Simple Poverty Scorecard® Poverty-Assessment Tool Jordan

Mark Schreiner

21 October 2010

This document is at `SimplePovertyScorecard.com`

Abstract

The Simple Poverty Scorecard-brand poverty-assessment tool uses 10 low-cost indicators from Jordan's 2006 Household Income and Expenditure 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 Jordan to measure poverty rates, to track changes in poverty rates over time, and to segment clients for differentiated treatment.

Acknowledgements

This paper is funded by the United Nations Relief and Works Agency for Palestine Refugees in the Near East via a contract with Microfinanza SRL. Data are from Jordan's Department of Statistics. Thanks go to Zeinab Al-Dabbagh, Maha Dawass, Abdel Al-Fattah Jaradat, Micol Guarneri, Fabio Malanchini, Alex Pollock, Safwat Radaideh, and Elsa Valli.

Simple Poverty Scorecard® Poverty-Assessment Tool

			Name	Identifier
Interview ID:				
Interview date:		Participant:		
Country:	JOR	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. Nine or more	0	
	B. Eight	4	
	C. Seven	7	
	D. Six	13	
	E. Five	15	
	F. Four	23	
	G. Three	30	
	H. One or two	38	
2. How many household members worked at least one hour in the past week?	A. None	0	
	B. One	1	
	C. Two	2	
	D. Three or more	3	
3. Does any household member work as a legislator, senior official, manager, professional, technician, or associated professional?	A. No	0	
	B. Yes	3	
4. How many rooms (not counting the kitchen) does the residence of the household have?	A. One or two	0	
	B. Three	6	
	C. Four	7	
	D. Five	11	
	E. Six or more	18	
5. Does the household own a combined gas stove with oven?	A. No	0	
	B. Yes	3	
6. Does the household own a vacuum cleaner?	A. No	0	
	B. Yes	3	
7. Does the household own an air conditioner?	A. No	0	
	B. Yes	5	
8. Does the household have a computer connected to the internet?	A. No	0	
	B. Only computer	3	
	C. Computer and internet	6	
9. How many land-line and/or mobile telephones does the household have?	A. None	0	
	B. One mobile, but no land-lines	6	
	C. One or more land-lines, but no mobile	7	
	D. One or more landlines, and one mobile	8	
	E. Two or more mobiles, but no land-lines	10	
	F. One or more landlines, and two or more mobiles	15	
10. Does the household own a private car?	A. No	0	
	B. Yes	6	

Score:

Simple Poverty Scorecard® Poverty-Assessment Tool Jordan

1. Introduction

Pro-poor programs in Jordan 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, Jordan's 2006 Household Income and Expenditure Survey (HIES) runs more than 60 pages. Enumerators visit households quarterly and also weekly for 30 weeks, applying a consumption module with hundreds of questions.

In contrast, the indirect approach via the scorecard is simple, quick, and inexpensive. It uses ten verifiable indicators (such as "How many rooms (not counting the kitchen) does the residence of the household have?" and "Does the household have a vacuum cleaner?") to get a score that is correlated with poverty status as measured by consumption from the exhaustive 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.

The scorecard would be useful for organizations that want to know what share of their participants are below a poverty line, perhaps because they want to relate their participants' poverty status to the Millennium Development Goals' $1.25/day poverty line at 2005 purchase-power parity (PPP). The scorecard could also be useful to organizations that 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. The scorecard would also be useful for to organizations that want to measure movement across a poverty line (Daley-Harris, 2009). In short, the scorecard is a 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 poverty-ssessment tool.

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 often have complex indicators and 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, transparent 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 2006 HIES conducted by Jordan's Department of Statistics. 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 an appropriate 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 a poverty line that is 150% of Jordan's national poverty line.[1] Scores from this scorecard are calibrated to poverty likelihoods for nine poverty lines, including the national line.

The scorecard is constructed and calibrated using two sub-samples from the 2006 HIES, and its accuracy is validated on a third sub-sample. While all three scoring estimators are unbiased when applied to the population from which they were derived

[1] For sample-size reasons, this line works better than the national line.

(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.[2]

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.[3] Of course, these assumptions—ubiquitous and inevitable in predictive modeling—hold only partly.

When applied to the validation sample for Jordan with the 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 +0.2 percentage points. Across all nine lines, the average absolute difference is 0.4 percentage points, and the maximum absolute difference is 1.0 percentage points.

Because the validation sample is representative of the same population as the data that is used in construction/calibration and because all the data come from the same time period, the scorecard estimators are unbiased and these observed differences are due to sampling variation. That is, the average difference would be zero if the 2006

[2] Important examples of "different populations" are nationally representative samples at another point in time or non-representative sub-groups (Tarozzi and Deaton, 2007).
[3] Bias may also result from changes over time in data quality or in poverty lines.

HIES were to be repeatedly redone 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 ± 0.7 percentage points or less. For $n = 1{,}024$, these intervals are ± 2.7 percentage points or less.

Section 2 below documents data, poverty lines, and poverty rates for Jordan. Sections 3 and 4 describe scorecard construction and offer practical guidelines for implementation. 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 discusses an existing similar exercise for Jordan. The final section is a summary.

2. Data and poverty lines

This section discusses the data used to construct and validate the Simple scorecard. It also defines the poverty lines to which scores are calibrated.

2.1 Data

The scorecard is based on data from the 23,278 households in the 2006 HIES. This is the most recent national consumption survey available for Jordan.[4] For scoring, the data 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

[4] There is a 2008 HIES, but it is not available to this project.

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 Jordan's 2006 HIES and household-level lines. Scores are calibrated to household-level poverty likelihoods, and accuracy is measured for household-level rates.

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 2 shows household-level poverty rates for the three sub-samples used in scorecard construction, calibration, and validation. Figure 3 shows nine poverty lines and poverty rates (household level and person level) for Jordan as a whole and for each governorate.

Jordan's food poverty line is defined as the cost of a food basket that on average provides 2,340 Calories per person per day, adjusted for prices in a given governorate and for the age and sex of people in a given household (World Bank, 2008a). The prices used are "poverty prices" in that they are based on a basket of 193 food items and 259 non-food items whose consumption shares derived from households in the bottom consumption quintile in the 2006 HIES. On average, the food line is JOD0.66 per person per day. The scorecard is not calibrated to this food line, as almost no one in Jordan has consumption below it.[5]

After an adjustment for economies of household size in non-food consumption (World Bank, 2008a), Jordan's national poverty line is defined as the per-capita total

[5] The scorecard is calibrated to $1.25/day 2005 PPP, which is the same as the average food line, see below.

consumption estimated via regression for six-person households whose food consumption is set equal to the food poverty line. Thus, the national line includes the cost of the caloric minimum and the cost of non-food goods and services that people consume even before they reach the caloric minimum. For Jordan as a whole, the average national line was JOD1.52 per person per day (Figure 2), giving a household-level poverty rate of 9.6 percent and a person-level rate of 13.0 percent.

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

- National
- 150% of national
- 200% of national
- 250% of national
- USAID "extreme"
- USD1.25/day 2005 PPP
- USD2.50/day 2005 PPP
- USD3.75/day 2005 PPP
- USD5.00/day 2005 PPP

Scorecard construction uses 150% of the national line, as the number of poor households under the national line is too low for the methods used here.

The USAID "extreme" line is defined as the median consumption of people (not households) below the national line (U.S. Congress, 2002). It is calculated by governorate and averages JOD1.23 overall.

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

- 2005 PPP exchange rate for "individual consumption expenditure by households" (World Bank, 2008b): JOD0.49 per $1.00
- Monthly consumer price indices from Jordan's Department of Statistics: 101.27 on average for June–December 2006 when the HIES was in the field, and 94.11 for all of 2005

The $1.25/day 2005 PPP line for Jordan as a whole for June–December 2006 is then (Sillers, 2006):

$$(\text{2005 PPP exchange rate}) \cdot \$1.25 \cdot \frac{\text{CPI}_{\text{June-December 2006}}}{\text{CPI}_{\text{Ave. 2005}}} =$$

$$\left(\frac{\text{JOD0.49}}{\$1.00}\right) \cdot \$1.25 \cdot \frac{101.27}{94.11} = \text{JOD0.66}.$$

This $1.25/day 2005 PPP line applies to Jordan as a whole. It is adjusted for differences in cost-of-living across governorates using the 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 of JOD0.66, multiplied by that household's specific national line, and divided by the average all-Jordan national line.

3. Scorecard construction

For the Jordan scorecard, about 70 potential indicators are initially prepared in the areas of:

- Family composition (such as household size)
- Education (such as education of the female head/spouse)
- Employment (such as the number of household members who work)
- Housing (such as the number of rooms)
- Ownership of durable goods (such as vacuum cleaners or private cars)

Figure 4 lists all the candidate indicators, ranked by the entropy-based "uncertainty coefficient", a measure of how well an 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 computer 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 150% of the 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 its 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 theory, experience, 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 with 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 Jordan. Tests for Mexico and India (Schreiner, 2006a and 2006b), Sri Lanka (Narayan and Yoshida, 2005), and Jamaica (Grosh and Baker, 1995) suggest that segmenting poverty-assessment tools by

urban/rural does not improve targeting 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 in 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. High-quality outputs require high-quality 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).[6] IRIS Center (2007a) and Toohig (2008) are useful nuts-and-bolts guides for planning, budgeting, training field

[6] 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 at the central office.

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 are 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 an example from Mexico, Martinelli and Parker (2007) find that in the first stage of targeting a 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 Mexico does in the second stage of its 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 a sub-group of interest for a particular question

If not determined by other factors, the number of participants to be scored can be derived from sample-size formulas (presented later) for a desired 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 are recorded on paper in the field before being sent to a central office to be entered into a database and scored. The sampling plans of ASA and BRAC cover 50,000–100,000 participants each, which is far more than would be required to inform most relevant questions 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 Jordan, scores range from 0 to 100. While higher scores indicate less likelihood of being below a poverty line, the scores themselves have only relative units. For example, doubling the score does not double the likelihood of being above a poverty line.

To get absolute units, scores must be converted to *poverty likelihoods*, that is, probabilities of being below a poverty line. This is done via simple look-up tables. For the example of the national line with the 2006 HIES, scores of 15–19 have a poverty likelihood of 39.5 percent, and scores of 20–24 have a poverty likelihood of 22.6 percent (Figure 5).

Naturally, the poverty likelihood associated with a score varies by poverty line. For example, scores of 15–19 are associated with a poverty likelihood of 39.5 percent for the national line but 78.5 percent for the 150% of the national line.[7]

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.

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

For the example of the national line (Figure 6), there are 5,010 (normalized) households in the calibration sub-sample with a score of 15–19, of whom 1,978 (normalized) are below the poverty line. The estimated poverty likelihood associated with a score of 15–19 is then 39.5 percent, as 1,978 ÷ 5,010 = 0.395.

To illustrate further with the national line and a score of 20–24, there are 8,959 (normalized) households in the calibration sample, of whom 2,021 (normalized) are below the line (Figure 6). Thus, the poverty likelihood for this score is 2,021 ÷ 8,959 = 22.6 percent.

The same method is used to calibrate scores with estimated poverty likelihoods for all nine 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 15–19 falls in the following ranges with probability:

- 2.1 percent below the $1.25/day 2005 PPP line
- 20.2 percent between the $1.25/day 2005 PPP and the USAID "extreme" lines
- 4.4 percent between the USAID "extreme" and the $2.50/day 2005 PPP lines
- 12.9 percent between the $2.50/day 2005 PPP and 100% of the national lines
- 28.7 percent between 100% of the national and $3.75/day 2005 PPP lines
- 10.3 percent between the $3.75/day 2005 PPP and 150% of the national lines
- 12.1 percent between 150% of the national and $5.00/day 2005 PPP lines
- 5.5 percent between $5.00/day 2005 PPP and 200% of the national and lines
- 2.3 percent between 200% of the national and 250% of the national lines
- 1.5 percent above 250% of the national 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, consumption-based poverty lines and survey data. 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 Jordan's scorecard are transformed coefficients from a Logit regression, scores are not converted to poverty likelihoods via the Logit formula of $2.718281828^{\text{score}}$ x $(1+ 2.718281828^{\text{score}})^{-1}$. 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.[8]

The relationship between indicators and poverty, however, does change with time and also across sub-groups in Jordan's population. Thus, the scorecard is generally biased when applied after the end date of fieldwork for the 2006 HIES (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).

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 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 range, 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 range, report the average difference between estimated and true poverty likelihoods across the 1,000 bootstrap samples
- For each score range, report the two-sided interval containing the central 900, 950, or 990 differences between estimated and true poverty likelihoods

[8] This follows because these estimates of groups' poverty rates are linear functions of the unbiased estimates of households' 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 national line in the validation sample, the average poverty likelihood across bootstrap samples for scores of 15–19 is too low by 11.0 percentage points. For scores of 20–24, the estimate is too low by 1.8 percentage points.[9]

The 90-percent confidence interval for the differences for scores of 15–19 is ±7.3 percentage points (Figure 8). This means that in 900 of 1,000 bootstraps, the difference between the estimate and the true value is between –18.3 and –3.7 percentage points (because –11.0 – 7.3 = –18.3, and –11.0 + 7.3 = –3.7). In 950 of 1,000 bootstraps (95 percent), the difference is –11.0 ±7.6 percentage points, and in 990 of 1,000 bootstraps (99 percent), the difference is –11.0 ±8.5 percentage points.

For some 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 Jordan'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 just below the targeting cut-off. This

[9] 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 process of construction and calibration.

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 December 2006 end of field work for the 2006 HIES. That is, the scorecard may fit the data from the 2006 HIES so closely that it captures not only real patterns but also some random patterns that, due to sampling variation, show up only in the 2006 HIES. 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 2006 HIES data but rather also considering experience, judgment, and theory. Of course, the scorecard here does just that. 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 relationships between indicators and poverty, sampling variation, changes in poverty lines, changes in data quality, and inconsistencies in cost-of-living adjustments. These factors can be addressed only by improving data quantity and quality (which is beyond the scope of the scorecard), 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 22.6, 7.7, and 1.5 percent (national line, Figure 5). The group's estimated poverty rate is the households' average poverty likelihood of $(22.6 + 7.7 + 1.5) \div 3 = 10.6$ percent.[10]

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 Jordan scorecard applied to 1,000 bootstrap samples from the validation sample.

Figure 9 summarizes Figure 10 across poverty lines and years for $n = 16{,}384$. It shows that the absolute differences between the estimated poverty rate and the true rate for the scorecard applied to the validation sample are 1.0 percentage points or less. The average absolute difference across the nine poverty lines is 0.4 percentage points.

[10] 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 7.7 percent. This is not the 10.6 percent found as the average of the three poverty likelihoods associated with each of the three scores.

In terms of precision, the 90-percent confidence interval for a group's estimated poverty rate at a point in time in 2006 with n = 16,384 and for all poverty lines is ±0.7 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 0.7 percentage points or less.

In the specific case of the national line, 90 percent of all samples of n = 16,384 produce estimates that differ from the true value in the range of +0.2 – 0.5 = –0.3 to +0.2 + 0.5 = +0.7 percentage points. This is because +0.2 is the average difference and ±0.5 is its 90-percent confidence interval. The average difference is +0.2 because the average scorecard estimate is too high by 0.2 percentage points; the scorecard tends to estimate a poverty rate of 9.6 percent for the validation sample, but the true value is 9.4 percent (Figure 2). Future accuracy will depend on how closely the period of application resembles the second half of 2006.

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 9.4 percent (the true rate in the validation sample for the national line in Figure 2), the confidence interval c is

$$+/- z \cdot \sqrt{\frac{p \cdot (1-p)}{n}} = +/- 1.64 \cdot \sqrt{\frac{0.094 \cdot (1-0.094)}{16{,}384}} = \pm 0.374 \text{ percentage points.}$$

The scorecard, however, does not measure poverty directly, so this formula is not applicable. To derive a formula for the Jordan 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 national line, and the validation sub-sample, the 90-percent confidence

interval is ±0.460 percentage points.[11] Thus, the ratio of confidence intervals with the scorecard versus direct measurement is 0.460 ÷ 0.374 = 1.23.

Now consider the same case, but with n = 8,192. The confidence interval under direct measurement is $+/-1.64 \cdot \sqrt{\frac{0.094 \cdot (1-0.094)}{8{,}192}} = \pm 0.529$ percentage points. The empirical confidence interval with the Jordan scorecard for the national line (Figure 10) is ±0.655 percentage points. Thus for n = 8,192, the ratio is 0.655 ÷ 0.529 = 1.24.

Across all sample sizes of 256 or more in Figure 10, the average ratio turns out to be 1.18, implying that confidence intervals for indirect estimates of poverty rates via the Jordan scorecard and the national poverty line are about 18-percent wider than those for direct estimates. This 1.18 appears in Figure 9 as the "α factor" because if α = 1.18, then the formula relating confidence intervals c and standard errors for the Jordan scorecard is $c = +/- z \cdot \alpha \cdot \sigma$. The standard error 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 greater than 1.00, it means that the scorecard is less precise than direct measurement. This occurs in all nine cases in Figure 9.

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.[12] If

[11] Due to rounding, Figure 10 displays 0.5, not 0.460.

$\hat{p}$ 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 $c = 0.03380$ and $z = 1.64$ (90-percent confidence), and $\hat{p} = 0.0975$ (the average poverty rate for the national line in the construction and calibration sub-samples, Figure 2). Then the formula gives $n = \left(\frac{1.18 \cdot 1.64}{0.03380}\right)^2 \cdot 0.0975 \cdot (1 - 0.0975) = 289$, not too far from the sample size of 256 observed for these parameters in Figure 10.

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

In practice after the end of the HIES field work in December 2006, an organization would select a poverty line (say, the national line), select a desired confidence level (say, 90 percent, or $z = 1.64$), select a desired confidence interval (say, ± 2.0 percentage points, or $c = 0.02$), make an assumption about $\hat{p}$ (perhaps based on a previous measurement such as the 9.6-percent average for the national line in the 2006

[12] 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.

HIES in Figure 2), look up α (here, 1.18), assume that the scorecard is still valid in the future and/or for non-nationally representative sub-groups,[13] and then compute the required sample size. In this illustration, $n = \left(\frac{1.18 \cdot 1.64}{0.02}\right)^2 \cdot 0.096 \cdot (1 - 0.096) = 813$.

[13] This paper reports accuracy for the scorecard applied to the validation sample, but it cannot test accuracy for later years or other groups. Performance will deteriorate with time to the extent that the relationships between indicators and poverty change.

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 2006 HIES, this paper cannot test the accuracy of estimates of change over time for Jordan, 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 22.6, 7.7, and 1.5 percent (national line, Figure 5). The group's baseline estimated poverty rate is the households' average poverty likelihood of (22.6 + 7.7 + 1.5) ÷ 3 = 10.6 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 16.0, 3.3, and 0.2 percent, national line, Figure 5). Their average poverty likelihood at follow-up is (16.0 + 3.3 + 0.2) ÷ 3 = 6.5 percent, an improvement of 10.6 – 6.5 = 4.1 percentage points.[14]

This suggests that about one of 20 participants crossed above the poverty line in 2010. (This is a net figure; some people start above the line and end below it, and vice versa.) Compared with the share who started below the line, more than one in three

[14] 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.

(4.1 ÷ 10.6 = 38.6 percent) 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 Jordan

With only the 2006 HIES, 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 Jordan 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,[15] and α is the average (across a range of sample sizes) of the ratio of the observed

[15] 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.

bootstrap confidence intervals from a scorecard and the theoretical confidence intervals from the textbook formula for direct measurement.

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

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

For countries for which this α has been measured (Schreiner, 2010, 2009a, 2009b, 2009c, 2009d, 2009e, and 2008b; Schreiner and Woller, 2010a and 2010b; and Chen and Schreiner, 2009), the average of α (first averaged across poverty lines and years for a given country, and then averaged across countries) is 1.19. This is as reasonable a figure as any to use for Jordan.

To illustrate the use of the formula above to determine sample size for estimating changes in poverty rates across two independent samples, suppose the desired confidence level is 90 percent ($z = 1.64$), the desired confidence interval is 2 percentage points ($c = 0.02$), the poverty line is the national line, $\alpha = 1.19$, and $\hat{p} = 0.096$ (from Figure 2). Then the baseline sample size is $n = 2 \cdot \left(\frac{1.19 \cdot 1.64}{0.02} \right)^2 \cdot 0.096 \cdot (1 - 0.096) =$ 1,653, and the follow-up sample size is also 1,653.

7.5 Accuracy for estimated change for one sample, scored twice

The general 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:[16]

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

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:

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

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

$\hat{p}_*$ could be anything between 0 to 0.5, 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 Jordan scorecard is applied twice (once after the end of field work for the 2006 HIES 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 national line, and the sample will be scored first in 2010 and then again in 2013 ($y = 3$). The before-baseline poverty rate is 9.6 percent ($p_{2006} = 0.096$, Figure 2), and suppose $\alpha = 1.30$. Then the baseline sample size is $n = 2 \cdot \left(\frac{1.30 \cdot 1.64}{0.02} \right)^2 \cdot \{ -0.02 + 0.016 \cdot 3 + 0.47 \cdot [0.096 \cdot (1 - 0.096)] \} = 1{,}564$. The same group of 1,564 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 19 or less and the scorecard applied to the validation sample, outcomes for the national line are:

- Inclusion: 4.3 percent are below the line and correctly targeted
- Undercoverage: 5.0 percent are below the line and mistakenly not targeted
- Leakage: 4.6 percent are above the line and mistakenly targeted
- Exclusion: 86.1 percent are above the line and correctly not targeted

Increasing the cut-off to 24 or less improves inclusion and undercoverage but worsens leakage and exclusion:

- Inclusion: 6.4 percent are below the line and correctly targeted
- Undercoverage: 3.0 percent are below the line and mistakenly not targeted
- Leakage: 11.5 percent are above the line and mistakenly targeted
- Exclusion: 79.2 percent are above the line and correctly not targeted

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

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

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

- Assign benefits and costs to possible outcomes, based on its values and mission
- Tally total net benefits for each cut-off using Figure 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 and exclusion versus errors of undercoverage and leakage. It is healthy to go through a process of thinking explicitly and intentionally about how possible targeting outcomes are valued.

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

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

Figure 12 shows "Total Accuracy" for all cut-offs for Jordan's scorecard. For the national line in the validation sample, total net benefit is greatest (91.2) for a cut-off of 9 or less, with about nine in ten households in Jordan 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).[17]

[17] Figure 12 also reports "BPAC", the Balanced Poverty Accuracy Criteria adopted 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 targeting inclusion. After normalizing by the number of people below the poverty line, BPAC = (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 households in Jordan who score at or below a given cut-off. For the example of the national line and the validation sample, targeting households who score 19 or less would target 8.9 percent of all households (second column) and lead to a poverty rate among those targeted of 48.7 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 national line and the validation sample with a cut-off of 19 or less, 46.4 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 national line, the validation sample, and a cut-off of 19 or less, covering 0.9 poor households means leaking to 1 non-poor household.

9. Context of poverty-assessment tools for Jordan

The only other poverty-assessment tool for Jordan is Gwatkin *et al.* (2007). They apply an approach used in 56 countries with Demographic and Health Surveys (Rutstein and Johnson, 2004). They use Principal Components Analysis to make an asset index from simple, low-cost indicators available for the 7,335 households in the Jordan 1997 DHS. The PCA index is like the scorecard except that, because the DHS does not collect data on income or consumption, it is based on a different conception of poverty, its accuracy vis-à-vis consumption-based poverty is unknown, and it can only be assumed to be a proxy for long-term wealth/economic status.[18] Well-known examples of the PCA asset-index approach include Sahn and Stifel (2000 and 2003), and Filmer and Pritchett (2001).

The 21 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 floors
  - — Presence of electricity
  - — Source of drinking water
  - — Type of toilet arrangement
  - — Number of people per sleeping room

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

- Ownership of consumer durables:
  - — Radios
  - — Telephones
  - — Televisions
  - — Video players
  - — Satellite dishes
  - — Refrigerators
  - — Air conditioners
  - — Solar water heaters
  - — Bicycles
  - — Motorcycles
  - — Cars
  - — Commercial cars
  - — Pick-up trucks
  - — Agricultural tractors
  - — Other modes of transport
- Whether anyone in the household works their own or family's agricultural land

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 posts 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 similar to those of the scorecard here.

Still, the Gwatkin *et al.* index is more difficult and costly because it cannot be computed by hand in the field, as it has 75 point values, half of them negative, and all with five decimal places.

Unlike the PCA index, the scorecard here 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.

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

10. Conclusion

Pro-poor programs in Jordan 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 Jordan that want to improve how they monitor and manage their social performance.

The scorecard is built with a sub-sample of data from the 2006 HIES, tested on a different sub-sample from the 2006 HIES, and calibrated to nine 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 validation sample with $n = 16{,}384$, the absolute difference between estimates and true poverty rates at a point in time is 1.0 percentage points or less and averages—across the nine poverty lines—0.4 percentage points. With 90-percent confidence, the precision of these differences is ± 0.7 percentage points or less. The scorecard is less 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 Jordan to measure poverty rates, track changes in poverty rates over time, and target services, provided that the scorecard is applied in a time period similar to that of the second half of 2006, the 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.

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 21 October 2010.

Chen, Shiyuan; and Mark Schreiner. (2009) "Simple Poverty Scorecard Poverty-Assessment Tool: Vietnam", `SimplePovertyScorecard.com/BGD_2010_ENG.pdf`, retrieved 9 July 2016.

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

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

Daley-Harris, Sam. (2009) *State of the Microcredit Summit Campaign Report 2009*, `microcreditsummit.org/state_of_the_campaign_report/`, retrieved 21 October 2010.

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

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

Falkenstein, Eric. (2008) "DefProb™: A Corporate Probability-of-Default Model", `papers.ssrn.com/sol3/papers.cfm?abstract_id=1103404`, retrieved 21 October 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 21 October 2009.

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 21 October 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 21 October 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 21 October 2010.

Gwatkin, Davidson R.; Rutstein, Shea; Johnson, Kiersten; Suliman, Eldaw; Wagstaff, Adam; and Agbessi Amouzou. (2007) "Socio-Economic Differences in Health, Nutrition, and Population: Jordan", World Bank Country Reports on HNP and Poverty, go.worldbank.org/T6LCN5A340, retrieved 21 October 2010.

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

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

IRIS Center. (2007a) "Manual for the Implementation of USAID Poverty Assessment Tools", povertytools.org/training_documents/Manuals/USAID_PAT_Manual_Eng.pdf, retrieved 21 October 2010.

_____. (2007b) "Introduction to Sampling for the Implementation of PATs", povertytools.org/training_documents/Sampling/Introduction_Sampling.ppt, retrieved 21 October 2010.

_____. (2005) "Notes on Assessment and Improvement of Tool Accuracy", povertytools.org/other_documents/AssessingImproving_Accuracy.pdf, retrieved 21 October 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 21 October 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 21 October 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, www.mfc.org.pl/doc/Research/ImpAct/SN/MFC_SN04_eng.pdf, retrieved 21 October 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, documents.worldbank.org/curated/en/2005/07/6209268/proxy-means-test-targeting-welfare-benefits-sri-lanka, retrieved 5 May 2016.

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

Rutstein, Shea Oscar; and Kiersten Johnson. (2004) "The DHS Wealth Index", DHS Comparative Reports No. 6, measuredhs.com/pubs/pdf/CR6/CR6.pdf, retrieved 21 October 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 21 October 2010.

Schreiner, Mark. (2013) "Simple Poverty Scorecard Poverty-Assessment Tool: Bangladesh", SimplePovertyScorecard.com/BGD_2010_ENG.pdf, retrieved 9 July 2016.

_____. (2010) "Simple Poverty Scorecard Poverty-Assessment Tool: Honduras", SimplePovertyScorecard.com/HND_2007_ENG.pdf, retrieved 9 July 2016.

_____. (2009a) "Simple Poverty Scorecard Poverty-Assessment Tool: Peru", SimplePovertyScorecard.com/PER_2007_ENG.pdf, retrieved 9 July 2016.

_____. (2009b) "Simple Poverty Scorecard Poverty-Assessment Tool: Philippines", SimplePovertyScorecard.com/PHL_2002_ENG.pdf, retrieved 9 July 2016.

_____. (2009c) "Simple Poverty Scorecard Poverty-Assessment Tool: Pakistan", SimplePovertyScorecard.com/PAK_2005_ENG.pdf, retrieved 9 July 2016.

_____. (2009d) "Simple Poverty Scorecard Poverty-Assessment Tool: Bolivia", SimplePovertyScorecard.com/BOL_2007_ENG.pdf, retrieved 9 July 2016.

_____. (2009e) "Simple Poverty Scorecard Poverty-Assessment Tool: Mexico", SimplePovertyScorecard.com/MEX_2008_ENG.pdf, retrieved 9 July 2016.

_____. (2008a) "Simple Poverty Scorecard Poverty-Assessment Tool: Peru", SimplePovertyScorecard.com/PER_2003_ENG.pdf, retrieved 9 July 2016.

_____. (2008b) "Simple Poverty Scorecard Poverty-Assessment Tool: India", SimplePovertyScorecard.com/IND_2005_ENG.pdf, retrieved 9 July 2016.

_____. (2006a) "La Herramienta del Índice de Calificación de la Pobreza™: México", SimplePovertyScorecard.com/MEX_2002_SPA.pdf, retrieved 9 July 2016.

_____. (2006b) "Is One Simple Poverty Scorecard Poverty-Assessment Tool Enough for India?", microfinance.com/English/Papers/Scoring_Poverty_India_Segments.pdf, retrieved 21 October 2010.

_____. (2005) "IRIS Questions on the Simple Poverty Scorecard Poverty-Assessment Tool", microfinance.com/English/Papers/coring_Poverty_Response_to_IRIS.pdf, retrieved 21 October 2010.

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

_____; Matul, Michal; Pawlak, Ewa; and Sean Kline. (2004) "Poverty Scoring: Lessons from a Microlender in Bosnia-Herzegovina", microfinance.com/English/apers/Scoring_Poverty_in_BiH_Short.pdf, retrieved 21 October 2010.

_____; and Gary Woller. (2010a) "Simple Poverty Scorecard Poverty-Assessment Tool: Ghana", SimplePovertyScorecard.com/GHA_2005_ENG.pdf, retrieved 9 July 2016.

_____; and Gary Woller. (2010b) "Simple Poverty Scorecard®: Guatemala", SimplePovertyScorecard.com/GTM_2006_ENG.pdf, retrieved 9 July 2016.

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

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 21 October 2010.

Toohig, Jeff. (2008) "PPI Pilot Training Guide", `progressoutofpoverty.org/toolkit`, retrieved 32 October 2010.

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

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. (2008a) *Jordan Poverty Update Volume II: Appendices*, `go.worldbank.org/O2GAJ0JGE1`, retrieved 21 October 2010.

_____. (2008b) "International Comparison Project: Tables of Results", `siteresources.worldbank.org/ICPINT/Resources/icp-final-tables.pdf`, retrieved 21 October 2010.

Zeller, Manfred. (2004) "Review of Poverty Assessment Tools", `pdf.usaid.gov/pdf_docs/PNADH120.pdf`, retrieved 13 May 2016.

Figure 2: Poverty lines and poverty rates

	Sample size	Poverty rates (% with expenditure below a poverty line)								
		National line				USAID	Intl. 2005 PPP (per person)			
		100%	150%	200%	250%	'Extreme'	$1.25/day	$2.50/day	$3.75/day	$5.00/day
All Jordan:										
Poverty line (JOD/person/day)	N/A	1.52	2.29	3.05	3.81	1.23	0.66	1.32	1.98	2.64
Household-level poverty rate	23,278	9.6	30.0	50.2	64.2	4.7	0.2	5.7	21.2	40.1
Person-level poverty rate	23,278	13.0	36.9	58.2	71.5	6.7	0.4	7.9	26.9	47.7
Household-level poverty rates in scoring sub-samples										
Construction										
Selecting indicators and points	7,695	9.5	29.8	50.1	63.3	4.8	0.2	5.5	20.8	39.9
Calibration										
Associating scores with likelihoods	7,758	10.0	30.4	50.6	64.9	4.9	0.2	6.3	21.9	40.5
Validation										
Measuring accuracy	7,825	9.4	29.7	50.0	64.4	4.5	0.2	5.4	21.1	39.9
Change in poverty rate (percentage points)										
Construction/calibration to validation	N/A	+0.4	+0.4	+0.3	–0.3	+0.3	–0.0	+0.5	+0.3	+0.3

Source: 2006 HIES.

Figure 3: Poverty lines and poverty rates for All-Jordan and by governorate, household- and person-level

	Line or		Poverty line (JOD/person/day) and poverty rate (%)								
			National line				USAID	Intl. 2005 PPP (per person)			
Governorate	rate	Level	100%	150%	200%	250%	'Extreme'	$1.25/day	$2.50/day	$3.75/day	$5.00/day
All Jordan	Line		1.52	2.29	3.05	3.81	1.23	0.66	1.32	1.98	2.64
	Rate	Household	9.6	30.0	50.2	64.2	4.7	0.2	5.7	21.2	40.1
	Rate	Person	13.0	36.9	58.2	71.5	6.7	0.4	7.9	26.9	47.7
Amman	Line		1.59	2.39	3.18	3.98	1.30	0.69	1.38	2.07	2.76
	Rate	Household	6.7	22.8	39.6	53.2	3.3	0.1	3.7	15.6	31.0
	Rate	Person	9.4	29.4	47.8	61.4	4.9	0.2	5.3	20.7	38.5
Balqa	Line		1.49	2.23	2.97	3.71	1.22	0.64	1.29	1.93	2.57
	Rate	Household	11.0	33.6	53.8	68.3	5.7	0.1	6.5	24.4	44.8
	Rate	Person	15.3	41.5	62.0	75.5	8.7	0.2	9.6	31.3	52.7
Zarqa	Line		1.50	2.25	3.00	3.76	1.25	0.65	1.30	1.95	2.60
	Rate	Household	10.9	36.0	62.0	73.8	5.2	0.1	5.9	25.0	49.6
	Rate	Person	14.9	43.8	70.1	80.8	7.3	0.3	8.4	31.4	57.8
Madaba	Line		1.48	2.23	2.97	3.71	1.17	0.64	1.28	1.93	2.57
	Rate	Household	7.8	28.2	44.8	61.0	3.9	0.0	5.1	16.6	34.6
	Rate	Person	10.0	35.6	55.0	69.3	4.9	0.1	6.4	20.6	44.0
Irbid	Line		1.5	2.2	3.0	3.7	1.2	0.6	1.3	1.9	2.6
	Rate	Household	9.7	30.7	52.3	66.7	4.9	0.3	6.1	22.9	41.2
	Rate	Person	12.1	36.1	58.9	72.6	6.3	0.4	7.8	27.3	47.3
Mafraq	Line		1.4	2.2	2.9	3.6	1.1	0.6	1.2	1.9	2.5
	Rate	Household	12.9	35.3	59.5	74.1	6.8	1.3	11.0	25.8	43.2
	Rate	Person	15.4	38.5	64.4	78.8	8.6	1.4	13.0	29.8	46.7

Figure 3 (cont.): Poverty lines and poverty rates for All-Jordan and by governorate, household- and person-level

	Line or		Poverty line (JOD/person/day) and poverty rate (%)								
			National line				USAID	Intl. 2005 PPP (per person)			
Governorate	rate	Level	100%	150%	200%	250%	'Extreme'	$1.25/day	$2.50/day	$3.75/day	$5.00/day
Jarash	Line		1.5	2.2	2.9	3.7	1.1	0.6	1.3	1.9	2.5
	Rate	Household	12.5	36.4	54.4	70.1	6.8	0.1	7.9	27.3	47.3
	Rate	Person	16.7	43.0	62.9	77.7	8.7	0.1	10.2	35.3	55.8
Ajlun	Line		1.5	2.2	2.9	3.6	1.1	0.6	1.3	1.9	2.5
	Rate	Household	14.6	40.3	66.2	77.7	7.1	0.4	9.6	28.7	55.4
	Rate	Person	17.7	44.6	72.4	82.4	8.6	0.5	11.2	32.4	61.6
Karak	Line		1.5	2.2	3.0	3.7	1.1	0.6	1.3	1.9	2.6
	Rate	Household	14.8	36.6	60.4	79.3	7.3	0.2	10.0	27.2	49.2
	Rate	Person	21.7	46.1	67.3	85.6	11.2	0.3	14.6	35.7	57.7
Tafiela	Line		1.5	2.2	3.0	3.7	1.2	0.6	1.3	1.9	2.6
	Rate	Household	14.6	48.0	70.9	86.2	6.5	0.0	7.9	33.3	58.3
	Rate	Person	19.1	56.7	78.2	89.9	9.3	0.0	11.2	40.6	67.0
Ma'an	Line		1.5	2.2	2.9	3.7	1.1	0.6	1.3	1.9	2.5
	Rate	Household	9.6	32.7	53.5	68.6	4.3	0.3	6.1	22.9	44.4
	Rate	Person	12.7	41.7	65.9	79.2	6.2	0.5	8.2	29.3	55.9
Aqaba	Line		1.5	2.2	3.0	3.7	1.1	0.6	1.3	1.9	2.6
	Rate	Household	12.9	35.3	59.5	74.1	6.8	1.3	11.0	25.8	43.2
	Rate	Person	15.4	38.5	64.4	78.8	8.6	1.4	13.0	29.8	46.7

Source: 2006 HIES.

Figure 4: Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly indicative of poverty)
760	Does the household own a microwave? (No; Yes)
666	Does the residence have a dining room? (No; Yes)
644	Does the household own a vacuum cleaner? (No; Yes)
623	Does the household own a combined gas stove with oven? (No; Yes)
519	Does the household own a gas stove (separate from any gas oven)? (Yes; No)
511	Does the household own a private car? (No; Yes)
495	Does the household have a refrigerator and/or freezer? (None; Refrigerator, but no freezer; Freezer, but no refrigerator; Both)
457	What is the main construction material of the exterior walls of the residence? (Bricks, or other; Concrete; Stone and concrete; Cut stone)
398	Does the household have a computer connected to the internet? (No; Only computer; Computer and internet)
396	Does the household have any land-line telephones? (No; Yes)
367	What is the source of heating for the residence? (Wood, none, or other; Kerosene; Gas, electric, or central)
361	Does the household own any video players (CDs, VCRs, or DVDs)? (No; Yes)
360	What type of bathroom does the household have? (Flush to public sewer system, other, or none; Toilet with shower; More than one toilet and/or more than one shower)
342	How many household members are 17-years-old or younger? (Five or more; Four; Three; Two; One; None)
340	How many household members are 18-years-old or younger? (Six or more; Five; Four; Three; Two; One; None)
333	How many household members are 16-years-old or younger? (Five or more; Four; Three; Two; One; None)
325	How many household members are 15-years-old or younger? (Five or more; Four; Three; Two; One; None)

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

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly associated with poverty)
324	Does the residence of the household have a guest room? (No; Yes)
324	Does any household member work as a legislator, senior official, manager, professional, technician, or associated professional? (No; Yes)
308	How many land-line and/or mobile telephones does the household have? (None; One mobile, but no land-lines; One or more land-lines, but no mobile; One or more landlines, and one mobile; Two or more mobiles, but no land-lines; One or more landlines, and two or more mobiles)
307	How many household members are 14-years-old or younger? (Five or more; Four; Three; Two; One; None)
305	Does the household own a combined gas stove with oven? (No; Yes)
299	How many members does the household have? (Nine or more; Eight; Seven; Six; Five; Four; Three; One or two)
291	How many household members are 13-years-old or younger? (Five or more; Four; Three; Two; One; None)
285	Does the household have a water filter? (No; Yes)
271	How many household members are 12-years-old or younger? (Four or more; Three; Two; One; None)
266	Does the household own any satellite receivers? (No; Yes)
265	What is the rent (in dinar) that the household pays (or that it would pay, if it paid rent for this same house)? (Up to 49; 50 to 59; 60 to 69; 70 to 89; 90 to 109; 110 or more)
261	Does the household own any satellite receivers and/or video players (CDs, VCRs, or DVDs)? (No; Satellite receiver, but no video player; Video player, but no satellite receiver; Both)
259	Does the household own an air conditioner? (No; Yes)
254	How many household members are 11-years-old or younger? (Four or more; Three; Two; One; None)
253	How many rooms (not counting the kitchen) does the residence of the household have? (One or two; Three; Four; Five; Six or more)
245	Do any household members work as legislators, senior officials, managers, professionals, technicians, associated professionals, or clerks? (No; Yes)

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

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly associated with poverty)
244	Does the household have a solar water heater? (No; Yes)
238	What is the source of drinking water for the household? (Not mineral water; Mineral water)
236	Does the residence have a sitting room? (No; Yes)
223	How many mobile telephones does the household have? (None; One; Two or more)
211	How many bedrooms does the residence of the household have? (One; Two; Three or more)
196	Does the household own a television? (No; Yes)
187	How many household members are from a country other than Jordan? (Two or more; One; None)
186	Does the household own a radio or a radio-cassette player? (No; Yes)
186	What is the highest level of education completed by any family member? (Preparatory; Basic, or vocational apprenticeship; Literate, or elementary; Secondary; Intermediate bachelor; Illiterate; Bachelor, higher diploma, MA/MSc, or PhD)
170	How many household members are 11-years-old or younger? (Three or more; Two; One; None)
170	What type of residence does the household live in? (*Dar*; Villa, or apartment)
160	What is the area (in square meters) of the residence? (Up to 79; 80 to 109; 110 to 129; 130 to 169; 170 or more)
149	What is the structure of household headship? (Male head/spouse with more than one female head/spouse; Both male and female heads/spouses; Only female head/spouse; Only male head/spouse)
144	Is any household member from an Arab country other than Egypt, Syria, or Iraq? (Yes; No)
139	What is the highest level of education completed by the female head/spouse? (More than one female head/spouse; Illiterate; Elementary; Preparatory; Basic; Vocational apprenticeship; Literate; Secondary; Intermediate bachelor; No female head/spouse; Bachelor, higher diploma, MA/MSc, PhD)

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

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly associated with poverty)
124	What is the highest level of education completed by the male head/spouse? (Illiterate; Elementary; Preparatory, basic, or vocational apprenticeship; Literate; Secondary; Intermediate bachelor; No male head/spouse; Bachelor, higher diploma, MA/MSc, PhD)
113	What is the main occupation of the male head/spouse? (Other; Does not work, or armed forces; Legislators, senior managers, managers, or professionals)
100	In what sector does the female head/spouse work? (More than one female head/spouse; Does not work; Works; No female head/spouse)
89	How does the household dispose of liquid wastes? (Not public sewer network; Public sewer network)
85	Do any household members work in elementary occupations? (Yes; No)
79	What is the marital status of the male head/spouse? (Married; Not married)
77	How old is the female head/spouse? More than one female head/spouse; 30 to 39; 40 to 44; 25 to 29; 45 to 49; 25 or younger; 50 to 55; 56 or older; No female head/spouse)
75	What is the country of origin of the female head/spouse? (Palestine, Egypt, or other; Jordan; More than one female head/spouse; Syria, Iraq, or No female head/spouse)
72	Do any household members work in elementary occupations or as skilled workers in agriculture and fisheries? (Yes; No)
69	Did the female head/spouse work at least one hour in the past week? (No; More than one female head/spouse; Yes; No female head/spouse)
64	How old is the male head/spouse? (35 to 44; 30 to 34; 45 to 49; 50 to 54; 29 or younger; 55 or older; No male head/spouse)
55	What is the country of origin of the male head/spouse? (Egypt, or Palestine; Jordan; No male head/spouse; Syria, Iraq, or Other)

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

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly associated with poverty)
48	What is the marital status of the female head/spouse? (Married; More than one female head/spouse; Other)
38	What is the employment status of the male head/spouse in his main line of work? (Self-employed without employees, unpaid family worker, or unpaid apprentice; Salaried; Does not work; No male head/spouse; Self-employed with employees)
38	Does the household own a sewing machine? (No; Yes)
34	What is the tenancy status of the household in its residence? (Not rent or own; Rent; Own)
28	Does the household live in a rural area? (Yes; No)
27	What is the main economic activity of the male head/spouse? (Does not work; Construction, or manufacturing; Other; Wholesale and retail trade, repair of vehicles and goods, or public administration and defense, including compulsory social security; No male head/spouse)
25	Does the household own a gas oven (separate from any gas stove)? (Yes; No)
20	In what sector does the male head/spouse work? (Does not work; Public; Private; International, outside, or no male head/spouse)
13	How many household members worked at least one hour in the past week? (None; One; Two; Three or more)
10	Are any household members are self-employed without employees? (Yes; No)
7	How many household members are salaried or are self-employed with employees? (None; One; Two or more)
5	Do any household members work in the public sector? (No; Yes)
3	Did the male head/spouse work at least one hour in the past week? (No; Yes; No male head/spouse)

Source: 2006 HIES and the 150% of the national poverty line.

100% of the National Poverty Line

(and tables pertaining to all poverty lines)

Figure 5 (100% of national line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	86.7
5–9	65.7
10–14	63.7
15–19	39.5
20–24	22.6
25–29	16.0
30–34	7.7
35–39	3.3
40–44	1.5
45–49	0.2
50–54	0.9
55–59	0.0
60–64	0.0
65–69	0.0
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Figure 6 (100% of national line): Derivation of estimated poverty likelihoods associated with scores

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	287	÷	331	=	86.7
5–9	694	÷	1,056	=	65.7
10–14	1,602	÷	2,515	=	63.7
15–19	1,978	÷	5,010	=	39.5
20–24	2,021	÷	8,959	=	22.6
25–29	1,562	÷	9,742	=	16.0
30–34	899	÷	11,617	=	7.7
35–39	432	÷	13,061	=	3.3
40–44	174	÷	11,230	=	1.5
45–49	19	÷	9,860	=	0.2
50–54	77	÷	8,565	=	0.9
55–59	0	÷	6,328	=	0.0
60–64	0	÷	4,789	=	0.0
65–69	0	÷	3,004	=	0.0
70–74	0	÷	2,025	=	0.0
75–79	0	÷	839	=	0.0
80–84	0	÷	732	=	0.0
85–89	0	÷	269	=	0.0
90–94	0	÷	55	=	0.0
95–100	0	÷	12	=	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 adjacent poverty lines									
	<$1.25/day	=>$1.25/day and <USAID	=>USAID and <$2.50/day	=>$2.50/day and <100% Natl.	=>100% Natl. and <$3.75/day	=>$3.75/day and <150% Natl.	<150% Natl. and <$5.00/day	=>$5.00/day and <200% Natl.	=>200% Natl. and <250% Natl.	=>250% Natl.
Score	<JOD0.66	=>JOD0.66 and <JOD1.23	=>JOD1.23 and <JOD1.32	=>JOD1.32 and <JOD1.52	=>JOD1.52 and <JOD1.98	=>JOD1.98 and <JOD2.29	=>JOD2.29 and <JOD2.64	=>JOD2.64 and <JOD3.05	=>JOD3.05 and <JOD3.81	=>JOD3.81
0–4	21.8	63.9	0.0	1.0	13.3	0.0	0.0	0.0	0.0	0.0
5–9	2.7	42.1	6.9	14.1	24.7	4.0	2.2	2.6	0.8	0.0
10–14	1.2	41.5	4.0	17.1	15.0	8.3	6.5	4.4	1.4	0.8
15–19	2.1	20.2	4.4	12.9	28.7	10.3	12.1	5.5	2.3	1.5
20–24	0.0	9.5	4.5	8.6	31.1	12.5	11.3	9.1	7.4	5.9
25–29	0.0	6.6	2.2	7.2	21.1	14.9	11.2	11.3	15.1	10.5
30–34	0.0	2.1	1.7	4.0	19.6	15.6	17.0	13.6	14.2	12.1
35–39	0.0	0.7	1.4	1.3	10.7	13.3	15.6	14.4	19.0	23.8
40–44	0.0	0.5	0.0	1.1	7.0	8.3	13.5	16.1	22.1	31.5
45–49	0.0	0.0	0.2	0.0	3.1	3.6	9.0	13.4	21.4	49.3
50–54	0.0	0.0	0.2	0.7	0.5	2.2	6.3	10.1	19.8	60.2
55–59	0.0	0.0	0.0	0.0	0.1	2.0	3.8	3.6	15.1	75.4
60–64	0.0	0.0	0.0	0.0	0.1	0.5	0.7	2.9	11.0	84.7
65–69	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.8	4.7	94.2
70–74	0.0	0.0	0.0	0.0	0.0	0.0	0.4	1.6	2.5	95.5
75–79	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
80–84	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
85–89	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
90–94	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
95–100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0

All poverty likelihoods in percentage units.

Figure 8 (100% of national line): Bootstrapped differences between estimated and true household poverty likelihoods with confidence intervals in a large sample (n = 16,384), scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Score	**Diff.**	**90-percent**	**95-percent**	**99-percent**
0–4	+14.3	11.7	14.0	17.3
5–9	+15.2	9.0	10.8	13.1
10–14	+14.3	5.6	6.8	8.7
15–19	–11.0	7.3	7.6	8.5
20–24	–1.8	2.5	2.9	4.0
25–29	+3.9	1.7	1.9	2.6
30–34	+0.1	1.1	1.3	1.8
35–39	+0.1	0.6	0.8	1.0
40–44	+0.5	0.4	0.5	0.6
45–49	–0.7	0.5	0.6	0.6
50–54	+0.8	0.1	0.1	0.1
55–59	+0.0	0.0	0.0	0.0
60–64	+0.0	0.0	0.0	0.0
65–69	+0.0	0.0	0.0	0.0
70–74	+0.0	0.0	0.0	0.0
75–79	+0.0	0.0	0.0	0.0
80–84	+0.0	0.0	0.0	0.0
85–89	+0.0	0.0	0.0	0.0
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 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, scorecard applied to the validation sample

	Poverty line									
	National line				USAID	Intl. 2005 PPP (per person)				
	100%	150%	200%	250%	'Extreme'	$1.25/day	$2.50/day	$3.75/day	$5.00/day
Estimate minus true value										
Scorecard applied to validation sample	+0.2	–0.1	–0.5	–1.0	+0.2	+0.0	+0.7	+0.4	+0.1
Precision of difference										
Scorecard applied to validation sample	0.5	0.6	0.7	0.6	0.4	0.1	0.4	0.5	0.7
α for sample size										
Scorecard applied to validation sample	1.18	1.07	1.03	1.03	1.29	1.32	1.31	1.01	1.06
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 (100% of 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, scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+1.1	50.0	61.7	80.7
4	+0.6	25.4	32.8	46.1
8	+0.1	18.4	23.0	33.1
16	+0.0	12.5	16.0	21.7
32	+0.3	9.6	11.4	15.0
64	+0.4	6.9	8.1	10.9
128	+0.1	4.9	5.7	7.4
256	+0.1	3.4	3.9	5.1
512	+0.2	2.5	3.0	3.9
1,024	+0.2	1.8	2.0	2.7
2,048	+0.2	1.2	1.5	2.0
4,096	+0.2	0.8	1.0	1.3
8,192	+0.2	0.7	0.8	1.0
16,384	+0.2	0.5	0.5	0.7

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 (100% of national line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.2	9.1	0.1	90.6	90.8	–93.9
5–9	1.0	8.4	0.4	90.2	91.2	–74.9
10–14	2.2	7.2	1.7	88.9	91.1	–35.2
15–19	4.3	5.0	4.6	86.1	90.4	+41.6
20–24	6.4	3.0	11.5	79.2	85.6	–22.5
25–29	7.7	1.7	19.9	70.7	78.4	–112.9
30–34	8.6	0.8	30.7	60.0	68.5	–227.7
35–39	9.1	0.3	43.2	47.4	56.5	–361.7
40–44	9.2	0.1	54.3	36.3	45.6	–480.0
45–49	9.3	0.0	64.0	26.6	35.9	–584.1
50–54	9.4	0.0	72.6	18.1	27.4	–675.4
55–59	9.4	0.0	78.9	11.7	21.1	–743.0
60–64	9.4	0.0	83.7	6.9	16.3	–794.1
65–69	9.4	0.0	86.7	3.9	13.3	–826.2
70–74	9.4	0.0	88.7	1.9	11.3	–847.8
75–79	9.4	0.0	89.6	1.1	10.4	–856.8
80–84	9.4	0.0	90.3	0.3	9.7	–864.6
85–89	9.4	0.0	90.6	0.1	9.4	–867.5
90–94	9.4	0.0	90.6	0.0	9.4	–868.1
95–100	9.4	0.0	90.6	0.0	9.4	–868.2

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

Figure 13 (100% of national line): Households below the poverty line and all households at a given score or at or below a given score cut-off, scorecard applied to the 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	0.3	73.6	2.6	2.8:1
5–9	1.4	69.5	10.3	2.3:1
10–14	3.9	55.4	23.1	1.2:1
15–19	8.9	48.7	46.4	0.9:1
20–24	17.9	35.8	68.4	0.6:1
25–29	27.6	27.8	82.0	0.4:1
30–34	39.2	21.8	91.4	0.3:1
35–39	52.3	17.3	96.9	0.2:1
40–44	63.5	14.5	98.6	0.2:1
45–49	73.4	12.7	99.8	0.1:1
50–54	81.9	11.4	100.0	0.1:1
55–59	88.3	10.6	100.0	0.1:1
60–64	93.1	10.1	100.0	0.1:1
65–69	96.1	9.7	100.0	0.1:1
70–74	98.1	9.5	100.0	0.1:1
75–79	98.9	9.5	100.0	0.1:1
80–84	99.7	9.4	100.0	0.1:1
85–89	99.9	9.4	100.0	0.1:1
90–94	100.0	9.4	100.0	0.1:1
95–100	100.0	9.4	100.0	0.1:1

150% of the National Poverty Line

Figure 5 (150% of national line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	94.4
10–14	86.9
15–19	78.5
20–24	66.2
25–29	52.0
30–34	43.0
35–39	27.3
40–44	16.8
45–49	6.9
50–54	3.6
55–59	2.1
60–64	0.7
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 (150% of national line): Bootstrapped differences between estimated and true household poverty likelihoods with confidence intervals in a large sample (n = 16,384), scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	**Diff.**	**90-percent**	**95-percent**	**99-percent**
0–4	+0.0	0.0	0.0	0.0
5–9	–3.1	2.2	2.3	2.4
10–14	+0.7	2.9	3.4	4.5
15–19	–6.2	4.3	4.6	4.9
20–24	+0.2	2.6	3.0	4.4
25–29	+2.5	2.7	3.1	3.9
30–34	+4.3	2.4	2.8	3.4
35–39	–0.7	1.9	2.2	3.1
40–44	+0.7	1.9	2.2	2.9
45–49	–1.4	1.3	1.6	1.9
50–54	–5.0	3.5	3.7	4.1
55–59	+1.1	0.6	0.7	0.8
60–64	–0.2	0.7	0.8	1.1
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 (150% of 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, scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+1.1	62.4	74.7	87.5
4	+0.3	36.6	45.4	56.5
8	+0.2	26.3	32.8	42.6
16	+0.4	18.6	21.9	29.6
32	+0.2	14.4	17.0	22.3
64	+0.1	9.9	12.1	15.2
128	–0.1	6.9	8.3	11.3
256	–0.1	4.9	6.0	8.2
512	–0.1	3.7	4.3	5.4
1,024	–0.1	2.6	3.1	4.0
2,048	–0.1	1.8	2.1	2.8
4,096	–0.1	1.2	1.5	1.9
8,192	–0.2	0.9	1.0	1.3
16,384	–0.1	0.6	0.7	0.9

Figure 12 (150% of national line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.3	29.4	0.0	70.3	70.6	–97.8
5–9	1.3	28.4	0.1	70.2	71.6	–90.8
10–14	3.4	26.3	0.5	69.8	73.2	–75.5
15–19	7.4	22.3	1.5	68.8	76.2	–45.0
20–24	13.2	16.5	4.6	65.6	78.9	+4.6
25–29	17.9	11.8	9.7	60.6	78.6	+53.3
30–34	22.6	7.1	16.6	53.7	76.4	+44.2
35–39	26.4	3.3	25.9	44.4	70.8	+12.9
40–44	28.2	1.5	35.3	35.0	63.2	–18.8
45–49	29.1	0.6	44.2	26.0	55.2	–48.9
50–54	29.6	0.1	52.3	18.0	47.6	–76.1
55–59	29.7	0.0	58.6	11.7	41.4	–97.2
60–64	29.7	0.0	63.3	6.9	36.7	–113.2
65–69	29.7	0.0	66.4	3.9	33.6	–123.3
70–74	29.7	0.0	68.4	1.9	31.6	–130.1
75–79	29.7	0.0	69.2	1.1	30.8	–132.9
80–84	29.7	0.0	69.9	0.3	30.1	–135.4
85–89	29.7	0.0	70.2	0.1	29.8	–136.3
90–94	29.7	0.0	70.3	0.0	29.7	–136.5
95–100	29.7	0.0	70.3	0.0	29.7	–136.5

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

Figure 13 (150% of national line): Households below the poverty line and all households at a given score or at or below a given score cut-off, scorecard applied to the 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	0.3	100.0	1.1	Only poor targeted
5–9	1.4	96.1	4.5	24.6:1
10–14	3.9	86.8	11.4	6.6:1
15–19	8.9	83.4	25.0	5.0:1
20–24	17.9	74.0	44.5	2.8:1
25–29	27.6	65.0	60.4	1.9:1
30–34	39.2	57.7	76.2	1.4:1
35–39	52.3	50.5	88.9	1.0:1
40–44	63.5	44.4	95.0	0.8:1
45–49	73.4	39.7	98.0	0.7:1
50–54	81.9	36.1	99.7	0.6:1
55–59	88.3	33.6	99.9	0.5:1
60–64	93.1	31.9	100.0	0.5:1
65–69	96.1	30.9	100.0	0.4:1
70–74	98.1	30.3	100.0	0.4:1
75–79	98.9	30.0	100.0	0.4:1
80–84	99.7	29.8	100.0	0.4:1
85–89	99.9	29.7	100.0	0.4:1
90–94	100.0	29.7	100.0	0.4:1
95–100	100.0	29.7	100.0	0.4:1

200% of the National Poverty Line

Figure 5 (200% of national line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	99.2
10–14	97.8
15–19	96.2
20–24	86.7
25–29	74.4
30–34	73.6
35–39	57.2
40–44	46.4
45–49	29.3
50–54	20.0
55–59	9.5
60–64	4.2
65–69	1.1
70–74	2.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Figure 8 (200% of national line): Bootstrapped differences between estimated and true household poverty likelihoods with confidence intervals in a large sample (n = 16,384), scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+0.0	0.0	0.0	0.0
5–9	–0.2	0.6	0.7	0.8
10–14	+4.4	2.2	2.5	3.3
15–19	+0.6	1.1	1.3	1.7
20–24	–3.7	2.5	2.6	2.8
25–29	–6.1	4.0	4.2	4.6
30–34	+3.5	2.1	2.5	3.4
35–39	–0.8	2.2	2.6	3.6
40–44	+0.8	2.4	2.8	3.8
45–49	+1.5	2.3	2.7	3.4
50–54	–3.3	2.9	3.1	3.8
55–59	+1.0	1.8	2.1	2.8
60–64	+0.2	1.3	1.6	2.0
65–69	–3.6	3.1	3.4	4.2
70–74	+2.0	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 (200% of 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, scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+0.3	63.6	72.6	93.4
4	+0.4	39.4	45.5	55.6
8	–0.1	27.6	32.9	42.8
16	+0.1	20.2	24.7	32.3
32	–0.1	15.0	17.7	25.1
64	–0.3	10.2	12.5	16.1
128	–0.5	7.3	8.6	11.3
256	–0.5	5.3	6.3	8.2
512	–0.5	3.7	4.4	5.6
1,024	–0.5	2.7	3.2	4.1
2,048	–0.5	1.9	2.2	3.0
4,096	–0.5	1.3	1.5	1.9
8,192	–0.5	0.9	1.1	1.4
16,384	–0.5	0.7	0.8	1.0

ecked

Figure 12 (200% of national line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.3	49.7	0.0	50.0	50.3	–98.7
5–9	1.4	48.7	0.0	49.9	51.3	–94.5
10–14	3.7	46.4	0.2	49.7	53.4	–84.9
15–19	8.4	41.7	0.6	49.4	57.8	–65.5
20–24	16.2	33.9	1.7	48.2	64.4	–32.0
25–29	23.9	26.2	3.8	46.2	70.1	+2.9
30–34	32.2	17.9	7.1	42.9	75.0	+42.6
35–39	39.6	10.4	12.7	37.3	76.9	+74.7
40–44	44.8	5.3	18.8	31.2	76.0	+62.5
45–49	47.6	2.4	25.8	24.2	71.8	+48.5
50–54	49.2	0.8	32.7	17.2	66.5	+34.6
55–59	49.8	0.3	38.5	11.5	61.2	+23.1
60–64	50.0	0.1	43.1	6.9	56.9	+13.9
65–69	50.0	0.0	46.0	3.9	54.0	+8.0
70–74	50.0	0.0	48.0	1.9	52.0	+4.0
75–79	50.0	0.0	48.9	1.1	51.1	+2.3
80–84	50.0	0.0	49.6	0.3	50.4	+0.9
85–89	50.0	0.0	49.9	0.1	50.1	+0.3
90–94	50.0	0.0	49.9	0.0	50.1	+0.2
95–100	50.0	0.0	50.0	0.0	50.0	+0.2

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

Figure 13 (200% of national line): Households below the poverty line and all households at a given score or at or below a given score cut-off, scorecard applied to the 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	0.3	100.0	0.7	Only poor targeted
5–9	1.4	98.7	2.7	78.6:1
10–14	3.9	94.2	7.3	16.2:1
15–19	8.9	93.8	16.7	15.1:1
20–24	17.9	90.5	32.3	9.5:1
25–29	27.6	86.4	47.7	6.4:1
30–34	39.2	82.0	64.2	4.5:1
35–39	52.3	75.8	79.2	3.1:1
40–44	63.5	70.5	89.5	2.4:1
45–49	73.4	64.9	95.2	1.8:1
50–54	81.9	60.1	98.4	1.5:1
55–59	88.3	56.4	99.5	1.3:1
60–64	93.1	53.7	99.9	1.2:1
65–69	96.1	52.1	100.0	1.1:1
70–74	98.1	51.0	100.0	1.0:1
75–79	98.9	50.6	100.0	1.0:1
80–84	99.7	50.2	100.0	1.0:1
85–89	99.9	50.1	100.0	1.0:1
90–94	100.0	50.1	100.0	1.0:1
95–100	100.0	50.0	100.0	1.0:1

250% of the National Poverty Line

Figure 5 (250% of national line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	99.2
15–19	98.5
20–24	94.1
25–29	89.5
30–34	87.9
35–39	76.2
40–44	68.5
45–49	50.7
50–54	39.8
55–59	24.6
60–64	15.3
65–69	5.8
70–74	4.5
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Figure 8 (250% of national line): Bootstrapped differences between estimated and true household poverty likelihoods with confidence intervals in a large sample (n = 16,384), scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+0.0	0.0	0.0	0.0
5–9	+0.2	0.2	0.3	0.4
10–14	+0.9	0.9	1.1	1.5
15–19	+0.7	0.9	1.0	1.3
20–24	–2.3	1.5	1.6	1.8
25–29	–3.9	2.5	2.6	2.8
30–34	+5.8	2.0	2.3	3.1
35–39	–0.9	1.7	2.0	2.6
40–44	–1.1	2.0	2.4	3.3
45–49	–4.4	3.5	3.7	4.1
50–54	–6.3	4.6	4.8	5.4
55–59	–0.3	3.2	3.7	4.6
60–64	–0.1	2.6	3.1	4.2
65–69	–1.1	2.5	3.0	4.1
70–74	+4.2	0.4	0.4	0.6
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 (250% of 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, scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+0.7	62.8	74.0	89.4
4	+0.2	36.0	44.3	59.9
8	+0.0	26.9	32.5	42.1
16	–0.7	18.4	22.6	29.0
32	–0.9	14.4	16.9	21.3
64	–1.1	9.8	11.5	15.8
128	–1.1	7.0	8.1	11.0
256	–1.0	5.3	6.4	8.1
512	–1.0	3.7	4.3	5.7
1,024	–1.0	2.6	3.0	3.9
2,048	–1.0	1.7	2.1	2.8
4,096	–1.0	1.2	1.4	1.9
8,192	–1.0	0.9	1.0	1.3
16,384	–1.0	0.6	0.7	0.9

Figure 12 (250% of national line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.3	64.0	0.0	35.6	36.0	–99.0
5–9	1.4	63.0	0.0	35.6	37.0	–95.7
10–14	3.8	60.5	0.1	35.6	39.4	–88.0
15–19	8.7	55.7	0.2	35.4	44.1	–72.6
20–24	17.2	47.1	0.7	35.0	52.2	–45.5
25–29	26.2	38.2	1.4	34.2	60.4	–16.4
30–34	36.1	28.3	3.1	32.5	68.6	+17.1
35–39	45.9	18.5	6.4	29.3	75.2	+52.6
40–44	53.4	11.0	10.2	25.5	78.8	+81.6
45–49	58.7	5.7	14.7	21.0	79.6	+77.2
50–54	62.2	2.2	19.8	15.9	78.0	+69.3
55–59	63.5	0.8	24.8	10.9	74.4	+61.5
60–64	64.2	0.2	28.9	6.8	70.9	+55.1
65–69	64.3	0.0	31.7	3.9	68.3	+50.7
70–74	64.4	0.0	33.7	1.9	66.3	+47.6
75–79	64.4	0.0	34.6	1.1	65.4	+46.3
80–84	64.4	0.0	35.3	0.3	64.7	+45.1
85–89	64.4	0.0	35.6	0.1	64.4	+44.7
90–94	64.4	0.0	35.6	0.0	64.4	+44.6
95–100	64.4	0.0	35.6	0.0	64.4	+44.6

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

Figure 13 (250% of national line): Households below the poverty line and all households at a given score or at or below a given score cut-off, scorecard applied to the 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	0.3	100.0	0.5	Only poor targeted
5–9	1.4	99.4	2.1	165.8:1
10–14	3.9	98.0	5.9	49.0:1
15–19	8.9	97.5	13.5	39.1:1
20–24	17.9	96.3	26.7	25.8:1
25–29	27.6	94.9	40.7	18.5:1
30–34	39.2	92.0	56.1	11.5:1
35–39	52.3	87.8	71.3	7.2:1
40–44	63.5	84.0	82.9	5.2:1
45–49	73.4	80.0	91.2	4.0:1
50–54	81.9	75.9	96.6	3.1:1
55–59	88.3	71.9	98.7	2.6:1
60–64	93.1	69.0	99.7	2.2:1
65–69	96.1	67.0	100.0	2.0:1
70–74	98.1	65.6	100.0	1.9:1
75–79	98.9	65.1	100.0	1.9:1
80–84	99.7	64.6	100.0	1.8:1
85–89	99.9	64.4	100.0	1.8:1
90–94	100.0	64.4	100.0	1.8:1
95–100	100.0	64.4	100.0	1.8:1

USAID "Extreme" Poverty Line

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	85.7
5–9	44.7
10–14	42.7
15–19	22.2
20–24	9.5
25–29	6.6
30–34	2.1
35–39	0.7
40–44	0.5
45–49	0.0
50–54	0.0
55–59	0.0
60–64	0.0
65–69	0.0
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

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

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+16.1	11.8	14.1	17.7
5–9	+5.2	7.8	9.7	12.0
10–14	+23.1	3.8	4.5	5.6
15–19	–0.8	3.2	4.0	5.1
20–24	–5.9	4.0	4.2	4.8
25–29	+2.4	0.8	1.0	1.2
30–34	–0.1	0.6	0.7	0.9
35–39	–0.7	0.6	0.6	0.8
40–44	+0.4	0.1	0.1	0.1
45–49	–0.0	0.0	0.0	0.0
50–54	+0.0	0.0	0.0	0.0
55–59	+0.0	0.0	0.0	0.0
60–64	+0.0	0.0	0.0	0.0
65–69	+0.0	0.0	0.0	0.0
70–74	+0.0	0.0	0.0	0.0
75–79	+0.0	0.0	0.0	0.0
80–84	+0.0	0.0	0.0	0.0
85–89	+0.0	0.0	0.0	0.0
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 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, scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+1.1	11.1	50.0	69.1
4	+0.5	19.2	28.9	40.6
8	+0.4	15.1	20.4	28.4
16	+0.4	10.8	13.4	18.9
32	+0.4	7.4	10.0	14.1
64	+0.3	5.6	6.8	8.9
128	+0.2	4.1	4.9	6.7
256	+0.2	2.7	3.2	4.3
512	+0.2	1.9	2.3	3.1
1,024	+0.2	1.3	1.6	2.3
2,048	+0.2	1.0	1.2	1.5
4,096	+0.2	0.7	0.8	1.0
8,192	+0.2	0.5	0.6	0.8
16,384	+0.2	0.4	0.4	0.5

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

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.2	4.3	0.1	95.4	95.6	–87.7
5–9	0.7	3.8	0.6	94.9	95.6	–52.6
10–14	1.4	3.1	2.5	93.0	94.4	+18.0
15–19	2.4	2.1	6.5	89.0	91.5	–43.9
20–24	3.5	1.0	14.4	81.1	84.6	–220.0
25–29	4.0	0.5	23.6	71.9	75.9	–425.1
30–34	4.3	0.2	35.0	60.5	64.8	–677.5
35–39	4.5	0.0	47.8	47.7	52.1	–963.6
40–44	4.5	0.0	59.0	36.5	41.0	–1,212.8
45–49	4.5	0.0	68.9	26.6	31.1	–1,431.9
50–54	4.5	0.0	77.5	18.1	22.5	–1,622.4
55–59	4.5	0.0	83.8	11.7	16.2	–1,763.1
60–64	4.5	0.0	88.6	6.9	11.4	–1,869.6
65–69	4.5	0.0	91.6	3.9	8.4	–1,936.4
70–74	4.5	0.0	93.6	1.9	6.4	–1,981.4
75–79	4.5	0.0	94.4	1.1	5.6	–2,000.1
80–84	4.5	0.0	95.2	0.3	4.8	–2,016.4
85–89	4.5	0.0	95.4	0.1	4.6	–2,022.4
90–94	4.5	0.0	95.5	0.0	4.5	–2,023.6
95–100	4.5	0.0	95.5	0.0	4.5	–2,023.9

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, scorecard applied to the 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	0.3	66.6	4.9	2.0:1
5–9	1.4	53.6	16.5	1.2:1
10–14	3.9	36.0	31.2	0.6:1
15–19	8.9	27.4	54.3	0.4:1
20–24	17.9	19.5	77.5	0.2:1
25–29	27.6	14.5	89.0	0.2:1
30–34	39.2	10.9	94.9	0.1:1
35–39	52.3	8.5	99.3	0.1:1
40–44	63.5	7.1	99.8	0.1:1
45–49	73.4	6.1	100.0	0.1:1
50–54	81.9	5.5	100.0	0.1:1
55–59	88.3	5.1	100.0	0.1:1
60–64	93.1	4.8	100.0	0.1:1
65–69	96.1	4.7	100.0	0.0:1
70–74	98.1	4.6	100.0	0.0:1
75–79	98.9	4.5	100.0	0.0:1
80–84	99.7	4.5	100.0	0.0:1
85–89	99.9	4.5	100.0	0.0:1
90–94	100.0	4.5	100.0	0.0:1
95–100	100.0	4.5	100.0	0.0:1

$1.25/day 2005 PPP Poverty Line

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

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	21.8
5–9	2.7
10–14	1.2
15–19	2.1
20–24	0.0
25–29	0.0
30–34	0.0
35–39	0.0
40–44	0.0
45–49	0.0
50–54	0.0
55–59	0.0
60–64	0.0
65–69	0.0
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Figure 8 ($1.25/day line): Bootstrapped differences between estimated and true household poverty likelihoods with confidence intervals in a large sample (n = 16,384), scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+14.6	6.7	8.1	9.3
5–9	+1.0	1.3	1.6	2.2
10–14	+0.2	0.6	0.7	1.0
15–19	+1.5	0.3	0.4	0.5
20–24	–1.4	1.1	1.2	1.5
25–29	–0.0	0.0	0.0	0.0
30–34	+0.0	0.0	0.0	0.0
35–39	+0.0	0.0	0.0	0.0
40–44	+0.0	0.0	0.0	0.0
45–49	+0.0	0.0	0.0	0.0
50–54	+0.0	0.0	0.0	0.0
55–59	+0.0	0.0	0.0	0.0
60–64	+0.0	0.0	0.0	0.0
65–69	+0.0	0.0	0.0	0.0
70–74	+0.0	0.0	0.0	0.0
75–79	+0.0	0.0	0.0	0.0
80–84	+0.0	0.0	0.0	0.0
85–89	+0.0	0.0	0.0	0.0
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 10 ($1.25/day line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	–0.1	1.0	1.0	1.3
4	+0.0	0.5	0.7	7.6
8	+0.0	0.4	0.5	8.9
16	–0.0	0.4	2.2	7.9
32	+0.0	1.0	1.8	4.4
64	–0.0	1.0	2.1	2.8
128	–0.0	1.0	1.2	1.4
256	+0.0	0.6	0.7	1.1
512	+0.0	0.4	0.5	0.8
1,024	+0.0	0.3	0.4	0.5
2,048	+0.0	0.2	0.3	0.4
4,096	+0.0	0.2	0.2	0.3
8,192	+0.0	0.1	0.1	0.2
16,384	+0.0	0.1	0.1	0.1

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

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.0	0.2	0.3	99.5	99.5	–33.4
5–9	0.1	0.2	1.3	98.4	98.5	–476.4
10–14	0.1	0.1	3.8	96.0	96.1	–1,546.5
15–19	0.2	0.1	8.8	91.0	91.2	–3,696.4
20–24	0.2	0.0	17.6	82.1	82.3	–7,548.4
25–29	0.2	0.0	27.4	72.4	72.6	–11,767.0
30–34	0.2	0.0	39.0	60.8	61.0	–16,801.5
35–39	0.2	0.0	52.1	47.7	47.9	–22,461.9
40–44	0.2	0.0	63.3	36.5	36.7	–27,328.7
45–49	0.2	0.0	73.2	26.6	26.8	–31,601.8
50–54	0.2	0.0	81.7	18.1	18.3	–35,313.9
55–59	0.2	0.0	88.0	11.7	12.0	–38,056.3
60–64	0.2	0.0	92.8	6.9	7.2	–40,131.5
65–69	0.2	0.0	95.8	3.9	4.2	–41,433.2
70–74	0.2	0.0	97.9	1.9	2.1	–42,310.9
75–79	0.2	0.0	98.7	1.1	1.3	–42,674.4
80–84	0.2	0.0	99.4	0.3	0.6	–42,991.6
85–89	0.2	0.0	99.7	0.1	0.3	–43,108.4
90–94	0.2	0.0	99.8	0.0	0.2	–43,132.3
95–100	0.2	0.0	99.8	0.0	0.2	–43,137.6

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

Figure 13 ($1.25/day line): Households below the poverty line and all households at a given score or at or below a given score cut-off, scorecard applied to the 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	0.3	7.0	10.0	0.1:1
5–9	1.4	4.1	24.9	0.0:1
10–14	3.9	2.7	44.9	0.0:1
15–19	8.9	1.7	66.1	0.0:1
20–24	17.9	1.2	96.6	0.0:1
25–29	27.6	0.8	100.0	0.0:1
30–34	39.2	0.6	100.0	0.0:1
35–39	52.3	0.4	100.0	0.0:1
40–44	63.5	0.4	100.0	0.0:1
45–49	73.4	0.3	100.0	0.0:1
50–54	81.9	0.3	100.0	0.0:1
55–59	88.3	0.3	100.0	0.0:1
60–64	93.1	0.2	100.0	0.0:1
65–69	96.1	0.2	100.0	0.0:1
70–74	98.1	0.2	100.0	0.0:1
75–79	98.9	0.2	100.0	0.0:1
80–84	99.7	0.2	100.0	0.0:1
85–89	99.9	0.2	100.0	0.0:1
90–94	100.0	0.2	100.0	0.0:1
95–100	100.0	0.2	100.0	0.0:1

$2.50/day 2005 PPP Poverty Line

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

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	85.7
5–9	51.6
10–14	46.6
15–19	26.6
20–24	13.9
25–29	8.8
30–34	3.7
35–39	2.0
40–44	0.4
45–49	0.2
50–54	0.2
55–59	0.0
60–64	0.0
65–69	0.0
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

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

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Score	**Diff.**	**90-percent**	**95-percent**	**99-percent**
0–4	+16.0	11.7	14.0	17.7
5–9	+7.0	8.6	10.1	12.8
10–14	+24.4	4.0	4.8	6.0
15–19	–4.6	4.1	4.4	5.3
20–24	–4.3	3.4	3.6	4.0
25–29	+3.9	0.9	1.0	1.5
30–34	+1.6	0.5	0.6	0.8
35–39	+0.5	0.5	0.5	0.7
40–44	+0.3	0.1	0.1	0.1
45–49	–0.1	0.2	0.2	0.3
50–54	+0.2	0.0	0.0	0.0
55–59	+0.0	0.0	0.0	0.0
60–64	+0.0	0.0	0.0	0.0
65–69	+0.0	0.0	0.0	0.0
70–74	+0.0	0.0	0.0	0.0
75–79	+0.0	0.0	0.0	0.0
80–84	+0.0	0.0	0.0	0.0
85–89	+0.0	0.0	0.0	0.0
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

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

Sample Size	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+1.5	13.3	60.0	71.4
4	+1.0	21.1	28.9	41.5
8	+0.8	16.1	21.4	30.1
16	+0.8	11.3	14.1	20.0
32	+0.9	8.0	10.3	13.8
64	+0.8	6.0	7.1	9.2
128	+0.6	4.1	5.0	6.6
256	+0.7	3.0	3.5	4.4
512	+0.7	2.1	2.4	3.4
1,024	+0.7	1.5	1.8	2.2
2,048	+0.7	1.1	1.3	1.7
4,096	+0.7	0.8	0.9	1.1
8,192	+0.7	0.5	0.6	0.9
16,384	+0.7	0.4	0.5	0.6

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

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.2	5.2	0.1	94.5	94.7	–89.7
5–9	0.8	4.6	0.6	94.0	94.9	–59.0
10–14	1.6	3.8	2.3	92.3	93.8	+1.4
15–19	2.8	2.5	6.1	88.5	91.4	–12.4
20–24	4.1	1.3	13.7	80.9	85.0	–154.6
25–29	4.8	0.6	22.8	71.8	76.5	–323.5
30–34	5.1	0.3	34.1	60.5	65.5	–533.1
35–39	5.3	0.1	47.0	47.6	53.0	–770.8
40–44	5.4	0.0	58.2	36.4	41.8	–978.4
45–49	5.4	0.0	68.0	26.6	32.0	–1,160.5
50–54	5.4	0.0	76.6	18.1	23.4	–1,319.3
55–59	5.4	0.0	82.9	11.7	17.1	–1,436.6
60–64	5.4	0.0	87.7	6.9	12.3	–1,525.4
65–69	5.4	0.0	90.7	3.9	9.3	–1,581.1
70–74	5.4	0.0	92.7	1.9	7.3	–1,618.6
75–79	5.4	0.0	93.5	1.1	6.5	–1,634.2
80–84	5.4	0.0	94.3	0.3	5.7	–1,647.7
85–89	5.4	0.0	94.5	0.1	5.5	–1,652.7
90–94	5.4	0.0	94.6	0.0	5.4	–1,653.8
95–100	5.4	0.0	94.6	0.0	5.4	–1,654.0

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

Figure 13 ($2.50/day line): Households below the poverty line and all households at a given score or at or below a given score cut-off, scorecard applied to the 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	0.3	67.2	4.1	2.1:1
5–9	1.4	59.4	15.3	1.5:1
10–14	3.9	40.1	29.0	0.7:1
15–19	8.9	32.0	52.8	0.5:1
20–24	17.9	23.2	76.7	0.3:1
25–29	27.6	17.3	88.4	0.2:1
30–34	39.2	13.0	94.2	0.1:1
35–39	52.3	10.2	98.6	0.1:1
40–44	63.5	8.4	99.2	0.1:1
45–49	73.4	7.4	100.0	0.1:1
50–54	81.9	6.6	100.0	0.1:1
55–59	88.3	6.1	100.0	0.1:1
60–64	93.1	5.8	100.0	0.1:1
65–69	96.1	5.6	100.0	0.1:1
70–74	98.1	5.5	100.0	0.1:1
75–79	98.9	5.5	100.0	0.1:1
80–84	99.7	5.4	100.0	0.1:1
85–89	99.9	5.4	100.0	0.1:1
90–94	100.0	5.4	100.0	0.1:1
95–100	100.0	5.4	100.0	0.1:1

$3.75/day 2005 PPP Poverty Line

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

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	90.4
10–14	78.7
15–19	68.2
20–24	53.7
25–29	37.1
30–34	27.3
35–39	14.0
40–44	8.6
45–49	3.3
50–54	1.4
55–59	0.1
60–64	0.1
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 ($3.75/day line): Bootstrapped differences between estimated and true household poverty likelihoods with confidence intervals in a large sample (n = 16,384), scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+3.3	2.9	3.7	4.6
5–9	–5.8	3.8	3.9	4.2
10–14	–5.1	4.0	4.3	4.9
15–19	–6.1	4.6	4.8	5.3
20–24	+1.7	2.7	3.3	4.4
25–29	+6.1	2.4	2.8	3.7
30–34	+6.1	1.8	2.1	2.7
35–39	–2.4	1.9	2.1	2.4
40–44	+1.8	1.0	1.3	1.6
45–49	–2.1	1.6	1.7	1.9
50–54	–1.5	1.2	1.2	1.4
55–59	–0.7	0.6	0.7	0.8
60–64	+0.1	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 ($3.75/day line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+0.7	63.2	77.1	88.2
4	+0.2	33.4	39.9	50.3
8	+0.4	22.5	27.1	37.1
16	+0.5	16.5	19.8	25.9
32	+0.6	11.9	14.6	19.5
64	+0.5	8.2	9.7	13.2
128	+0.4	6.0	7.3	9.4
256	+0.4	4.2	5.2	7.0
512	+0.4	2.9	3.5	4.7
1,024	+0.5	2.2	2.6	3.2
2,048	+0.4	1.5	1.8	2.3
4,096	+0.5	1.1	1.3	1.7
8,192	+0.4	0.7	0.9	1.2
16,384	+0.4	0.5	0.6	0.8

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

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.3	20.7	0.0	78.9	79.2	–96.9
5–9	1.3	19.8	0.1	78.8	80.1	–87.3
10–14	3.2	17.8	0.7	78.3	81.5	–66.1
15–19	6.7	14.4	2.3	76.7	83.3	–26.1
20–24	11.2	9.9	6.7	72.2	83.4	+37.8
25–29	14.3	6.7	13.3	65.7	80.0	+36.9
30–34	17.0	4.0	22.2	56.7	73.7	–5.5
35–39	19.3	1.8	33.0	45.9	65.2	–56.8
40–44	20.2	0.8	43.3	35.7	55.9	–105.6
45–49	20.8	0.3	52.6	26.3	47.1	–149.8
50–54	21.0	0.0	60.9	18.0	39.0	–189.4
55–59	21.1	0.0	67.2	11.7	32.8	–219.2
60–64	21.1	0.0	72.0	6.9	28.0	–242.0
65–69	21.1	0.0	75.0	3.9	25.0	–256.2
70–74	21.1	0.0	77.0	1.9	23.0	–265.9
75–79	21.1	0.0	77.9	1.1	22.1	–269.9
80–84	21.1	0.0	78.6	0.3	21.4	–273.3
85–89	21.1	0.0	78.9	0.1	21.1	–274.6
90–94	21.1	0.0	78.9	0.0	21.1	–274.9
95–100	21.1	0.0	78.9	0.0	21.1	–274.9

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

Figure 13 ($3.75/day line): Households below the poverty line and all households at a given score or at or below a given score cut-off, scorecard applied to the 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	0.3	94.5	1.5	17.2:1
5–9	1.4	93.0	6.1	13.3:1
10–14	3.9	82.8	15.4	4.8:1
15–19	8.9	74.7	31.6	2.9:1
20–24	17.9	62.4	53.0	1.7:1
25–29	27.6	51.9	68.0	1.1:1
30–34	39.2	43.4	80.8	0.8:1
35–39	52.3	36.9	91.6	0.6:1
40–44	63.5	31.8	96.1	0.5:1
45–49	73.4	28.3	98.7	0.4:1
50–54	81.9	25.6	99.8	0.3:1
55–59	88.3	23.9	100.0	0.3:1
60–64	93.1	22.6	100.0	0.3:1
65–69	96.1	21.9	100.0	0.3:1
70–74	98.1	21.5	100.0	0.3:1
75–79	98.9	21.3	100.0	0.3:1
80–84	99.7	21.1	100.0	0.3:1
85–89	99.9	21.1	100.0	0.3:1
90–94	100.0	21.1	100.0	0.3:1
95–100	100.0	21.1	100.0	0.3:1

$5.00/day 2005 PPP Poverty Line

Figure 5 ($5.00/day line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	96.6
10–14	93.4
15–19	90.7
20–24	77.6
25–29	63.2
30–34	60.0
35–39	42.8
40–44	30.3
45–49	15.9
50–54	9.9
55–59	5.8
60–64	1.3
65–69	0.4
70–74	0.4
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Figure 8 ($5.00/day line): Bootstrapped differences between estimated and true household poverty likelihoods with confidence intervals in a large sample (n = 16,384), scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+0.0	0.0	0.0	0.0
5–9	–2.8	1.7	1.7	1.7
10–14	+1.0	2.2	2.5	3.3
15–19	+1.6	2.3	2.8	3.5
20–24	–2.8	2.5	2.7	3.6
25–29	+1.0	2.4	2.9	3.8
30–34	+0.7	2.2	2.7	3.7
35–39	+1.2	2.2	2.5	3.3
40–44	–0.4	2.1	2.6	3.5
45–49	+0.1	1.7	2.0	2.6
50–54	–1.6	2.0	2.4	3.2
55–59	+1.2	1.5	1.7	2.3
60–64	–0.1	0.8	1.0	1.2
65–69	+0.4	0.0	0.0	0.0
70–74	+0.4	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 ($5.00/day line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+0.0	66.4	73.7	91.4
4	+0.3	38.9	45.6	57.6
8	+0.4	28.8	33.9	41.5
16	+0.5	20.0	24.4	30.4
32	+0.5	15.1	17.6	23.4
64	+0.4	10.3	12.0	16.0
128	+0.1	7.2	8.5	12.0
256	+0.1	5.3	6.2	8.9
512	+0.1	3.9	4.6	5.7
1,024	+0.1	2.7	3.2	4.1
2,048	+0.1	1.9	2.2	2.9
4,096	+0.0	1.3	1.6	2.0
8,192	+0.1	0.9	1.1	1.5
16,384	+0.1	0.7	0.8	1.0

Figure 12 ($5.00/day line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.3	39.6	0.0	60.1	60.4	–98.3
5–9	1.4	38.5	0.0	60.1	61.4	–93.1
10–14	3.6	36.3	0.3	59.8	63.4	–81.1
15–19	8.0	31.9	0.9	59.2	67.2	–57.5
20–24	15.1	24.8	2.8	57.3	72.4	–17.4
25–29	21.2	18.7	6.4	53.7	74.9	+22.4
30–34	28.1	11.8	11.1	49.0	77.1	+68.7
35–39	33.7	6.3	18.6	41.5	75.1	+53.3
40–44	37.1	2.8	26.4	33.6	70.7	+33.7
45–49	38.8	1.1	34.6	25.5	64.2	+13.3
50–54	39.6	0.3	42.4	17.7	57.3	–6.2
55–59	39.8	0.1	48.4	11.6	51.5	–21.4
60–64	39.9	0.0	53.2	6.9	46.8	–33.2
65–69	39.9	0.0	56.2	3.9	43.8	–40.7
70–74	39.9	0.0	58.2	1.9	41.8	–45.8
75–79	39.9	0.0	59.0	1.1	41.0	–47.9
80–84	39.9	0.0	59.8	0.3	40.2	–49.7
85–89	39.9	0.0	60.0	0.1	40.0	–50.4
90–94	39.9	0.0	60.1	0.0	39.9	–50.5
95–100	39.9	0.0	60.1	0.0	39.9	–50.6

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

Figure 13 ($5.00/day line): Households below the poverty line and all households at a given score or at or below a given score cut-off, scorecard applied to the 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	0.3	100.0	0.8	Only poor targeted
5–9	1.4	98.7	3.4	78.6:1
10–14	3.9	92.9	9.1	13.1:1
15–19	8.9	90.1	20.1	9.1:1
20–24	17.9	84.5	37.8	5.4:1
25–29	27.6	76.9	53.2	3.3:1
30–34	39.2	71.7	70.4	2.5:1
35–39	52.3	64.4	84.3	1.8:1
40–44	63.5	58.4	92.9	1.4:1
45–49	73.4	52.8	97.1	1.1:1
50–54	81.9	48.3	99.1	0.9:1
55–59	88.3	45.1	99.8	0.8:1
60–64	93.1	42.9	100.0	0.8:1
65–69	96.1	41.5	100.0	0.7:1
70–74	98.1	40.7	100.0	0.7:1
75–79	98.9	40.3	100.0	0.7:1
80–84	99.7	40.0	100.0	0.7:1
85–89	99.9	39.9	100.0	0.7:1
90–94	100.0	39.9	100.0	0.7:1
95–100	100.0	39.9	100.0	0.7:1