yes, color)
454	Does the household own any color televisions? (No; Yes)
448	Does the household own any electric irons? (No; Yes)
442	What is the main source of energy for cooking? (Kerosene, electricity, wood, firewood, other; Gas cylinders; Natural gas (piped))
421	In what sector does the female head/spouse work in her main job? (Non-profit NGOs, foreign/joint, or outside organizations; Private or cooperative; No data; No female head/spouse; Government, economic authorities, or public and public business)
419	Does the household own a television (color or black-and-white) and a videocassette recorder? (None; television only; Both)
413	Do any household members earn any income from private agricultural business? (Yes; No)
409	Do all children ages 6 to 13 attend school? (No; Yes; No children 6 to 13)

Figure 4 (cont.): Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly associated with poverty)
394	What is the principle economic activity of the female head/spouse in her main job? (Agriculture, forestry, and fishing, mining and quarrying; Does not work; No female head/spouse; Other)
373	In their main occupation, how many household members are professionals, technicians and associate professionals, clerks, or legislators, senior officials, and managers? (None; One; Two or more)
369	In their main line of work, do any household members have non-permanent (temporary, seasonal, or irregular) wage jobs? (Yes; No)
363	Does the household own any refrigerators? (No; Yes)
354	What is the principle economic activity of the male head/spouse in his main job? (Agriculture, forestry, and fishing, mining and quarrying; Construction; Other; No male head/spouse, information and communication, finance and insurance, or administrative and support services)
336	Does the household own any videocassette recorders? (No; Yes)
335	Do all children ages 6 to 12 attend school? (No; Yes; No children 6 to 12)
330	How many members ages 14 or younger does the household have? (Three or more; Two; One; None)
321	Does the household own any satellite dishes/informal cable connection? (No; Yes)
313	In their main line of work, how many household members are in outside organizations? (Two or more; One; None)
306	How many members ages 13 or younger does the household have? (Three or more; Two; One; None)
285	What is the highest educational level that a household member has completed? (Illiterate, no educational certificate, but can read and write, less than intermediate certificate, or intermediate certificate; More than intermediate but less than college; College or graduate degree (masters or Ph.D.))
271	Does the household own any black-and-white televisions? (Yes; No)
263	Do all children ages 6 to 11 attend school? (No; Yes; No children 6 to 11)
260	Does the household own any personal computers? (No; Yes)
260	Does the household own any private cars? (No; Yes)
259	Does the household own any cassette players (normal/stereo/radio)? (No; Yes)

Figure 4 (cont.): Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly associated with poverty)
254	In their main line of work, how many household members are in the private sector? (Two or more; One; None)
240	What is the material of the walls of the residence? (Stones, mud, wood, tin, asbestos, or other; Bricks with mortar; Concrete)
233	How many members ages 12 or younger does the household have? (Three or more; Two; One; None)
228	How old is the female head/spouse in complete years? (40 to 49; 50 to 59; 31 to 39; 30 or less; 61 or more; No female head/spouse)
227	What kind of ration card does the household have? (Full; Partial; None)
222	How old is the male head/spouse in complete years? (56 to 60; 46 to 55; 40 to 45; 61 or more; 36 to 39; 31 to 35; 30 or less; No male head/spouse)
219	Does the household own any normal cameras or video cameras? (No; Yes)
207	What is the employment status of the male head/spouse in his main job? (Non-permanent (temporary, seasonal, or irregular) wage job, or unpaid family labor; Employer; Self-employed; Does not work; Permanent wage job; No male head/spouse;)
203	In what sector does the male head/spouse work in his main job? (Cooperative, non-profit NGOs, foreign/joint, or outside organizations; Private; No data; Government; No male head/spouse, economic authorities, or public and public business)
197	How many members ages 11 or younger does the household have? (Three or more; Two; One; None)
183	What is the tenancy status of the household in its residence? (Owned, in-kind rent, or other; Rent-free; Ordinary/standard rent (without furniture), new rent (without furniture), or ordinary/standard or old rent with furniture)
170	In their main line of work, are any household members self-employed or business owners who employ others? (Yes; No)

Figure 4 (cont.): Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly associated with poverty)
149	Does the household own any air conditioners? (No; Yes)
148	Does the household own any deep freezers? (No; Yes)
135	What type of bathroom does the household use? (No bathroom used, or share; Private)
135	Does the household own any gas or electric stoves or microwaves? (No; Yes)
117	Does the household own any water filters? (No; Yes)
116	What is the current marital status of the male head/spouse? (Married; No male head/spouse; Single/never-married, engaged, divorced, or widowed)
115	Does the household own any gas, electric, or kerosene heaters? (No; Yes)
114	What is the structure of household headship? (Both male and female heads/spouses; Female head/spouse only; Male head/spouse only)
109	What is the current marital status of the female head/spouse? (Married, or engaged; Widowed; No female head/spouse, single/never-married, divorced)
98	Does the household own any electric fans? (No; Yes)
88	How many rooms does the residence of the household have (including parlor/reception hall)? (One; Two; Three; Four or more)
85	What is the area of the dwelling of the household in square meters? (45 or less; 46 to 60; 61 to 89; 90 to 120; 121 or more)
84	Was the female head/spouse employed in the past week? (Employed, unemployed (but had previously worked), student who works, or housewife who works outside the home or in agriculture; Newly unemployed, out of the labor force, or out of labor force; No female head/spouse)
79	In their main line of work, how many household members are in the government sector? (None; One; Two or more)

Figure 4 (cont.): Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly associated with poverty)
78	Does the household have a connection to the internet? (No; Yes)
73	Was the male head/spouse employed in the past week? (Employed, unemployed (but had previously worked), student who works, or housewife who works outside the home or in agriculture; Newly unemployed, out of the labor force, or out of labor force; No male head/spouse)
64	In their main line of work, are any household members self-employed with no employees? (Yes; No)
61	Does the household own any dishwashers? (No; Yes)
58	In their main line of work, do any household members own businesses that employ others? (Yes; No)
55	Does the household own any non-automatic clothes-washing machines? (No; Yes)
37	How many members ages 6 or younger does the household have? (Two or more; One; None)
23	In their main line of work, how many household members have permanent wage jobs? (None; One; Two or more)
17	Do any former household members now work abroad? (No; Yes)
17	Do any household members earn any income from private non-agricultural business? (No; Yes)
6	Does the household own any motorcycles? (No; Yes)
2	Does the household own any bicycles? (Yes; No)

Source: 2004/5 HIECS and the upper national poverty line.

Upper National Poverty Line

2004/5 Scorecard Applied to 2004/5 Validation Sample

(and tables pertaining to all poverty lines)

Figure 5 (Upper national line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	100.0
15–19	91.9
20–24	86.5
25–29	76.8
30–34	65.2
35–39	50.9
40–44	44.6
45–49	37.1
50–54	26.9
55–59	17.6
60–64	9.9
65–69	8.3
70–74	3.6
75–79	1.6
80–84	0.7
85–89	0.0
90–94	0.0
95–100	0.0

Figure 6 (Upper national line): Derivation of estimated poverty likelihoods associated with scores

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	93	÷	93	=	100.0
5–9	247	÷	247	=	100.0
10–14	899	÷	899	=	100.0
15–19	1,763	÷	1,919	=	91.9
20–24	2,967	÷	3,431	=	86.5
25–29	3,316	÷	4,320	=	76.8
30–34	4,619	÷	7,086	=	65.2
35–39	4,664	÷	9,164	=	50.9
40–44	3,955	÷	8,870	=	44.6
45–49	3,505	÷	9,452	=	37.1
50–54	2,293	÷	8,522	=	26.9
55–59	1,807	÷	10,282	=	17.6
60–64	848	÷	8,551	=	9.9
65–69	635	÷	7,628	=	8.3
70–74	252	÷	7,086	=	3.6
75–79	69	÷	4,374	=	1.6
80–84	31	÷	4,329	=	0.7
85–89	0	÷	1,404	=	0.0
90–94	0	÷	993	=	0.0
95–100	0	÷	1,349	=	0.0

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

Figure 7 (All poverty lines): Distribution of household poverty likelihoods across ranges demarcated by poverty lines

	Likelihood of having expenditure in range demarcated by poverty lines per day per capita							
	<$1.25/day	=>$1.25/day and <Food	=>Food and <Lower natl.	=>Lower natl. and <USAID	=>USAID and <$2.50/day	=>$2.50/day and <Upper natl.	=>Upper natl. and <$3.75/day	=>$3.75/day
Score	<EGP2.49	=>EGP2.49 and <EGP2.73	=>EGP2.73 and <EGP3.90	=>EGP3.90 and <EGP3.98	=>EGP3.98 and <EGP4.98	=>EGP4.98 and <EGP5.08	=>EGP5.08 and <EGP7.47	=>EGP7.47
0–4	0.0	34.3	65.8	0.0	0.0	0.0	0.0	0.0
5–9	66.3	0.3	33.4	0.0	0.0	0.0	0.0	0.0
10–14	24.1	10.3	48.3	3.0	14.3	0.0	0.0	0.0
15–19	13.0	9.6	37.9	1.2	24.9	5.3	6.3	1.8
20–24	6.5	9.6	42.5	0.0	27.9	0.0	12.8	0.7
25–29	5.0	2.5	38.4	0.0	28.3	2.7	20.4	2.9
30–34	2.1	2.4	27.7	0.0	30.6	2.4	29.3	5.6
35–39	1.4	0.7	17.3	1.4	27.2	2.8	39.0	10.1
40–44	0.3	1.2	12.9	1.8	27.1	1.4	42.7	12.7
45–49	1.7	0.3	9.8	0.3	23.3	1.7	46.4	16.6
50–54	0.0	1.0	8.5	1.7	14.1	1.6	48.2	24.9
55–59	0.0	0.0	3.6	2.7	9.1	2.2	41.2	41.3
60–64	0.3	0.2	1.3	1.1	6.1	1.0	40.4	49.7
65–69	0.0	0.4	2.0	0.0	5.2	0.8	27.7	64.0
70–74	0.0	0.0	0.0	1.0	1.4	1.1	17.8	78.6
75–79	0.0	0.0	0.0	0.0	1.6	0.0	6.6	91.8
80–84	0.0	0.0	0.7	0.0	0.0	0.0	9.9	89.4
85–89	0.0	0.0	0.0	0.0	0.0	0.0	5.0	95.0
90–94	0.0	0.0	0.0	0.0	0.0	0.0	2.2	97.8
95–100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0

Note: All poverty likelihoods in percentage units.

Figure 8 (Upper national line): Bootstrapped differences between estimated and true household poverty likelihoods with confidence intervals in a large sample (n = 16,384), 2004/5 scorecard applied to the 2004/5 validation sample

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+0.0	0.0	0.0	0.0
5–9	+8.7	7.0	7.9	11.0
10–14	+2.4	2.0	2.4	3.3
15–19	+1.6	2.6	3.2	4.3
20–24	+12.5	2.9	3.5	4.5
25–29	–6.8	4.5	4.7	5.1
30–34	–1.3	2.3	2.8	3.7
35–39	–4.8	3.5	3.6	4.0
40–44	–2.9	2.6	2.9	3.6
45–49	+2.8	2.0	2.4	3.2
50–54	+6.2	1.8	2.0	2.8
55–59	–0.6	1.6	1.8	2.4
60–64	–1.2	1.4	1.6	2.2
65–69	+3.8	0.9	1.1	1.4
70–74	+0.8	0.8	0.9	1.3
75–79	+1.6	0.0	0.0	0.0
80–84	+0.7	0.0	0.0	0.0
85–89	–3.9	3.1	3.3	3.7
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 9 (All poverty lines): Differences, precision of differences, and the α factor for bootstrapped estimates of poverty rates for groups of households at a point in time, 2004/5 scorecard applied to the 2004/5 validation sample

	Poverty line								
	National			USAID	International 2005 PPP				
	Upper	Lower	Food	'Extreme'	$1.25/day	$2.50/day	$3.75/day		
Estimate minus true value									
2004/5 scorecard applied to 2004/5 validation	+0.4	+0.3	+0.5	–0.1	+0.3	+0.3	+1.2		
Precision of difference									
2004/5 scorecard applied to 2004/5 validation	0.5	0.4	0.2	0.4	0.1	0.5	0.5		
α factor for sample size									
2004/5 scorecard applied to 2004/5 validation	0.81	0.88	0.94	0.90	0.98	0.81	0.77		
Precision is measured as 90-percent confidence intervals in units of +/– percentage points.									
Differences and precision estimated from 500 bootstraps of size n = 16,384.									
α is estimated from 1,000 bootstrap samples of n = 256, 512, 1,024, 2,048, 4,096, 8,192, and 16,384.									

Figure 10 (Upper national line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, 2004/5 scorecard applied to the 2004/5 validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+1.2	56.9	73.8	83.4
4	–0.1	30.3	35.6	48.2
8	–0.5	22.3	25.7	34.4
16	+0.1	15.8	18.2	24.1
32	+0.0	11.0	13.6	17.1
64	+0.1	7.9	9.9	12.7
128	+0.3	5.7	6.9	8.9
256	+0.3	4.0	4.8	6.5
512	+0.4	2.7	3.3	4.4
1,024	+0.4	1.9	2.3	3.2
2,048	+0.4	1.4	1.7	2.3
4,096	+0.4	1.0	1.1	1.5
8,192	+0.5	0.7	0.8	1.2
16,384	+0.4	0.5	0.6	0.8

Figure 11 (All poverty lines): Possible types of outcomes from targeting by poverty score

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

Figure 12 (Upper national line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, 2004/5 scorecard applied to the 2004/5 validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.1	31.4	0.0	68.5	68.6	–99.4
5–9	0.3	31.2	0.0	68.5	68.8	–97.9
10–14	1.2	30.3	0.0	68.5	69.7	–92.3
15–19	2.9	28.6	0.2	68.3	71.2	–80.7
20–24	5.5	26.0	1.1	67.4	72.9	–61.7
25–29	9.1	22.4	1.8	66.7	75.8	–36.5
30–34	13.8	17.7	4.2	64.3	78.1	+1.0
35–39	18.9	12.6	8.3	60.3	79.2	+46.3
40–44	23.1	8.4	13.0	55.6	78.6	+58.8
45–49	26.3	5.2	19.2	49.4	75.7	+39.1
50–54	28.1	3.4	25.9	42.6	70.7	+17.7
55–59	29.9	1.5	34.4	34.2	64.1	–9.1
60–64	30.9	0.6	42.0	26.6	57.4	–33.3
65–69	31.2	0.3	49.2	19.3	50.5	–56.4
70–74	31.4	0.1	56.1	12.4	43.8	–78.3
75–79	31.4	0.1	60.5	8.0	39.4	–92.2
80–84	31.4	0.1	64.8	3.7	35.1	–106.0
85–89	31.5	0.0	66.2	2.3	33.8	–110.3
90–94	31.5	0.0	67.2	1.3	32.8	–113.4
95–100	31.5	0.0	68.5	0.0	31.5	–117.7

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

Figure 13 (Upper national line): Households below the poverty line and all households at a given score or at or below a given score cut-off, 2004/5 scorecard applied to the 2004/5 validation sample

Targeting cut-off	% all households who are targeted	% targeted who are poor	% of poor who are targeted	Poor households targeted per non-poor household targeted
0–4	0.1	100.0	0.3	Only poor targeted
5–9	0.3	93.3	1.0	13.9:1
10–14	1.2	96.3	3.8	26.0:1
15–19	3.2	92.5	9.3	12.4:1
20–24	6.6	83.0	17.4	4.9:1
25–29	10.9	83.2	28.8	4.9:1
30–34	18.0	76.7	43.8	3.3:1
35–39	27.2	69.6	60.0	2.3:1
40–44	36.0	64.0	73.3	1.8:1
45–49	45.5	57.9	83.6	1.4:1
50–54	54.0	52.0	89.2	1.1:1
55–59	64.3	46.6	95.1	0.9:1
60–64	72.8	42.4	98.1	0.7:1
65–69	80.5	38.8	99.2	0.6:1
70–74	87.6	35.9	99.8	0.6:1
75–79	91.9	34.2	99.8	0.5:1
80–84	96.3	32.6	99.8	0.5:1
85–89	97.7	32.2	100.0	0.5:1
90–94	98.7	31.9	100.0	0.5:1
95–100	100.0	31.5	100.0	0.5:1

Lower National Poverty Line

2004/5 Scorecard Applied to 2004/5 Validation Sample

Figure 5 (Lower national line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	82.7
15–19	60.5
20–24	58.6
25–29	45.8
30–34	32.2
35–39	19.4
40–44	14.3
45–49	11.8
50–54	9.5
55–59	3.6
60–64	1.8
65–69	2.4
70–74	0.0
75–79	0.0
80–84	0.7
85–89	0.0
90–94	0.0
95–100	0.0

Figure 8 (Lower national line): Bootstrapped differences between estimated and true household poverty likelihoods with confidence intervals in a large sample (n = 16,384), 2004/5 scorecard applied to the 2004/5 validation sample

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	**Diff.**	**90-percent**	**95-percent**	**99-percent**
0–4	+0.0	0.0	0.0	0.0
5–9	+19.6	10.7	12.9	17.5
10–14	+3.0	5.7	6.9	8.5
15–19	+6.3	4.9	5.6	7.2
20–24	+9.1	3.5	4.2	5.6
25–29	–0.7	3.1	3.6	4.6
30–34	+2.0	2.2	2.7	3.3
35–39	–2.4	2.1	2.2	2.6
40–44	–1.1	1.6	1.9	2.6
45–49	–1.4	1.5	1.8	2.2
50–54	+1.6	1.2	1.4	1.8
55–59	+0.1	0.8	1.0	1.2
60–64	–0.1	0.6	0.7	0.9
65–69	+1.1	0.5	0.6	0.8
70–74	–0.3	0.3	0.3	0.4
75–79	+0.0	0.0	0.0	0.0
80–84	+0.7	0.0	0.0	0.0
85–89	–3.9	3.1	3.3	3.7
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 10 (Lower national line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, 2004/5 scorecard applied to the 2004/5 validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+0.3	56.4	65.7	75.5
4	+0.2	24.4	29.6	39.2
8	+0.2	17.5	20.7	26.0
16	+0.4	12.6	14.7	20.0
32	+0.3	8.6	10.1	13.0
64	+0.2	6.3	7.7	10.1
128	+0.3	4.5	5.6	7.0
256	+0.4	3.3	3.9	5.3
512	+0.3	2.2	2.8	3.6
1,024	+0.3	1.6	1.9	2.5
2,048	+0.3	1.1	1.4	1.8
4,096	+0.3	0.8	0.9	1.1
8,192	+0.3	0.5	0.6	0.9
16,384	+0.3	0.4	0.5	0.6

Figure 12 (Lower national line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, 2004/5 scorecard applied to the 2004/5 validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.1	13.8	0.0	86.1	86.2	–98.7
5–9	0.3	13.6	0.0	86.1	86.3	–95.5
10–14	1.0	12.9	0.2	85.9	86.9	–83.8
15–19	2.1	11.8	1.1	85.0	87.1	–62.4
20–24	3.8	10.1	2.8	83.3	87.0	–25.5
25–29	5.8	8.1	5.1	81.0	86.8	+20.1
30–34	7.9	6.0	10.1	76.0	84.0	+27.5
35–39	9.9	4.0	17.2	68.9	78.8	–24.1
40–44	11.3	2.6	24.7	61.4	72.6	–78.1
45–49	12.5	1.4	32.9	53.2	65.7	–137.1
50–54	13.2	0.7	40.8	45.3	58.5	–193.7
55–59	13.6	0.3	50.7	35.4	48.9	–265.1
60–64	13.7	0.2	59.1	27.0	40.7	–325.5
65–69	13.8	0.1	66.6	19.5	33.3	–379.7
70–74	13.8	0.1	73.7	12.4	26.2	–430.5
75–79	13.8	0.1	78.1	8.0	21.9	–462.0
80–84	13.8	0.1	82.4	3.7	17.5	–493.2
85–89	13.9	0.0	83.8	2.3	16.2	–502.9
90–94	13.9	0.0	84.8	1.3	15.2	–510.1
95–100	13.9	0.0	86.1	0.0	13.9	–519.8

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

Figure 13 (Lower national line): Households below the poverty line and all households at a given score or at or below a given score cut-off, 2004/5 scorecard applied to the 2004/5 validation sample

Targeting cut-off	% all households who are targeted	% targeted who are poor	% of poor who are targeted	Poor households targeted per non-poor household targeted
0–4	0.1	100.0	0.7	Only poor targeted
5–9	0.3	85.6	2.1	6.0:1
10–14	1.2	81.8	7.3	4.5:1
15–19	3.2	65.5	14.9	1.9:1
20–24	6.6	57.1	27.1	1.3:1
25–29	10.9	53.0	41.6	1.1:1
30–34	18.0	44.1	57.1	0.8:1
35–39	27.2	36.5	71.4	0.6:1
40–44	36.0	31.3	81.2	0.5:1
45–49	45.5	27.6	90.3	0.4:1
50–54	54.0	24.4	95.0	0.3:1
55–59	64.3	21.1	97.6	0.3:1
60–64	72.8	18.8	98.7	0.2:1
65–69	80.5	17.2	99.4	0.2:1
70–74	87.6	15.8	99.6	0.2:1
75–79	91.9	15.1	99.6	0.2:1
80–84	96.3	14.4	99.6	0.2:1
85–89	97.7	14.2	100.0	0.2:1
90–94	98.7	14.1	100.0	0.2:1
95–100	100.0	13.9	100.0	0.2:1

Food Poverty Line Tables

2004/5 Scorecard Applied to 2004/5 Validation Sample

Figure 5 (Food line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	34.3
5–9	66.6
10–14	34.4
15–19	22.6
20–24	16.1
25–29	7.5
30–34	4.5
35–39	2.2
40–44	1.4
45–49	2.0
50–54	1.0
55–59	0.0
60–64	0.5
65–69	0.4
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Figure 8 (Food line): Bootstrapped differences between estimated and true household poverty likelihoods with confidence intervals in a large sample (n = 16,384), 2004/5 scorecard applied to the 2004/5 validation sample

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	**Diff.**	**90-percent**	**95-percent**	**99-percent**
0–4	+8.9	18.1	21.5	29.5
5–9	+45.8	11.1	13.2	17.5
10–14	–10.1	8.6	9.2	10.5
15–19	+7.1	3.5	4.0	5.6
20–24	+6.5	2.0	2.4	3.3
25–29	–0.4	1.5	2.0	2.5
30–34	+0.6	0.9	1.1	1.5
35–39	–0.9	0.9	0.9	1.1
40–44	+0.4	0.4	0.5	0.7
45–49	+1.0	0.4	0.5	0.7
50–54	–0.3	0.5	0.6	0.8
55–59	+0.0	0.0	0.0	0.0
60–64	+0.2	0.3	0.3	0.4
65–69	+0.4	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 10 (Food line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, 2004/5 scorecard applied to the 2004/5 validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+1.1	8.1	11.3	64.2
4	+0.2	14.0	15.9	19.4
8	+0.3	7.9	8.9	16.0
16	+0.3	5.6	7.1	9.5
32	+0.4	4.0	5.0	6.5
64	+0.4	2.9	3.6	4.7
128	+0.4	2.1	2.5	3.4
256	+0.4	1.5	1.7	2.2
512	+0.4	1.1	1.2	1.6
1,024	+0.4	0.7	0.9	1.2
2,048	+0.5	0.5	0.7	0.8
4,096	+0.5	0.4	0.4	0.6
8,192	+0.5	0.3	0.3	0.4
16,384	+0.5	0.2	0.2	0.3

Figure 12 (Food line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, 2004/5 scorecard applied to the 2004/5 validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.0	2.4	0.1	97.5	97.6	–95.1
5–9	0.1	2.3	0.3	97.3	97.4	–82.6
10–14	0.5	1.9	0.8	96.9	97.3	–28.0
15–19	0.8	1.6	2.4	95.2	96.0	+0.6
20–24	1.1	1.3	5.5	92.1	93.2	–129.6
25–29	1.5	0.9	9.4	88.2	89.6	–295.6
30–34	1.8	0.6	16.2	81.4	83.1	–580.5
35–39	2.0	0.3	25.1	72.5	74.5	–952.3
40–44	2.1	0.2	33.9	63.7	65.9	–1,319.8
45–49	2.2	0.1	43.2	54.4	56.6	–1,711.6
50–54	2.4	0.0	51.6	46.0	48.3	–2,063.7
55–59	2.4	0.0	61.9	35.7	38.0	–2,494.4
60–64	2.4	0.0	70.5	27.2	29.5	–2,851.6
65–69	2.4	0.0	78.1	19.5	21.9	–3,171.1
70–74	2.4	0.0	85.2	12.4	14.8	–3,468.0
75–79	2.4	0.0	89.5	8.1	10.5	–3,651.2
80–84	2.4	0.0	93.9	3.7	6.1	–3,832.6
85–89	2.4	0.0	95.3	2.3	4.7	–3,891.4
90–94	2.4	0.0	96.3	1.3	3.7	–3,933.1
95–100	2.4	0.0	97.6	0.0	2.4	–3,989.6

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

Figure 13 (Food line): Households below the poverty line and all households at a given score or at or below a given score cut-off, 2004/5 scorecard applied to the 2004/5 validation sample

Targeting cut-off	% all households who are targeted	% targeted who are poor	% of poor who are targeted	Poor households targeted per non-poor household targeted
0–4	0.1	25.1	1.0	0.3:1
5–9	0.3	21.7	3.1	0.3:1
10–14	1.2	38.7	20.1	0.6:1
15–19	3.2	24.9	32.9	0.3:1
20–24	6.6	16.8	46.5	0.2:1
25–29	10.9	13.4	61.4	0.2:1
30–34	18.0	9.7	73.4	0.1:1
35–39	27.2	7.5	85.6	0.1:1
40–44	36.0	5.9	89.7	0.1:1
45–49	45.5	4.9	93.9	0.1:1
50–54	54.0	4.4	98.9	0.0:1
55–59	64.3	3.7	98.9	0.0:1
60–64	72.8	3.3	100.0	0.0:1
65–69	80.5	3.0	100.0	0.0:1
70–74	87.6	2.7	100.0	0.0:1
75–79	91.9	2.6	100.0	0.0:1
80–84	96.3	2.5	100.0	0.0:1
85–89	97.7	2.4	100.0	0.0:1
90–94	98.7	2.4	100.0	0.0:1
95–100	100.0	2.4	100.0	0.0:1

USAID "Extreme" Poverty Line

2004/5 Scorecard Applied to 2004/5 Validation Sample

Figure 5 (USAID "extreme" line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	91.9
10–14	85.7
15–19	61.7
20–24	56.4
25–29	45.6
30–34	30.3
35–39	20.9
40–44	16.1
45–49	12.1
50–54	11.2
55–59	6.3
60–64	2.8
65–69	1.2
70–74	1.0
75–79	0.0
80–84	0.7
85–89	0.0
90–94	0.0
95–100	0.0

Figure 8 (USAID "extreme" line): Bootstrapped differences between estimated and true household poverty likelihoods with confidence intervals in a large sample (n = 16,384), 2004/5 scorecard applied to the 2004/5 validation sample

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+0.0	0.0	0.0	0.0
5–9	+20.6	11.8	13.7	20.6
10–14	+9.4	5.7	6.9	9.1
15–19	+6.1	4.7	5.8	7.5
20–24	+6.3	3.4	4.0	5.7
25–29	–2.2	3.1	3.7	4.6
30–34	–2.4	2.3	2.6	3.7
35–39	–0.9	1.7	2.0	2.8
40–44	–2.5	2.1	2.3	2.9
45–49	–2.1	1.8	2.0	2.3
50–54	+4.5	1.1	1.3	1.8
55–59	–1.0	1.1	1.2	1.6
60–64	–0.4	0.8	0.9	1.3
65–69	–0.9	0.8	0.8	1.0
70–74	+0.6	0.3	0.4	0.4
75–79	+0.0	0.0	0.0	0.0
80–84	+0.7	0.0	0.0	0.0
85–89	–3.9	3.1	3.3	3.7
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 10 (USAID "extreme" line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, 2004/5 scorecard applied to the 2004/5 validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+0.2	54.7	64.7	77.7
4	–0.6	26.2	31.4	41.6
8	–0.9	18.3	21.4	27.2
16	–0.5	13.1	14.8	19.1
32	–0.5	8.9	10.1	13.8
64	–0.2	6.6	7.8	10.1
128	–0.2	4.7	5.8	7.2
256	–0.1	3.5	4.1	4.9
512	–0.1	2.3	2.7	3.9
1,024	–0.1	1.7	2.0	2.5
2,048	–0.1	1.2	1.4	1.9
4,096	–0.1	0.8	1.0	1.3
8,192	–0.1	0.6	0.7	0.9
16,384	–0.1	0.4	0.5	0.6

Figure 12 (USAID "extreme" line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, 2004/5 scorecard applied to the 2004/5 validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.1	14.8	0.0	85.1	85.2	–98.7
5–9	0.3	14.6	0.1	85.0	85.3	–95.9
10–14	1.0	13.9	0.3	84.8	85.8	–85.3
15–19	2.0	12.9	1.1	84.0	86.0	–65.3
20–24	3.7	11.2	2.9	82.2	86.0	–30.7
25–29	5.8	9.1	5.1	80.0	85.8	+12.2
30–34	8.1	6.8	9.8	75.3	83.4	+33.9
35–39	10.1	4.8	17.0	68.1	78.2	–14.2
40–44	11.8	3.1	24.3	60.8	72.6	–63.0
45–49	13.1	1.8	32.4	52.7	65.8	–117.5
50–54	13.7	1.2	40.3	44.8	58.4	–170.8
55–59	14.4	0.5	49.9	35.2	49.6	–234.9
60–64	14.7	0.2	58.2	26.9	41.6	–290.5
65–69	14.8	0.1	65.6	19.5	34.3	–340.7
70–74	14.8	0.1	72.7	12.4	27.2	–388.0
75–79	14.8	0.1	77.1	8.0	22.9	–417.4
80–84	14.8	0.1	81.4	3.7	18.5	–446.5
85–89	14.9	0.0	82.8	2.3	17.2	–455.5
90–94	14.9	0.0	83.8	1.3	16.2	–462.2
95–100	14.9	0.0	85.1	0.0	14.9	–471.3

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

Figure 13 (USAID "extreme" line): Households below the poverty line and all households at a given score or at or below a given score cut-off, 2004/5 scorecard applied to the 2004/5 validation sample

Targeting cut-off	% all households who are targeted	% targeted who are poor	% of poor who are targeted	Poor households targeted per non-poor household targeted
0–4	0.1	100.0	0.6	Only poor targeted
5–9	0.3	78.7	1.8	3.7:1
10–14	1.2	76.7	6.4	3.3:1
15–19	3.2	63.9	13.5	1.8:1
20–24	6.6	56.7	25.1	1.3:1
25–29	10.9	53.2	39.0	1.1:1
30–34	18.0	45.3	54.7	0.8:1
35–39	27.2	37.3	68.1	0.6:1
40–44	36.0	32.6	78.9	0.5:1
45–49	45.5	28.8	87.8	0.4:1
50–54	54.0	25.3	91.7	0.3:1
55–59	64.3	22.4	96.7	0.3:1
60–64	72.8	20.1	98.4	0.3:1
65–69	80.5	18.4	99.5	0.2:1
70–74	87.6	17.0	99.6	0.2:1
75–79	91.9	16.1	99.6	0.2:1
80–84	96.3	15.4	99.6	0.2:1
85–89	97.7	15.3	100.0	0.2:1
90–94	98.7	15.1	100.0	0.2:1
95–100	100.0	14.9	100.0	0.2:1

$1.25/day 2005 PPP Poverty Line

2004/5 Scorecard Applied to 2004/5 Validation Sample

Figure 5 ($1.25/day 2005 PPP line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	0.0
5–9	66.3
10–14	24.1
15–19	13.0
20–24	6.5
25–29	5.0
30–34	2.1
35–39	1.4
40–44	0.3
45–49	1.7
50–54	0.0
55–59	0.0
60–64	0.3
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 8 ($1.25/day 2005 PPP line): Bootstrapped differences between estimated and true household poverty likelihoods with confidence intervals in a large sample (n = 16,384), 2004/5 scorecard applied to the 2004/5 validation sample

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	–25.4	22.3	25.0	29.5
5–9	+55.7	8.4	9.6	11.9
10–14	–7.5	7.1	7.8	10.7
15–19	+5.6	2.5	3.1	3.7
20–24	–0.5	1.9	2.2	2.8
25–29	+0.8	1.2	1.4	1.9
30–34	+0.1	0.7	0.8	1.0
35–39	+0.4	0.4	0.5	0.7
40–44	+0.0	0.2	0.2	0.3
45–49	+1.3	0.3	0.3	0.4
50–54	–0.6	0.4	0.5	0.5
55–59	+0.0	0.0	0.0	0.0
60–64	+0.3	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 10 ($1.25/day 2005 PPP line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, 2004/5 scorecard applied to the 2004/5 validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+0.7	3.3	6.5	57.2
4	+0.2	9.9	14.1	16.9
8	+0.1	6.8	7.7	13.7
16	+0.3	3.9	4.7	7.8
32	+0.3	2.9	3.7	5.2
64	+0.3	2.0	2.5	3.4
128	+0.3	1.5	1.8	2.4
256	+0.3	1.1	1.3	1.6
512	+0.3	0.8	0.9	1.2
1,024	+0.3	0.6	0.7	0.9
2,048	+0.3	0.4	0.5	0.6
4,096	+0.3	0.3	0.3	0.5
8,192	+0.3	0.2	0.2	0.3
16,384	+0.3	0.1	0.2	0.2

Figure 12 ($1.25/day 2005 PPP line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, 2004/5 scorecard applied to the 2004/5 validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.0	1.2	0.1	98.7	98.7	–90.7
5–9	0.0	1.2	0.3	98.5	98.5	–69.0
10–14	0.3	0.9	0.9	97.8	98.2	+25.3
15–19	0.5	0.8	2.7	96.1	96.5	–113.2
20–24	0.7	0.5	5.9	92.9	93.6	–367.9
25–29	0.9	0.4	10.0	88.7	89.6	–696.6
30–34	1.0	0.2	17.0	81.8	82.8	–1,248.7
35–39	1.1	0.1	26.0	72.7	73.9	–1,970.2
40–44	1.2	0.1	34.9	63.9	65.0	–2,674.0
45–49	1.2	0.0	44.3	54.5	55.7	–3,422.3
50–54	1.3	0.0	52.7	46.0	47.3	–4,096.4
55–59	1.3	0.0	63.0	35.7	37.0	–4,914.4
60–64	1.3	0.0	71.6	27.2	28.4	–5,594.8
65–69	1.3	0.0	79.2	19.5	20.8	–6,201.6
70–74	1.3	0.0	86.3	12.4	13.7	–6,765.3
75–79	1.3	0.0	90.7	8.1	9.3	–7,113.3
80–84	1.3	0.0	95.0	3.7	5.0	–7,457.7
85–89	1.3	0.0	96.4	2.3	3.6	–7,569.5
90–94	1.3	0.0	97.4	1.3	2.6	–7,648.5
95–100	1.3	0.0	98.7	0.0	1.3	–7,755.8

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

Figure 13 ($1.25/day 2005 PPP line): Households below the poverty line and all households at a given score or at or below a given score cut-off, 2004/5 scorecard applied to the 2004/5 validation sample

Targeting cut-off	% all households who are targeted	% targeted who are poor	% of poor who are targeted	Poor households targeted per non-poor household targeted
0–4	0.1	25.1	1.9	0.3:1
5–9	0.3	14.4	3.9	0.2:1
10–14	1.2	27.1	26.7	0.4:1
15–19	3.2	15.2	38.1	0.2:1
20–24	6.6	10.7	56.3	0.1:1
25–29	10.9	8.2	71.4	0.1:1
30–34	18.0	5.8	83.0	0.1:1
35–39	27.2	4.2	90.6	0.0:1
40–44	36.0	3.2	92.5	0.0:1
45–49	45.5	2.7	96.2	0.0:1
50–54	54.0	2.3	100.0	0.0:1
55–59	64.3	2.0	100.0	0.0:1
60–64	72.8	1.7	100.0	0.0:1
65–69	80.5	1.6	100.0	0.0:1
70–74	87.6	1.4	100.0	0.0:1
75–79	91.9	1.4	100.0	0.0:1
80–84	96.3	1.3	100.0	0.0:1
85–89	97.7	1.3	100.0	0.0:1
90–94	98.7	1.3	100.0	0.0:1
95–100	100.0	1.3	100.0	0.0:1

$2.50/day 2005 PPP Poverty Line

2004/5 Scorecard Applied to 2004/5 Validation Sample

Figure 5 ($2.50/day 2005 PPP line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	100.0
15–19	86.6
20–24	86.5
25–29	74.1
30–34	62.8
35–39	48.1
40–44	43.2
45–49	35.3
50–54	25.3
55–59	15.4
60–64	8.9
65–69	7.6
70–74	2.5
75–79	1.6
80–84	0.7
85–89	0.0
90–94	0.0
95–100	0.0

Figure 8 ($2.50/day 2005 PPP line): Bootstrapped differences between estimated and true household poverty likelihoods with confidence intervals in a large sample (n = 16,384), 2004/5 scorecard applied to the 2004/5 validation sample

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+0.0	0.0	0.0	0.0
5–9	+8.7	7.0	7.9	11.0
10–14	+5.1	2.9	3.4	4.8
15–19	–2.3	2.9	3.4	4.7
20–24	+13.0	2.9	3.4	4.6
25–29	–7.3	4.8	5.0	5.4
30–34	–1.0	2.3	2.8	3.9
35–39	–5.8	4.0	4.1	4.5
40–44	–1.2	2.2	2.6	3.5
45–49	+3.1	2.0	2.4	3.1
50–54	+5.2	1.7	2.1	2.7
55–59	–1.0	1.5	1.7	2.2
60–64	–0.9	1.3	1.6	2.1
65–69	+3.4	0.9	1.1	1.4
70–74	–0.3	0.8	0.9	1.3
75–79	+1.6	0.0	0.0	0.0
80–84	+0.7	0.0	0.0	0.0
85–89	–3.9	3.1	3.3	3.7
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 10 ($2.50/day 2005 PPP line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, 2004/5 scorecard applied to the 2004/5 validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+1.3	56.4	68.7	82.6
4	–0.3	29.5	34.6	48.4
8	–0.4	22.0	25.3	33.0
16	+0.1	15.1	18.6	23.9
32	–0.1	11.0	13.2	17.9
64	–0.0	8.0	10.1	13.0
128	+0.2	5.8	7.1	9.3
256	+0.2	3.9	4.9	6.6
512	+0.3	2.7	3.3	4.5
1,024	+0.3	1.9	2.3	3.3
2,048	+0.3	1.4	1.6	2.3
4,096	+0.3	1.0	1.1	1.5
8,192	+0.3	0.7	0.8	1.1
16,384	+0.3	0.5	0.5	0.8

Figure 12 ($2.50/day 2005 PPP line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, 2004/5 scorecard applied to the 2004/5 validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.1	30.0	0.0	69.9	70.0	–99.4
5–9	0.3	29.8	0.0	69.9	70.2	–97.8
10–14	1.2	28.9	0.1	69.8	71.0	–92.0
15–19	2.9	27.2	0.3	69.6	72.5	–80.0
20–24	5.4	24.7	1.2	68.7	74.1	–60.2
25–29	8.9	21.2	2.0	67.9	76.8	–34.2
30–34	13.5	16.7	4.5	65.4	78.8	+4.5
35–39	18.4	11.7	8.8	61.1	79.5	+51.3
40–44	22.3	7.8	13.8	56.1	78.4	+54.3
45–49	25.3	4.8	20.2	49.7	75.1	+33.1
50–54	27.0	3.1	27.0	42.9	70.0	+10.4
55–59	28.7	1.4	35.6	34.3	63.0	–18.2
60–64	29.5	0.6	43.3	26.6	56.1	–43.9
65–69	29.9	0.3	50.6	19.3	49.1	–68.1
70–74	30.1	0.1	57.5	12.4	42.4	–91.0
75–79	30.1	0.1	61.9	8.0	38.1	–105.5
80–84	30.1	0.1	66.2	3.7	33.7	–119.9
85–89	30.1	0.0	67.6	2.3	32.4	–124.4
90–94	30.1	0.0	68.5	1.3	31.5	–127.7
95–100	30.1	0.0	69.9	0.0	30.1	–132.2

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

Figure 13 ($2.50/day 2005 PPP line): Households below the poverty line and all households at a given score or at or below a given score cut-off, 2004/5 scorecard applied to the 2004/5 validation sample

Targeting cut-off	% all households who are targeted	% targeted who are poor	% of poor who are targeted	Poor households targeted per non-poor household targeted
0–4	0.1	100.0	0.3	Only poor targeted
5–9	0.3	93.3	1.1	13.9:1
10–14	1.2	94.4	3.9	16.7:1
15–19	3.2	90.9	9.5	10.0:1
20–24	6.6	81.9	17.9	4.5:1
25–29	10.9	81.6	29.6	4.4:1
30–34	18.0	74.7	44.7	3.0:1
35–39	27.2	67.7	61.0	2.1:1
40–44	36.0	61.8	74.0	1.6:1
45–49	45.5	55.7	84.1	1.3:1
50–54	54.0	50.1	89.8	1.0:1
55–59	64.3	44.7	95.4	0.8:1
60–64	72.8	40.5	98.1	0.7:1
65–69	80.5	37.1	99.2	0.6:1
70–74	87.6	34.3	99.8	0.5:1
75–79	91.9	32.7	99.8	0.5:1
80–84	96.3	31.2	99.8	0.5:1
85–89	97.7	30.8	100.0	0.4:1
90–94	98.7	30.5	100.0	0.4:1
95–100	100.0	30.1	100.0	0.4:1

$3.75/day 2005 PPP Poverty Line

2004/5 Scorecard Applied to 2004/5 Validation Sample

Figure 5 ($3.75/day 2005 PPP line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	100.0
15–19	98.2
20–24	99.3
25–29	97.1
30–34	94.4
35–39	89.9
40–44	87.3
45–49	83.4
50–54	75.1
55–59	58.7
60–64	50.3
65–69	36.0
70–74	21.4
75–79	8.2
80–84	10.6
85–89	5.0
90–94	2.2
95–100	0.0

Figure 8 ($3.75/day 2005 PPP line): Bootstrapped differences between estimated and true household poverty likelihoods with confidence intervals in a large sample (n = 16,384), 2004/5 scorecard applied to the 2004/5 validation sample

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+0.0	0.0	0.0	0.0
5–9	+0.0	0.0	0.0	0.0
10–14	+0.0	0.0	0.0	0.0
15–19	–1.8	0.9	0.9	0.9
20–24	+0.5	0.8	0.9	1.2
25–29	–1.1	0.9	1.0	1.2
30–34	+0.4	1.1	1.4	1.8
35–39	–1.6	1.4	1.5	1.7
40–44	+1.9	1.5	1.8	2.6
45–49	+3.2	1.8	2.1	2.7
50–54	+3.9	2.0	2.4	3.0
55–59	+3.6	2.0	2.4	3.1
60–64	+2.7	2.2	2.6	3.2
65–69	+5.0	2.2	2.6	3.3
70–74	–2.9	2.5	2.7	3.1
75–79	–4.9	3.5	3.7	4.1
80–84	–0.8	2.1	2.4	3.1
85–89	+1.1	2.1	2.5	3.3
90–94	–0.0	1.9	2.1	2.7
95–100	+0.0	0.0	0.0	0.0

Figure 10 ($3.75/day 2005 PPP line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, 2004/5 scorecard applied to the 2004/5 validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+2.4	66.6	83.0	90.9
4	+1.5	31.0	37.2	49.3
8	+0.9	22.7	26.8	34.4
16	+1.0	15.6	18.4	25.8
32	+0.8	10.7	12.5	16.8
64	+0.9	7.7	9.4	11.7
128	+1.0	5.3	6.2	8.0
256	+1.0	4.0	4.7	6.0
512	+1.0	2.7	3.2	4.3
1,024	+1.1	1.9	2.4	3.0
2,048	+1.1	1.3	1.6	2.3
4,096	+1.2	0.9	1.1	1.6
8,192	+1.2	0.7	0.8	1.0
16,384	+1.2	0.5	0.6	0.8

Figure 12 ($3.75/day 2005 PPP line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, 2004/5 scorecard applied to the 2004/5 validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.1	61.9	0.0	38.0	38.1	–99.7
5–9	0.3	61.6	0.0	38.0	38.4	–98.9
10–14	1.2	60.7	0.0	38.0	39.3	–96.0
15–19	3.2	58.8	0.0	38.0	41.2	–89.8
20–24	6.5	55.4	0.0	38.0	44.5	–78.8
25–29	10.8	51.2	0.1	37.9	48.7	–65.0
30–34	17.5	44.5	0.5	37.5	55.0	–42.8
35–39	25.8	36.1	1.3	36.7	62.5	–14.5
40–44	33.4	28.6	2.6	35.4	68.8	+12.1
45–49	41.0	21.0	4.5	33.6	74.5	+39.6
50–54	47.1	14.9	7.0	31.1	78.2	+63.1
55–59	52.7	9.2	11.6	26.5	79.2	+81.3
60–64	56.7	5.2	16.1	22.0	78.7	+74.0
65–69	59.1	2.8	21.4	16.7	75.8	+65.5
70–74	60.8	1.1	26.7	11.3	72.1	+56.8
75–79	61.4	0.6	30.5	7.5	68.9	+50.7
80–84	61.9	0.1	34.4	3.7	65.5	+44.5
85–89	61.9	0.0	35.7	2.3	64.2	+42.3
90–94	62.0	0.0	36.7	1.3	63.3	+40.8
95–100	62.0	0.0	38.0	0.0	62.0	+38.6

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

Figure 13 ($3.75/day 2005 PPP line): Households below the poverty line and all households at a given score or at or below a given score cut-off, 2004/5 scorecard applied to the 2004/5 validation sample

Targeting cut-off	% all households who are targeted	% targeted who are poor	% of poor who are targeted	Poor households targeted per non-poor household targeted
0–4	0.1	100.0	0.2	Only poor targeted
5–9	0.3	100.0	0.5	Only poor targeted
10–14	1.2	100.0	2.0	Only poor targeted
15–19	3.2	100.0	5.1	Only poor targeted
20–24	6.6	99.3	10.6	146.3:1
25–29	10.9	98.9	17.4	89.2:1
30–34	18.0	97.0	28.2	32.0:1
35–39	27.2	95.1	41.7	19.4:1
40–44	36.0	92.7	53.9	12.7:1
45–49	45.5	90.1	66.2	9.1:1
50–54	54.0	87.1	76.0	6.8:1
55–59	64.3	82.0	85.1	4.6:1
60–64	72.8	77.9	91.6	3.5:1
65–69	80.5	73.5	95.4	2.8:1
70–74	87.6	69.5	98.2	2.3:1
75–79	91.9	66.8	99.1	2.0:1
80–84	96.3	64.3	99.9	1.8:1
85–89	97.7	63.4	100.0	1.7:1
90–94	98.7	62.8	100.0	1.7:1
95–100	100.0	62.0	100.0	1.6:1