Simple Poverty Scorecard® Poverty-Assessment Tool Romania

Mark Schreiner, Dean Caire, and Gary Woller

28 September 2009

This document is at `SimplePovertyScorecard.com`

Abstract

The Simple Poverty Scorecard-brand poverty-assessment tool uses 10 low-cost indicators from Romania's 2007 Household Budget Survey to estimate the likelihood that a household has income or 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 Romania to measure poverty rates, to track changes in poverty rates over time, and to segment clients for differentiated treatment.

Acknowledgements

This paper was funded by the Consultative Group to Assist the Poorest under the CGAP/Ford Social Indicators Project. Data come from Romania's National Institute of Statistics and was purchased by Microfinance Risk Management, L.L.C. Thanks go to Malika Anand, Ewa Bankowska, Nigel Biggar, Frank DeGiovanni, Ioan Goreac, Micol Guarneri, Syed Hashemi, Katarzyna Pawlak, Roxana Preda-Ramniceanu, Tony Sheldon, Jeff Toohig, and especially Lucian Pop. "Simple Poverty Scorecard" is a Registered Trademark of Microfinance Risk Management, L.L.C. for its brand of poverty-assessment tools.

Simple Poverty Scorecard® Poverty-Assessment Tool

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

Indicator	Response	Points	Score
1. How many people live in the household?	A. Five or more	0	
	B. Four	5	
	C. Three	11	
	D. Two	18	
	E. One	28	
2. What is the highest educational level completed by any household member?	A. No formal school, primary (Class 1–4), or secondary (Class 5–8)	0	
	B. Vocational, complementary, apprenticeship, or lower high school (Class 9-10)	2	
	C. High school (Class 9-12)	4	
	D. Specialty post-high school, technical foreman, or university short-term (college)	5	
	E. University long-term	10	
3. What type of sewage system does the residence have?	A. Own system, not available, or no data	0	
	B. Public system	1	
4. How many working land-line and/or mobile telephones does the household own?	A. No land-line, and no mobile	0	
	B. No land-line, but one mobile	7	
	C. One or more land-lines, but no mobile	9	
	D. No land-line, but two or more mobiles	13	
	E. One or more land-lines, and one mobile	16	
	F. One or more land-lines, and two or more mobiles	18	
5. Does the household own any working gas stoves?	A. No	0	
	B. Yes	3	
6. What type of working washing machine does the household own?	A. None	0	
	B. Only non-automatic	3	
	C. Automatic (regardless of owning non-automatic)	5	
7. What type of working refrigerator/freezer does the household own?	A. None	0	
	B. Refrigerator only or freezer only	5	
	C. Refrigerator and freezer, or refrigerating combine	10	
8. Does the household own any color TV sets?	A. No	0	
	B. Yes	5	
9. Does the household own any radios?	A. No	0	
	B. Yes	2	
10. Does the household own a working car?	A. No	0	
	B. Yes	18	

Score:

Simple Poverty Scorecard® Poverty-Assessment Tool Romania

1. Introduction

Pro-poor programs in Romania 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, Romania's Household Budget Survey (HBS) asks households about the purchase and use of more than 500 food, non-food and service items in the past month.

In contrast, the indirect approach via the scorecard is simple, quick, and inexpensive. It uses ten verifiable indicators (such as "What type of sewage system does the residence have?" or "Does the household own any color TV sets?") to get a score that is highly correlated with poverty status as measured by 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 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). Measurements from these approaches are not comparable across organizations nor across countries, they may be costly, and their accuracy and precision are unknown.

Suppose an organization wants to know what share of its participants are below a poverty line (such as the Millennium Development Goals' $1.25/day at 2005 purchase-power parity). Or—as required of USAID microenterprise partners—an organization might want to report how many of its participants are among the poorest half of people below the national poverty line. Or an organization might want to measure movement across a poverty line (see, for example, Daley-Harris, 2009). In all these cases, what is needed is a data-based, objective tool with known accuracy. While household budget surveys are costly even for governments, local organizations may be able to implement an inexpensive poverty-assessment tool that can serve for monitoring, management, and targeting.

The statistical approach here aims to be understood by non-specialists. After all, if managers are to adopt the scorecard on their own and apply it to inform their decisions, 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, not because they do not work, but because they are presented (when they are presented at all) as tables of regression coefficients incomprehensible to non-specialists (with cryptic indicator names such as "`LGHHSZ_2`", negative values, and many decimal

places). Thanks to the predictive-modeling phenomenon known as the "flat maximum", simple, transparent scorecards are usually 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 these techniques are simple, they have rarely or never been applied to poverty-assessment tools.

The scorecard is based on the 2007 HBS conducted by the România Institutul Național de Statistică. Indicators are selected to be:

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

All points in the scorecard are non-negative integers, and total scores range from 0 (most likely below a poverty line) to 100 (least likely below a poverty line). Non-specialists can collect data and tally scores on paper in the field in 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-adult-equivalent income or consumption (or per-capita consumption) below a given poverty line.

Second, the scorecard can be used to estimate the poverty rate of a group of households at a point in time. This is estimated as the average poverty likelihood among the households in the group.

Third, the scorecard can be used to estimate changes in the poverty rate for a 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. To help managers choose the most appropriate targeting cut-off for their purposes, this paper reports several measures of targeting accuracy for a range of possible cut-offs.

This paper presents a single scorecard whose indicators and points are derived from data on household consumption and a poverty line equivalent to 150% of Romania's national poverty line. Scores from this one scorecard are calibrated to poverty likelihoods for eight poverty lines.

The scorecard is constructed and calibrated using some of the data from the 2007 HBS, and the scorecard's accuracy is validated on the rest of the data.

While all three scoring estimators are unbiased when applied to the population from which they were derived (that is, they match the true value on average in repeated samples from the same population from which the scorecard was built), they are—like all predictive models—biased to some extent when applied to a different population.[1]

Thus, while the indirect scoring approach is less costly than the direct survey approach, it is also biased. (The survey approach is unbiased by assumption.) There is

[1] Examples include nationally representative samples at a different point in time or non-nationally representative sub-groups (Tarozzi and Deaton, 2007).

bias because scoring must assume that the future relationship between indicators and poverty will be the same as in the data used to build the scorecard.[2] Of course, this assumption—ubiquitous and inevitable in predictive modeling—holds only partly.

When applied to the validation sample with bootstrap samples of $n = 16{,}384$, the difference between scorecard estimates of groups' poverty rates and the true rates at a point in time is +0.2 percentage points for the national line, and across all eight lines, the absolute average difference is 0.9 percentage points. These differences are due to sampling variation and not bias; the average of each difference would be zero if the whole 2007 HBS were to be repeatedly redrawn and divided into sub-samples before repeating the entire process of building and calibrating scorecards.

The 90-percent confidence intervals for these estimates are ±0.6 percentage points or less for estimates of a poverty rate at a point in time. For $n = 1{,}024$, the 90-percent intervals are ±2.1 percentage points or less.

Section 2 below describes data and poverty lines. Sections 3 and 4 describe scorecard construction and offer guidelines for use in practice. Sections 5 and 6 detail the estimation of households' poverty likelihoods and of groups' poverty rates at a point in time. Section 7 discusses estimating changes in poverty rates through time, Section 8 covers targeting, and Section 9 gives a summary.

[2] Bias may also result from changes in the quality of data collection, from changes in poverty lines, from imperfect adjustment of poverty lines to account for differences in cost-of-living across time or regions, or from sampling variation across surveys.

2. Data and poverty lines

This section discusses the data used to construct and test the scorecard. It also presents the poverty lines to which scores are calibrated.

2.1 Data

The scorecard is based on data from 31,597 households in the 2007 HBS conducted from 1 January to 31 December 2007. This is Romania's most recent available national consumption survey.

For the purposes of the scorecard, the households in the 2007 HBS are randomly divided into three sub-samples (Figure 2):

- *Construction* for selecting indicators and points
- *Calibration* for associating scores with poverty likelihoods
- *Validation* for measuring accuracy with 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 income or consumption (divided by the number of household members or by the number of adult equivalents) is below a given poverty line.

Beyond this general definition, there two special cases, *household-level poverty rates* and *person-level poverty rates*. With household-level rates, each household is counted as if it had only one person, regardless of true household size, so all households are counted equally. With person-level rates (the "head-count index"), each household is weighted by the number of people in it or by the number of adult equivalents in it, so larger households count more.

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 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 of $2 \div (1 + 2) = 67$ percent.

Whether the household-level rate or the person-level rate is 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 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 prefer to report household-level poverty rates.

Figure 2 reports poverty rates and poverty lines for Romania at both the household-level and the person-level. The scorecard is constructed using the 2007 HBS and household-level lines, scores are calibrated to household-level poverty likelihoods, and accuracy is measured for household-level rates. This use of household-level rates reflects the belief that they are relevant for most local, pro-poor organizations.

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

2.2.2 Poverty lines

Figure 2 shows the eight poverty lines used in this paper and their corresponding household- and person-level poverty rates.[3] The measure of consumption is described in World Bank (2003a).

The scorecard is constructed using a poverty line defined as 150% of Romania's national line. This national line is defined as the total consumption observed for households whose food consumption equals the cost of a 2,550-calorie food basket observed for people in the second and third quintiles of consumption (World Bank, 2007 and 2003b). In December 2007 prices, the national line is RON7.65 per day per adult equivalent, where the adult equivalents in a household is (Adults + 0.5 x Children)$^{0.9}$. In

[3] The measures of income and of consumption were graciously provided by Lucian Pop.

the 2007 HBS, this implies a household-level poverty rate of 7.7 percent and a person-level rate of 9.8 percent (Figure 2).[4]

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

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

The USAID "extreme" line is defined as the median consumption of people (not households) below the national line (U.S. Congress, 2002).

The USD1.25/day 2005 PPP line is derived from:

- 2005 PPP exchange rate for "individual consumption expenditure by households" (International Comparison Project, 2008): RON1.72 per USD1.00
- The 2005 average national monthly Consumer Price Index (RON 105.46), and the December 2007 CPI (RON 123.98).

[4] Income and consumption in December 2007 prices uses inflation data from `https://statistici.insse.ro/ipc/?page=ipc1&lang=en`, accessed 23 September 2009.

Given this, the USD2.50/day 2005 PPP line for Romania as a whole during the 2007 HBS is (Sillers, 2006):

$$(2005\ \text{PPP exchange rate}) \cdot \text{USD}2.50 \cdot \left(\frac{\text{CPI}_{\text{Dec 2007}}}{\text{CPI}_{\text{2005 average}}} \right) =$$

$$\left(\frac{\text{RON } 1.72}{\text{USD}1.00} \right) \cdot \text{USD}2.50 \cdot \left(\frac{123.98}{105.463} \right) = \text{RON } 5.06.$$

The USD3.75/day 2005 PPP and USD 5.00/day 2005 PPP lines are multiples of the USD2.50/day line.[5]

This paper also calibrates scores to poverty likelihoods for the Laeken poverty line, defined as 60 percent of the median per-adult-equivalent household income (Eurostat, 2003). The Laeken line uses its own definition of adult equivalents (1 for the first adult, 0.5 to other adults, and 0.3 to children), and, unlike the other lines here, it is relative and income-based, not absolute and consumption-based. Based on the 2007 HBS, the Laeken line is RON10.75 per day per adult equivalent (Figure 2).

[5] The USD1.25/day 2005 PPP line is not presented in this paper because almost no Romanian households are below it.

3. Scorecard construction

For Romania, about 60 potential indicators are initially prepared in the areas of:

- Family composition (such as household size)
- Education (such as highest level completed by any household member)
- Housing (such as the type of sewage system)
- Ownership of durable goods (such as gas stoves and colors television sets)

Each indicator is first reviewed with the entropy-based "uncertainty coefficient" (Goodman and Kruskal, 1979) that measures how well the indicator predicts poverty on its own. Figure 3 lists the candidate indicators, ranked by uncertainty coefficient. Responses for each indicator in Figure 3 are ordered starting with those most strongly linked with higher poverty likelihoods.

One aim of the scorecard is 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 gas stove 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. The first step is to use Logit to build one scorecard for each candidate indicator. Each scorecard's accuracy is taken as "c", a measure of ability to rank by poverty status (SAS Institute Inc., 2004).

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

A series of two-indicator scorecards are then built, each based on the one-indicator scorecard selected from the first step, 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.

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

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

The single scorecard here applies to all of Romania. Evidence from India and Mexico (Schreiner, 2006 and 2005a), Sri Lanka (Narayan and Yoshida, 2005), and Jamaica (Grosh and Baker, 1995) suggests that segmenting poverty-assessment tools by urban/rural does not improve targeting accuracy much, although it may improve estimates of poverty rates, as highlighted by Tarozzi and Deaton (2007).

4. Practical guidelines for scorecard use

The main challenge of scorecard design is not to maximize statistical accuracy but rather to improve the chances that scoring is actually used in practice (Schreiner, 2005b). When scoring projects fail, the reason is not usually statistical inaccuracy but rather the failure of an organization to decide to do what is needed to integrate scoring in its processes and to 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" (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 are more likely to adopt it and use it properly. Of course, accuracy matters, but it is balanced against 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 one 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 weights (non-negative integers, no arithmetic beyond addition)

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

- Record participant identifiers and household size
- Read each question from the scorecard
- Circle the 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 filing or data entry

Of course, field workers must be trained. Quality outputs depend on quality inputs. If organizations or field workers gather their own data and 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 budgeting, training field workers and supervisors, logistics, sampling, interviewing, piloting, recording data, and controlling quality.

[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 in a spreadsheet or database at the central office.

In particular, while collecting scorecard indicators is relatively easier than alternatives, it is still absolutely difficult. Training and explicit definitions of terms and concepts in the scorecard is essential. For the example of Nigeria, Onwujekwe, Hanson, and Fox-Rushby (2006) found distressingly low inter-rater and test-retest correlations for indicators as seemingly simple and obvious as whether the household owns an automobile. At the same time, Grosh and Baker (1995) find that gross underreporting of assets does not affect targeting. For the first stage of targeting in Mexico's *Oportunidades* conditional cash-transfer program, Martinelli and Parker (2007) find that "underreporting [of asset ownership] is widespread but not overwhelming, except for a few goods . . . [and] overreporting is common for a few goods, which implies that self-reporting may lead to the exclusion of deserving households" (pp. 24–25). Still, as is the practice of *Oportunidades* itself in the second stage of its targeting process, most false self-reports can be corrected by field agents who verify responses with a home visit, and this is the suggested procedure for the scorecard here.

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

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

In general, design should follow from the purposes that the organization intends for the scorecard.

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 downloaded to a database

The subjects to be scored can be:

- All participants (or all new participants)
- A representative sample of all participants (or of all new participants)
- All participants (or all new participants) in a representative sample of branches
- A representative sample of all participants (or of all new participants) in a representative sample of branches

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 level of confidence 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 change in poverty rates, it can be applied:

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

An example set of choices were made by BRAC and ASA, two Bangladeshi microlenders with more than 7 million participants each who are applying the scorecard

(Schreiner, 2013). Their design is that loan officers in a random sample of branches apply the scorecard with all participants 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. ASA's and BRAC's sampling plans cover 50,000–100,000 participants each.

5. Estimates of household poverty likelihoods

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

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, scores of 20–24 have a poverty likelihood of 32.1 percent, and scores of 25–29 have a poverty likelihood of 25.1 percent (Figure 4).

The poverty likelihood associated with a score varies by poverty line. For example, scores of 20–24 are associated with a poverty likelihood of 32.1 percent for the national line but 77.4 percent for 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 4, many figures have eight versions, one for each of eight poverty lines. To keep them straight, they are grouped by poverty line. Single tables that pertain to all poverty lines are placed with the first group of tables for the national line.

For the example of the national line (Figure 5), there are 3,525 (normalized) households in the calibration sub-sample with a score of 20–24, of whom 1,132 (normalized) are below the poverty line. The estimated poverty likelihood associated with a score of 20–24 is then 32.1 percent, because 1,132 ÷ 3,525 = 32.1 percent.

To illustrate with the national line and a score of 25–29, there are 5,882 (normalized) households in the calibration sample, of whom 1,476 (normalized) are below the line (Figure 5). Thus, the poverty likelihood for this score is 1,476 ÷ 5,882 = 25.1 percent.

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

Figure 6 shows, for all scores, the likelihood that consumption (per adult equivalent or per person) or income (per adult equivalent) falls in a range demarcated by two adjacent poverty lines.[8] For example, the daily consumption of someone with a score of 20–24 falls in the following ranges with probability:

- 16.0 percent below the USD2.50/day 2005 PPP line
- 43.6 percent between the USD2.50/day and USD3.75/day 2005 PPP lines
- 29.8 percent between the USD3.75/day and USD5.00/day 2005 PPP lines
- 10.6 percent above the USD5.00/day 2005 PPP line

Even though the scorecard is constructed partly based on judgment, the calibration process produces poverty likelihoods that are objective, that is, derived from

[8] Figure 6 has two parts, one for the per-adult-equivalent national lines and one for the per-person international 2005 PPP lines. The USAID "extreme" line and the Laeken line are not in Figure 6 because they are non-comparable to the other lines.

survey data on consumption (or income) and quantitative poverty lines. The poverty likelihoods would be objective even if indicators and/or points were selected without any data at all. In fact, objective scorecards of proven accuracy are often constructed using only expert 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 the Romania scorecard are transformed coefficients from a Logit regression, scores are not converted to poverty likelihoods via the Logit formula of 2.718281828^{score} x $(1+ 2.718281828^{score})^{-1}$. 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, going from scores to poverty likelihoods in this way requires no arithmetic at all, just a look-up table. This non-parametric calibration can also improve accuracy, especially with large 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, then this calibration process produces unbiased estimates of poverty likelihoods. *Unbiased* means that in repeated samples from the same population, the average estimate matches the true 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 over time.[9]

Of course, the relationship between indicators and poverty does change to some unknown extent with time—especially if there have been large changes in the macroeconomy, as in fact has occurred in Romania after 2007—and also across subgroups in Romania's population, so the scorecard will generally be biased when applied after December 2007 (the end date of fieldwork for the 2007 HBS) or when applied with non-nationally representative groups.

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

How accurate are estimates of households' poverty likelihoods? To measure, 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 and below a poverty line
- For each score, record the difference between the estimated poverty likelihood (Figure 4) and the true poverty likelihood in the bootstrap sample
- Repeat the previous three steps 1,000 times
- For each score, report the average difference between estimated and true poverty likelihoods across the 1,000 bootstrap samples
- For each score, report the two-sided interval containing the central 900, 950, or 990 differences between estimated and true poverty likelihoods

For each score range and for $n = 16{,}384$, Figure 7 shows the average difference between estimated and true poverty likelihoods, as well as confidence intervals for the differences.

For the national line, the average poverty likelihood across bootstrap samples for scores of 20–24 in the validation sample is too high by 1.0 percentage points (Figure 7). For scores of 25–29, the estimate is too low by 1.0 percentage points.[10]

The 90-percent confidence interval for the differences for scores of 20–24 is ±3.5 percentage points (Figure 7). This means that in 900 of 1,000 bootstraps, the difference between the estimate and the true value is between –2.5 and +4.5 percentage points

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

(because 1.0 – 3.5 = –2.5 and 1.0 + 3.5 = +4.5). In 950 of 1,000 bootstraps (95 percent), the difference is 1.0 ±4.2 percentage points, and in 990 of 1,000 bootstraps (99 percent), the difference is 1.0 ±5.5 percentage points.

For some scores of 39 or less, Figure 7 shows differences—sometimes large ones—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 Romania's population. For targeting, however, what matters is less the difference in all score ranges and more the difference in score ranges just above and below the targeting cut-off. This mitigates the effects of 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 cancel out. This is the case, as discussed in the next section.

Another possible source of differences between estimates and true values is overfitting. By construction, the scorecard here is unbiased, but it may still be *overfit* when applied after the end of the HBS fieldwork in December 2007. That is, it may fit the 2007 HBS data so closely that it captures not only some timeless patterns but also some random patterns that, due to sampling variation, show up only in the 2007 HBS. Or the scorecard may be overfit in the sense that it is not robust to changes in the

relationships between indicators and poverty over time or when it is applied to non-nationally representative samples.

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

Most errors in individual households' likelihoods cancel out in the estimates of groups' poverty rates (see next section). Furthermore, at least some of the differences come from non-scorecard sources that can be addressed only by improving data quantity and quality (which is beyond the scope of the scorecard) or by reducing overfitting (which likely has limited returns, given the scorecard's parsimony).

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

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

To illustrate, suppose a program samples three households on Jan. 1, 2009 and that they have scores of 20, 30, and 40, corresponding to poverty likelihoods of 32.1, 14.6, and 4.2 percent (national line, Figure 4). The group's estimated poverty rate is the households' average poverty likelihood of $(32.1 + 14.6 + 4.2) \div 3 = 17.0$ percent.[11]

6.1 Accuracy of estimated poverty rates at a point in time

For the Romania scorecard applied to the validation sample with $n = 16{,}384$, the absolute differences between the estimated poverty rate at a point in time and the true rate are 3.0 percentage points or less (Figure 9, summarizing Figure 8 across poverty lines). The average absolute difference across the eight poverty lines is 0.9 percentage points. At least part of these differences is due to sampling variation in the validation sample and in the random division of the 2007 HBS into three sub-samples.

In terms of precision, the 90-percent confidence interval for a group's estimated poverty rate at a point in time with $n = 16{,}384$ is ± 0.6 percentage points or less (Figure 9). This means that in 900 of 1,000 bootstraps of this size, the difference between the

[11] 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 a score of 30 is 14.6 percent. This is not the 17.0 percent found as the average of the three poverty likelihoods associated with each of the three scores.

estimate and the true value is within 0.6 percentage points of the average difference. In the specific case of the national line and the validation sample, 90 percent of all samples of $n = 16{,}384$ produce estimates that differ from the true value in the range of 0.2 + 0.4 = 0.6 to 0.2 – 0.4 = –0.2 percentage points. This is because 0.2 is the average difference, and ±0.4 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; it estimates a poverty rate of 7.8 percent for the validation sample, but the true value is 7.6 percent (Figure 2).

6.2 Formula for standard errors for estimates of poverty rates

How precise are the point-in-time estimates? Because they are averages of binary (0/1, or poor/non-poor) variables, the estimates (in "large" samples) have a Normal distribution and can be characterized by their average difference vis-à-vis true values together with the standard error of the average difference.

To derive a formula for the standard errors of estimated poverty rates at a point in time from indirect measurement via poverty-assessment tools (Schreiner, 2008a), note that the textbook formula (Cochran, 1977) that relates confidence intervals with standard errors in the case of direct measurement of poverty status 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, this implies that for a sample n of 16,384 with 90-percent confidence ($z = 1.64$) and a poverty rate p of 7.7 percent (the average poverty rate in the construction and calibration samples in Figure 2 for the national line), the confidence interval c is $+/- z \cdot \sqrt{\frac{p \cdot (1-p)}{n}} = +/- 1.64 \cdot \sqrt{\frac{0.077 \cdot (1-0.077)}{16{,}384}} = \pm 0.00342$, or ± 0.342 percentage points.

The scorecard, however, does not measure poverty directly, so this formula is not immediately applicable. To derive a formula for the Romania scorecard, consider Figure 8, 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 $n = 16{,}384$ and the national line, the 90-percent confidence interval is 0.360 percentage points.[12]

Thus, the 90-percent confidence interval with $n = 16{,}384$ is 0.342 percentage points for the Romania scorecard and 0.360 percentage points for direct measurement. The ratio of the two intervals is $0.360 \div 0.342 = 1.05$.

Now consider the same case, but with $n = 8{,}192$. The confidence interval under direct measurement is $+/- 1.64 \cdot \sqrt{\frac{0.077 \cdot (1-0.077)}{8{,}192}} = \pm 0.00483$, or about ± 0.483 percentage points. The empirical confidence interval with the Romania scorecard (Figure 8) is 0.490. Thus for $n = 8{,}192$, the ratio of the two intervals is $0.490 \div 0.483 = 1.01$.

[12] Due to rounding, Figure 8 displays 0.4, not 0.360.

This ratio of 1.01 for $n = 8{,}182$ is not far from the ratio of 1.05 for $n = 16{,}384$. Across all sample sizes of 256 or more in Figure 8, the average ratio turns out to be 1.04, implying that confidence intervals for indirect estimates of poverty rates via the Romania scorecard and this poverty line are about 4 percent wider than confidence intervals for direct estimates via the 2007 HBS. This 1.04 appears in Figure 9 as the "α factor" because if α is 1.04, then the formula relating confidence intervals c and standard errors σ for the Romania scorecard is $c = +/- z \cdot \alpha \cdot \sigma$. That is, formula for the standard error σ for point-in-time estimates of poverty rates via scoring is

$$\alpha \cdot \sqrt{\frac{p \cdot (1-p)}{n}}.$$

In general, α can be more or less than 1.00. When α is less than 1.00, it means that the scorecard is more precise than direct measurement. This occurs for five of the eight poverty lines 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 before measurement.[13] If $\hat{p}$ is the expected poverty rate before measurement, then the formula for sample size n

[13] IRIS Center (2007a and 2007b) says that a sample size of $n = 300$ is sufficient for USAID reporting. 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.

based on the desired confidence level that corresponds to z and the desired confidence interval $\pm c$ is $n = \left(\frac{\alpha \cdot z}{c}\right)^2 \cdot \hat{p} \cdot (1 - \hat{p})$.

To illustrate how to use this, suppose $c = 0.02805$ and $z = 1.64$ (90-percent confidence). Then the formula gives $n = \left(\frac{1.04 \cdot 1.64}{0.02805}\right)^2 \cdot 0.077 \cdot (1 - 0.077) = 263$, a bit higher than the sample size of 256 observed for these parameters in Figure 8 for the national line.

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

In practice after the end of fieldwork for the HBS in December 2007, 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 7.7 percent national average in the 2007 HBS in Figure 2), look up α (here, 1.04), assume that the scorecard will still work in the future and/or for non-nationally representative sub-groups,[14] and then compute the required sample size. In this illustration, $n = \left(\frac{1.04 \cdot 1.64}{0.02}\right)^2 \cdot 0.077 \cdot (1 - 0.077) = 517$.

[14] This paper reports accuracy for the scorecard applied to the validation sample, but it cannot test accuracy for later years or for other groups. Performance after December 2007 will resemble that in the 2007 HBS with deterioration to the extent that the relationships between indicators and poverty status change over time. Due to the recent post-2007 economic crisis, deterioration in performance will be greater than in the absence of such a large macroeconomic shock.

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 2007 HBS, this paper cannot test estimates of change over time for Romania, 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, poverty could get better or worse, and scoring does not indicate what caused change. This point is often forgotten or confused, so it bears repeating: the scorecard simply estimates change, and it does not, in and of itself, indicate the reason for the change. In particular, estimating the impact of program participation requires knowing what would have happened to participants if they had not been participants. Knowing this requires either strong assumptions or a control group that resembles participants in all ways except participation. To belabor the point, the scorecard can help estimate 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, 2009, a program samples three households who score 20, 30, and 40 and so have poverty likelihoods of 32.1, 14.6, and 4.2 percent (national line, Figure 4). The group's baseline estimated poverty rate is the households' average poverty likelihood of (32.1 + 14.6 + 4.2) ÷ 3 = 17.0 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, 2010, 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 25, 35, and 45 (poverty likelihoods of 25.1, 8.8, and 2.5 percent, national line, Figure 4). Their average poverty likelihood at follow-up is now (25.1 + 8.8 + 2.5) ÷ 3 = 12.1 percent, an improvement of 17.0 – 12.1 = 4.9 percentage points.[15]

This suggests that about one in twenty participants in this hypothetical example crossed the poverty line in 2009.[16] Among those who started below the line, about one in four (4.9 ÷ 17.0 = 28.8 percent) on net ended up above the line.[17]

[15] Of course, such a huge reduction in poverty in one year is unlikely, but this is just an example to show how the scorecard can be used to estimate change.
[16] This is a net figure; some people start above the line and end below it, and vice versa.

7.3 Accuracy for estimated change in two independent samples

With only the 2007 HBS, 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 Romania 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.

For two equal-sized independent samples, the same logic as above 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,[18] and α is the average (across a range of bootstrapped sample sizes) of the ratio of the observed confidence interval from a scorecard and the theoretical confidence interval under direct measurement.

---

[17] The scorecard does not reveal the reasons for this change.

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

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 the countries for which this α has been measured (Peru, the Philippines, India, Mexico, Pakistan, Vietnam, and Bangladesh, see Schreiner, 2009a, 2009b, 2009c, 2009d, and 2008b and Chen and Schreiner, 2009), the average α across poverty lines and years is 0.77, 0.77, 1.40, 1.48, 1.16, 0.68, and 1.03. The average across countries (1.04) is as reasonable a figure as any for Romania.

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, $\alpha = 1.04$, and $\hat{p} = 0.077$ (from Figure 2). Then the baseline sample size is $n = 2 \cdot \left(\frac{1.04 \cdot 1.64}{0.02} \right)^2 \cdot 0.077 \cdot (1 - 0.077) =$ 1,034, and the follow-up sample size is also 1,034.

7.4 Accuracy for estimated change for one sample, scored twice

Analogous to previous derivations, the general formula relating the confidence interval c to the standard error σ when using a scorecard to estimate change for a single group of households, all of whom are scored at two points in time, is:[19]

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

where z, c, and α are defined as usual, 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.

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

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

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

Because $\hat{p}_*$ could be anything between 0–1, more information is needed to apply 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 Romania scorecard is applied twice (once after December 2007 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 an 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 first be scored in 2009 and then again in 2012 ($y = 3$). The before-baseline poverty rate is 7.7 percent ($p_{2007} = 0.077$, Figure 2), and suppose $\alpha = 1.30$. Then the baseline sample size is $n = 2 \cdot \left(\frac{1.30 \cdot 1.64}{0.02} \right)^2 \cdot \{ -0.02 + 0.016 \cdot 3 + 0.47 \cdot [0.077 \cdot (1 - 0.077)] \} = 1{,}396$. The same group of 1,396 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* (consumption or income below a poverty line). Poverty status is a fact that depends on whether consumption or income 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 10 depicts these four possible targeting outcomes. Targeting accuracy varies by cut-off; a higher cut-off has better inclusion (but greater leakage), while a lower cut-off has better exclusion (but higher undercoverage).

It makes sense for a program to 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 11 shows the distribution of households by targeting outcome. For an example cut-off of 20–24, outcomes for the national line in the validation sample are:

- Inclusion: 3.5 percent are below the line and correctly targeted
- Undercoverage: 4.3 percent are below the line and mistakenly not targeted
- Leakage: 4.5 percent are above the line and mistakenly targeted
- Exclusion: 87.8 percent are above the line and correctly not targeted

Increasing the cut-off to 25–29 improves inclusion and undercoverage but worsens leakage and exclusion:

- Inclusion: 5.0 percent are below the line and correctly targeted
- Undercoverage: 2.8 percent are below the line and mistakenly not targeted
- Leakage: 8.8 percent are above the line and mistakenly targeted
- Exclusion: 83.4 percent are above the line and correctly not targeted

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

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

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

- Assign benefits and costs to possible outcomes, based on its values and mission
- Tally total net benefits for each cut-off using Figure 11 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. It makes sense for a program that uses targeting—with or without scoring—to consider thoughtfully how it values successful inclusion or exclusion versus errors of

undercoverage and leakage. It is healthy to go through a process of thinking explicitly and intentionally about how possible targeting outcomes are valued.

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

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

Figure 11 shows "Total Accuracy" for all cut-offs for the Romania scorecard. For the national line in the validation sample, total net benefit is greatest (93.2) for a cut-off of 10–14, with about 15 in 16 households in Romania 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).[20]

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

[20] Figure 11 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 both in terms of the estimated poverty rate and in terms of targeting inclusion. After normalizing by the number of people below the poverty line, the formula is:
BPAC = (Inclusion – |Undercoverage – Leakage|) x [100 ÷ (Inclusion+Undercoverage)].

achieve a desired poverty rate among targeted households. The third column of Figure 12 ("% targeted who are poor") shows, for the Romania scorecard applied to the validation sample, the expected poverty rate among households who score at or below a given cut-off. For the example of the national line in the validation sample, targeting households who score 29 or less would target 13.8 percent of all households (second column) and produce a poverty rate among those targeted of 35.9 percent (third column).

Figure 12 also reports two other measures of targeting accuracy. The first is a version of coverage ("% of poor who are targeted"). For the example of the national line in the validation sample and a cut-off of 29 or less, 64.0 percent of all poor households are covered.

The final targeting measure in Figure 12 is the number of successfully targeted poor households for each non-poor household mistakenly targeted (right-most column). For the national line in the validation sample and a cut-off of 29 or less, covering 0.6 poor households means leaking to 1 non-poor household.

9. Conclusion

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

The scorecard is built with part of the data from the 2007 HBS, tested on a different part of the 2007 HBS, and calibrated to eight poverty lines (national, 150% of national, 200% of national, USAID "extreme", USD2.50/day 2005 PPP, USD3.75/day 2005 PPP, USD5.00/day 2005 PPP, and the Laeken line).

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

When the scorecard is applied to the validation sample with $n = 16{,}384$, the absolute difference between estimates versus true poverty rates for groups of households at a point in time is 3.0 percentage points or less. The average absolute difference across the eight poverty lines is about 0.9 percentage points. For $n = 16{,}384$ and 90-percent confidence, the precision of these differences is ± 0.6 percentage points or better.

If a program wants to use the scorecard for targeting, then the results here provide the information needed to select a cut-off that fits its values and mission.

Although the statistical technique is innovative, and although technical accuracy is important, the design of the scorecard here 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 that are 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 Romania to monitor poverty rates, track changes in poverty rates over time, and target services. 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.

Caire, Dean. (2004) "Building Credit Scorecards for Small Business Lending in Developing Markets", `microfinance.com/English/Papers/Scoring_SMEs_Hybrid.pdf`, accessed 28 September 2009.

Chen, Shiyuan; and Mark Schreiner. (2009) "Simple Poverty Scorecard Poverty-Assessment Tool: Vietnam" `simplepovertyscorecard.com/VNM_2006_ENG.pdf`, accessed 11 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/`, accessed 28 September 2009.

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*.

Eurostat. (2003) "'Laeken' Indicators Detailed Calculation Methodology", European Commission Document E2/IPSE/2003, `cso.ie/eusilc/documents/Laeken%20Indicators%20-%20calculation%20algorithm.pdf`, accessed 23 September 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, accessed 28 September 2009.

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, worldbank.org/VPMWVLU8E0, accessed 28 September 2009.

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, accessed 28 September 2009.

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.

International Comparison Project. (2008) "Tables of Results", siteresources.worldbank.org/ICPINT/Resources/icp-final-tables.pdf, accessed 28 September 2009.

IRIS Center. (2007a) "Manual for the Implementation of USAID Poverty Assessment Tools", povertytools.org/training_documents/Manuals/USAID_PAT_Manual_Eng.pdf, accessed 28 September 2009.

_____. (2007b) "Introduction to Sampling for the Implementation of PATs", povertytools.org/training_documents/Sampling/Introduction_Sampling.ppt, accessed 28 September 2009.

_____. (2005) "Notes on Assessment and Improvement of Tool Accuracy", povertytools.org/other_documents/AssessingImproving_Accuracy.pdf, accessed 28 September 2009.

Johnson, Glenn. (2007) "Lesson 3: Two-Way Tables—Dependent Samples", www.stat.psu.edu/online/development/stat504/03_2way/53_2way_compare.htm, accessed 28 September 2009.

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

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

Martinelli, César; and Susan W. Parker. (2007) "Deception and Misreporting in a Social Program", `ciep.itam.mx/~martinel/lies4.pdf`, accessed 28 September 2009.

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`, accessed 28 September 2009.

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

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

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

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/59654/HTML/default/statug_logistic_sect035.htm`, accessed 28 September 2009.

Schreiner, Mark. (2013) "Simple Poverty Scorecard Poverty-Assessment Tool: Bangladesh", `simplepovertyscorecard.com/BGD_2010_ENG.pdf`, accessed 11 July 2016.

_____. (2009a) "Simple Poverty Scorecard Poverty-Assessment Tool: Peru", `simplepovertyscorecard.com/PER_2007_ENG.pdf`, accessed 11 July 2016.

_____. (2009b) "Simple Poverty Scorecard Poverty-Assessment Tool: Philippines", simplepovertyscorecard.com/PHL_2002_ENG.pdf, accessed 11 July 2016.

_____. (2009c) "Simple Poverty Scorecard Poverty-Assessment Tool: Mexico", simplepovertyscorecard.com/MEX_2008_ENG.pdf, 11 July 2016.

_____. (2009d) "Simple Poverty Scorecard Poverty-Assessment Tool: Pakistan", simplepovertyscorecard.com/PAK_2005_ENG.pdf, accessed 11 July 2016.

_____. (2008a) "Simple Poverty Scorecard Poverty-Assessment Tool: Peru", simplepovertyscorecard.com/PER_2003_ENG.pdf, accessed 11 July 2016.

_____. (2008b) "Simple Poverty Scorecard Poverty-Assessment Tool: India", simplepovertyscorecard.com/IND_2005_ENG.pdf, accessed 11 July 2016.

_____. (2006) "Is One Simple Poverty Scorecard Poverty-Assessment Tool Enough for India?", microfinance.com/English/Papers/Scoring_Poverty_India_Segments.pdf, accessed 28 September 2009.

_____. (2005a) "La Herramienta del Índice de Calificación de la Pobreza™: México", simplepovertyscorecard.com/MEX_2002_SPA.pdf, accessed 11 July 2016.

_____. (2005b) "IRIS Questions on the Simple Poverty Scorecard Poverty-Assessment Tool", microfinance.com/English/Papers/Scoring_Poverty_Response_to_IRIS.pdf, accessed 28 September 2009.

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

_____; Matul, Michal; Pawlak, Ewa; and Sean Kline. (2004) "Poverty Scoring: Lessons from a Microlender in Bosnia-Herzegovina", microfinance.com/English/Papers/Scoring_Poverty_in_BiH_Short.pdf, accessed 28 September 2009.

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, accessed 28 September 2009.

Toohig, Jeff. (2008) "PPI Pilot Training Guide", progressoutofpoverty.org/toolkit, accessed 28 September 2009.

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.

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

World Bank. (2007) "Romania Poverty Monitoring Analytical and Advisory Assistance Program, First Phase Report, Fiscal Year 2007", Report No. 40120–RO, www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2007/12/04/000020439_20071204114818/Rendered/PDF/401200RO.pdf, accessed 23 September 2009.

_____. (2003a) "Romania Poverty Assessment: Volume Two, Background Papers", Report No. 26169–RO, www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2003/10/29/000160016_20031029100844/Rendered/PDF/261691RO1vol02.pdf, accessed 23 September 2009.

_____. (2003b) "Romania Poverty Assessment: Volume One, Main Report", Report No. 26169–RO, siteresources.worldbank.org/INTROMANIA/Resources/PovertyAssessment_Eng.pdf, accessed 23 September 2009.

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

Figure 2: Sample sizes and household poverty rates by poverty line

			Poverty rates (% below a poverty line)							
				National		USAID	Intl. 2005 PPP			
Sub-sample	Level	Sample size	100%	150%	200%	'Extreme'	$2.50/day	$3.75/day	$5.00/day	Laeken
Poverty Rates:										
All Romania	Households	31,597	7.7	26.9	51.1	4.0	3.8	13.8	29.3	23.3
	People		9.8	31.3	55.8	6.3	6.0	19.3	37.4	23.9
Construction										
Selecting indicators	Households	10,584	7.7	26.9	51.3	3.9	3.8	13.8	29.4	22.9
and points	People		9.9	31.8	56.5	6.4	6.2	19.6	37.8	23.7
Calibration										
Associating scores	Households	10,454	7.7	27.0	51.4	3.8	3.7	14.0	29.3	23.7
with likelihoods	People		9.7	31.2	56.0	6.0	5.8	19.4	37.4	24.3
Validation										
Measuring accuracy	Households	10,559	7.6	26.9	50.6	4.1	3.8	13.6	29.2	23.3
	People		9.9	30.9	55.0	6.6	6.2	19.0	37.0	23.7
	Poverty lines:									
	All Romania		7.65	11.48	15.30	5.14	5.06	7.59	10.12	10.75

Source: 2007 Romania Household Budget Survey. Poverty lines are in RON in prices as of December 2007.

The national line is in units of expenditure of RON per adult equivalent per day.

The Laeken line is in units of income of RON per adult equivalent per day.

The national and the Laeken line use different definitions of "adult equivalent", see text.

The USAID "extreme" line and the international 2005 PPP lines are in units of expenditure in RON per person per day.

Almost no Romanian households are below the $1.25/day 2005 PPP line, so it is not displayed.

Figure 3: Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly linked with higher poverty likelihoods)
1516	How many working land-line and/or mobile telephones does the household own? (No land-line, and no mobile; No land-line, but one mobile; One or more land-lines, but no mobile; No land-line, but two or more mobiles; One or more land-lines, and one mobile; One or more land-lines, and two or more mobiles)
1385	How is hot water supplied to the residence? (No system available, disconnected, or no data; Own system; From public system)
1340	Does the residence have a bathroom? (No; Yes)
1308	Does the residence have an inside toilet? (No; Yes)
1295	What type of water supply does the residence have? (Other; Inside, own system; Inside from public system)
1270	What type of working washing machine does the household own? (None; Only non-automatic; Automatic (regardless of owning non-automatic))
1264	What was the highest grade completed by the female head/spouse? (No formal school, or primary (Class 1-4); Secondary (Class 5-8); No female head/spouse; Lower high school (Class 9-10); High school (Class 9-12); Vocational, complementary, apprenticeship, specialty post-high school, technical foreman, university short-term (college), or university long-term)
1248	How is the residence heated? (Wood/coal/oil stove; Natural gas stove, no heating, disconnected, or no data; Central heating; Own central heating)
1158	Does the household own any working automatic washing machines? (No; Yes)
1147	What type of working refrigerator/freezer does the household own? (None; Refrigerator only or freezer only; Refrigerator and freezer, or refrigerating combine)
1117	Does the household own a working vacuum cleaner? (No; Yes)
1081	What is the highest educational level completed by any household member? (No formal school, primary (Class 1–4), or secondary (Class 5–8); Vocational, complementary, apprenticeship, or lower high school (Class 9-10); High school (Class 9-12); Specialty post-high school, technical foreman, or university short-term (college); University long-term)

Figure 3 (cont.): Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly linked with higher poverty likelihoods)
1046	What is the highest-quality fuel that the household uses for cooking? (Wood/coal/oil, other, or no data; Gas cylinder; Electricity, or natural gas)
1042	What was the highest grade completed by the male head/spouse? (No formal school, or primary (Class 1-4); Secondary (Class 5-8); Vocational, complementary, apprenticeship, or lower high school (Class 9-10); No male head/spouse; High school (Class 9-12); Specialty post-high school, technical foreman, university short-term (college), or university long-term)
977	Does the household own a working car? (No; Yes)
961	Is the residence connected to the public sewer system? (No; Yes)
877	Does the household live in an apartment? (No; Yes)
874	What material is the residence made of? (Other or no data; Wood; Stone or bricks; Concrete)
849	Does the household own a working land-line telephone? (No; Yes)
682	How many working mobile telephones does the household own? (None; One; Two or more)
576	Does the household own a working PC? (No; Yes)
561	Does the household own a refrigerating combine? (No; Yes)
467	Does the household own a working gas stove? (No; Yes)
418	Does the household have a working internet connection? (No; Yes)
396	How many people live in the household? (Five or more; Four; Three; Two; One)
319	Does the household own any color TV sets? (No; Yes)
261	Does the household own a working CD player? (No; Yes)
238	In what year was the residence constructed? (1939 or before; 1940 to 1949; 1950 to 1959; 1960 to 1969; 1970 to 1979; 1980 to 1989; 1990 or later)
194	Does the household own a working music system? (No; Yes)
191	Does the household own a working freezer? (No; Yes)
190	How many people in the household are of ages 0 to 18? (Two or more; One; None)

Figure 3 (cont.): Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly linked with higher poverty likelihoods)
180	What is the structure of the household in terms of men 18-years-old or older, women 18-years-old or older, and children 17-years-old or younger? (One or more men, one or more women, two or more children; No men or women and only children, two or more men with no women and no children, one or more men with no women and one or more children, no men with two or more women and no children, or no men with one or more women and one or more children; One or more men, one or more women, no children; One or more men, one or more women, one child; No men, one woman, no children; One man, no women, no children)
165	Does the household own a working video player? (No; Yes)
157	How many people in the household are of ages 0 to 17? (Two or more; One; None)
151	How many people in the household are of ages 0 to 16? (Two or more; One; None)
150	How many people in the household are of ages 0 to 15? (Two or more; One; None)
140	Does the household own a working video camera? (No; Yes)
121	How many people in the household are of ages 0 to 14? (Two or more; One; None)
120	How many people in the household are of ages 0 to 13? (Two or more; One; None)
115	How many people in the household are of ages 0 to 12? (Two or more; One; None)
111	How many people in the household are of ages 0 to 11? (Two or more; One; None)
107	Does the household own any working black-and-white or color TV sets? (No; Yes)
100	What is the marital status of female head/spouse? (Married; Partner; Divorced; Widow; Unmarried; Separated; No female head/spouse)
91	What is the living area occupied by the household (square meters)? (19 or less; 20 to 29; 30 to 39; 40 to 49; 50 to 59; 60 or more)
79	How many people in the household are of ages 0 to 7? (One or more; None)
61	What is the marital status of male head/spouse? (Partner or separated; Widower; No male head/spouse; Married; Divorced or unmarried)
58	Does the household own a working sewing machine? (No; Yes)

Figure 3 (cont.): Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly linked with higher poverty likelihoods)
40	Does the household own a working tape recorder or radiocassette recorder? (No; Yes)
39	What energy source does the household use for illumination? (Electricity; Other)
30	How old is the female head/spouse? (55 or older; 46 to 55; 36 to 45; 25 or younger; No female head/spouse; 26 to 35)
28	How old is the male head/spouse? (55 or older; 46 to 55; No male head/spouse; 36 to 45; 35 or younger)
17	Does the household own a working refrigerator? (No; Yes)
16	Does the household own a working non-automatic washing machine? (No; Yes)
12	What is the household's occupancy status? (Personal property; Other)
11	Does the household own a working dishwasher? (No; Yes)
11	Does the household own a working radio? (No; Yes)
8	Does the household own a working motorcycle or motorbike? (No; Yes)
8	Does the household own a working motorcycle or motorbike? (No; Yes)
6	Does the household own a working van? (No; Yes)
6	How many rooms are occupied by the household? (One or two; Three or more)
0	What is the structure of household headship? (Not female head/spouse only; Female head/spouse only)
0	Does the household own a working bicycle? (No; Yes)

Source: 2007 Household Budget Survey and 150 percent of the national poverty line.

National Poverty Line Tables

(and Tables Pertaining to All Seven Poverty Lines)

Figure 4 (National poverty 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	77.9
5–9	68.1
10–14	50.2
15–19	46.7
20–24	32.1
25–29	25.1
30–34	14.6
35–39	8.8
40–44	4.2
45–49	2.5
50–54	0.8
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 5 (National poverty line): Derivation of estimated poverty likelihoods associated with scores

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	242	÷	311	=	77.9
5–9	437	÷	642	=	68.1
10–14	599	÷	1,194	=	50.2
15–19	1,050	÷	2,249	=	46.7
20–24	1,132	÷	3,525	=	32.1
25–29	1,476	÷	5,882	=	25.1
30–34	1,244	÷	8,537	=	14.6
35–39	864	÷	9,852	=	8.8
40–44	484	÷	11,576	=	4.2
45–49	266	÷	10,654	=	2.5
50–54	88	÷	10,710	=	0.8
55–59	7	÷	8,842	=	0.1
60–64	7	÷	7,232	=	0.1
65–69	0	÷	5,804	=	0.0
70–74	0	÷	4,898	=	0.0
75–79	0	÷	4,331	=	0.0
80–84	0	÷	2,424	=	0.0
85–89	0	÷	975	=	0.0
90–94	0	÷	231	=	0.0
95–100	0	÷	130	=	0.0

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

Figure 6: Distribution of household poverty likelihoods across consumption ranges demarcated by national poverty lines in RON per day per adult equivalent

	Likelihood of expenditure in range demarcated by poverty lines			
	<National	=>National and <150% of national	=>150% of national and <200% of national	=>200% of national
Score	<RON7.65	=>RON7.65 and <RON11.48	=>RON11.48 and <RON15.30	<RON15.30
0–4	77.9	22.1	0.0	0.0
5–9	68.1	31.9	0.0	0.0
10–14	50.2	32.6	14.3	2.9
15–19	46.7	38.6	10.9	3.8
20–24	32.1	45.3	18.2	4.4
25–29	25.1	44.2	22.2	8.5
30–34	14.6	41.0	29.6	14.9
35–39	8.8	34.7	34.2	22.3
40–44	4.2	26.8	37.0	32.0
45–49	2.5	15.3	36.0	46.1
50–54	0.8	11.0	31.0	57.2
55–59	0.1	6.2	24.0	69.7
60–64	0.1	3.1	18.2	78.6
65–69	0.0	0.7	9.9	89.4
70–74	0.0	0.8	4.7	94.5
75–79	0.0	0.0	2.4	97.6
80–84	0.0	0.0	0.0	100.0
85–89	0.0	0.0	0.0	100.0
90–94	0.0	0.0	0.0	100.0
95–100	0.0	0.0	0.0	100.0

All poverty likelihoods in percentage units.

The national, USAID "extreme", and Laeken lines all have different units and so cannot be compared with each other or with the international 2005 PPP lines here.

Figure 6 (cont.): Distribution of household poverty likelihoods across consumption ranges demarcated by international 2005 PPP poverty lines in RON per day per person

	Likelihood of expenditure in range demarcated by poverty lines			
	<$2.50/day	=>$2.50/day and <$3.75/day	=>$3.75/day and <$5.00/day	=>$5.00/day
Score	<RON5.06	=>RON5.06 and <RON7.59	=>RON7.59 and <RON10.12	<RON10.12
0–4	77.9	22.1	0.0	0.0
5–9	55.6	36.8	7.5	0.0
10–14	45.3	27.6	20.1	7.0
15–19	34.9	39.3	18.1	7.7
20–24	16.0	43.6	29.8	10.6
25–29	10.1	35.0	31.5	23.4
30–34	4.3	24.5	28.1	43.1
35–39	1.6	15.3	28.7	54.5
40–44	0.7	8.7	23.8	66.9
45–49	0.4	5.1	14.9	79.6
50–54	0.2	1.2	11.7	86.9
55–59	0.0	0.7	6.0	93.3
60–64	0.1	0.0	4.5	95.4
65–69	0.0	0.0	1.1	98.9
70–74	0.0	0.6	0.2	99.2
75–79	0.0	0.0	0.0	100.0
80–84	0.0	0.0	0.0	100.0
85–89	0.0	0.0	0.0	100.0
90–94	0.0	0.0	0.0	100.0
95–100	0.0	0.0	0.0	100.0

All poverty likelihoods in percentage units.

The national, USAID "extreme", and Laeken lines all have different units and so cannot be compared with each other or with the international 2005 PPP lines here.

Figure 7 (National poverty line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) with confidence intervals, 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	–19.8	11.1	11.1	11.1
5–9	–7.3	7.4	8.8	11.7
10–14	–16.0	10.7	11.3	12.3
15–19	+8.3	4.8	5.7	7.2
20–24	+1.0	3.5	4.2	5.5
25–29	–1.0	2.5	3.1	3.8
30–34	–1.5	1.8	2.2	2.7
35–39	+3.3	0.9	1.2	1.5
40–44	–0.0	0.9	1.1	1.3
45–49	+1.2	0.5	0.6	0.8
50–54	+0.5	0.2	0.2	0.3
55–59	–0.1	0.1	0.2	0.2
60–64	+0.1	0.1	0.1	0.1
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 8 (National poverty 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	39.2	53.5	71.9
4	+0.8	22.4	27.4	38.0
8	+0.4	15.2	18.9	25.8
16	+0.3	11.1	13.6	18.5
32	+0.2	7.7	9.3	12.1
64	+0.1	5.4	6.8	9.0
128	+0.2	3.9	4.7	6.3
256	+0.2	2.8	3.4	4.8
512	+0.2	2.0	2.4	3.0
1,024	+0.2	1.4	1.6	2.2
2,048	+0.2	1.0	1.2	1.5
4,096	+0.2	0.7	0.8	1.1
8,192	+0.2	0.5	0.6	0.7
16,384	+0.2	0.4	0.4	0.6

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							
		100%	200%	USAID	International 2005 PPP			
	Natl.	Natl.	Natl.	'Extreme'	$2.50/day	$3.75/day	$5.00/day	Laeken
Estimate minus true value								
Scorecard applied to validation sample	+0.2	+0.5	–0.4	–0.8	–0.6	–0.0	–1.8	+3.0
Precision of difference								
Scorecard applied to validation sample	0.4	0.5	0.6	0.3	0.3	0.4	0.6	0.5
α for sample size								
Scorecard applied to validation sample	1.04	0.91	0.88	1.13	1.13	0.93	0.93	0.90
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.								
See text for the units of the poverty lines.								

Figure 10 (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 11 (National poverty 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	7.5	0.0	92.2	92.5	–92.2
5–9	0.8	7.0	0.2	92.1	92.8	–77.7
10–14	1.5	6.2	0.6	91.7	93.2	–52.3
15–19	2.4	5.4	2.0	90.2	92.6	–12.6
20–24	3.5	4.3	4.5	87.8	91.3	+42.4
25–29	5.0	2.8	8.8	83.4	88.4	–14.2
30–34	6.4	1.4	16.0	76.3	82.7	–106.0
35–39	7.0	0.7	25.1	67.1	74.1	–224.7
40–44	7.5	0.2	36.3	56.0	63.5	–368.1
45–49	7.7	0.1	46.8	45.5	53.1	–503.7
50–54	7.7	0.0	57.4	34.8	42.6	–641.2
55–59	7.7	0.0	66.2	26.0	33.8	–755.1
60–64	7.7	0.0	73.5	18.8	26.5	–848.4
65–69	7.7	0.0	79.3	13.0	20.7	–923.3
70–74	7.7	0.0	84.2	8.1	15.8	–986.6
75–79	7.7	0.0	88.5	3.8	11.5	–1,042.5
80–84	7.7	0.0	90.9	1.3	9.1	–1,073.8
85–89	7.7	0.0	91.9	0.4	8.1	–1,086.3
90–94	7.7	0.0	92.1	0.1	7.9	–1,089.3
95–100	7.7	0.0	92.3	0.0	7.7	–1,091.0

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

Figure 12 (National poverty line): For a given score cut-off, the percentage of all households who are targeted (that is, have a score equal to or less than the cut-off), the percentage of targeted households who are poor (that is, below the poverty line), the percentage of poor households who are targeted, and the number of poor households who are successful targeted (coverage) per non-poor household mistakenly targeted (leakage), 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	95.1	3.8	19.4:1
5–9	1.0	81.1	10.0	4.3:1
10–14	2.1	72.1	20.0	2.6:1
15–19	4.4	53.9	30.6	1.2:1
20–24	7.9	43.7	44.7	0.8:1
25–29	13.8	35.9	64.0	0.6:1
30–34	22.3	28.6	82.5	0.4:1
35–39	32.2	21.9	90.9	0.3:1
40–44	43.8	17.2	97.0	0.2:1
45–49	54.4	14.1	98.9	0.2:1
50–54	65.1	11.9	99.7	0.1:1
55–59	74.0	10.5	99.9	0.1:1
60–64	81.2	9.5	100.0	0.1:1
65–69	87.0	8.9	100.0	0.1:1
70–74	91.9	8.4	100.0	0.1:1
75–79	96.2	8.0	100.0	0.1:1
80–84	98.7	7.9	100.0	0.1:1
85–89	99.6	7.8	100.0	0.1:1
90–94	99.9	7.8	100.0	0.1:1
95–100	100.0	7.7	100.0	0.1:1

150% of the National Poverty Line Tables

Figure 4 (150% of the national line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	82.8
15–19	85.3
20–24	77.4
25–29	69.3
30–34	55.6
35–39	43.5
40–44	31.0
45–49	17.8
50–54	11.9
55–59	6.3
60–64	3.2
65–69	0.7
70–74	0.8
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Figure 7 (150% of the national line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) with confidence intervals, 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	+1.8	2.2	2.5	3.4
5–9	+3.5	2.9	3.6	4.5
10–14	–9.4	6.3	6.6	7.0
15–19	–2.2	3.1	3.6	4.7
20–24	+6.5	3.5	4.1	5.4
25–29	+2.1	2.9	3.5	4.4
30–34	–3.1	2.7	3.0	3.9
35–39	+2.7	2.2	2.7	3.4
40–44	+2.8	2.0	2.4	2.8
45–49	–1.3	1.7	2.1	2.8
50–54	+0.3	1.5	1.7	2.3
55–59	+1.7	1.0	1.1	1.5
60–64	–2.3	1.8	2.0	2.2
65–69	+0.5	0.2	0.3	0.3
70–74	+0.7	0.2	0.3	0.3
75–79	–0.1	0.1	0.1	0.1
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 8 (150% of the 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	62.3	75.7	84.7
4	+0.7	31.4	39.3	51.8
8	+0.5	23.3	27.6	36.2
16	+0.4	16.4	19.6	25.0
32	+0.4	11.5	14.1	18.8
64	+0.3	8.4	9.8	13.0
128	+0.4	6.0	7.1	9.6
256	+0.4	4.2	5.2	6.6
512	+0.5	2.9	3.5	4.5
1,024	+0.5	2.1	2.4	3.0
2,048	+0.5	1.5	1.7	2.4
4,096	+0.5	1.1	1.3	1.7
8,192	+0.5	0.7	0.9	1.3
16,384	+0.5	0.5	0.6	0.8

Figure 11 (150% of the 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	26.7	0.0	73.0	73.3	–97.7
5–9	0.9	26.1	0.0	73.0	73.9	–93.1
10–14	2.0	25.0	0.1	72.9	74.9	–84.5
15–19	4.0	23.1	0.4	72.6	76.5	–69.1
20–24	6.5	20.5	1.4	71.6	78.1	–46.6
25–29	10.5	16.5	3.3	69.7	80.3	–9.9
30–34	15.6	11.4	6.7	66.3	81.9	+40.6
35–39	19.7	7.3	12.5	60.5	80.2	+53.7
40–44	23.1	3.9	20.7	52.3	75.3	+23.3
45–49	25.1	1.9	29.3	43.7	68.8	–8.6
50–54	26.3	0.8	38.9	34.1	60.4	–44.0
55–59	26.7	0.3	47.3	25.7	52.4	–75.1
60–64	27.0	0.0	54.2	18.8	45.7	–100.8
65–69	27.0	0.0	60.0	13.0	40.0	–122.3
70–74	27.0	0.0	64.9	8.1	35.1	–140.4
75–79	27.0	0.0	69.2	3.8	30.8	–156.4
80–84	27.0	0.0	71.7	1.3	28.3	–165.3
85–89	27.0	0.0	72.6	0.4	27.4	–169.0
90–94	27.0	0.0	72.9	0.1	27.1	–169.8
95–100	27.0	0.0	73.0	0.0	27.0	–170.3

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

Figure 12 (150% of the national line): For a given score cut-off, the percentage of all households who are targeted (that is, have a score equal to or less than the cut-off), the percentage of targeted households who are poor (that is, below the poverty line), the percentage of poor households who are targeted, and the number of poor households who are successful targeted (coverage) per non-poor household mistakenly targeted (leakage), 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	96.8	1.1	29.9:1
5–9	1.0	96.6	3.4	28.1:1
10–14	2.1	94.6	7.5	17.4:1
15–19	4.4	89.9	14.6	8.9:1
20–24	7.9	82.0	24.0	4.5:1
25–29	13.8	76.3	39.0	3.2:1
30–34	22.3	69.9	57.8	2.3:1
35–39	32.2	61.1	72.9	1.6:1
40–44	43.8	52.7	85.4	1.1:1
45–49	54.4	46.1	92.9	0.9:1
50–54	65.1	40.3	97.2	0.7:1
55–59	74.0	36.1	98.8	0.6:1
60–64	81.2	33.2	99.9	0.5:1
65–69	87.0	31.0	99.9	0.4:1
70–74	91.9	29.4	100.0	0.4:1
75–79	96.2	28.1	100.0	0.4:1
80–84	98.7	27.4	100.0	0.4:1
85–89	99.6	27.1	100.0	0.4:1
90–94	99.9	27.0	100.0	0.4:1
95–100	100.0	27.0	100.0	0.4:1

200% of the National Poverty Line Tables

Figure 4 (200% of the national line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	97.1
15–19	96.2
20–24	95.6
25–29	91.5
30–34	85.1
35–39	77.7
40–44	68.0
45–49	53.9
50–54	42.8
55–59	30.3
60–64	21.4
65–69	10.6
70–74	5.6
75–79	2.4
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Figure 7 (200% of the national line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) with confidence intervals, scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+0.0	0.0	0.0	0.0
5–9	+0.0	0.0	0.0	0.0
10–14	–1.6	1.5	1.6	2.1
15–19	–1.7	1.4	1.5	1.9
20–24	+2.0	1.8	2.2	2.8
25–29	+1.4	1.8	2.1	3.0
30–34	–0.8	1.8	2.2	2.8
35–39	–1.6	1.9	2.2	2.9
40–44	–0.6	2.0	2.4	3.1
45–49	–3.1	2.7	3.0	3.7
50–54	–1.5	2.3	2.6	3.6
55–59	+2.8	2.3	2.8	3.5
60–64	+0.2	2.3	2.6	3.4
65–69	+0.7	1.9	2.3	2.9
70–74	–1.2	1.7	2.1	2.8
75–79	+0.4	1.0	1.2	1.6
80–84	–1.1	1.1	1.3	1.4
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 8 (200% of the 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	+1.9	62.6	73.7	90.4
4	+0.1	34.8	42.9	57.2
8	+0.2	24.1	29.8	41.6
16	+0.3	17.6	21.0	26.9
32	+0.1	12.6	15.1	19.9
64	–0.1	9.4	11.5	15.3
128	–0.1	6.3	7.8	9.8
256	–0.3	4.6	5.5	7.7
512	–0.4	3.2	3.7	4.9
1,024	–0.4	2.3	2.5	3.2
2,048	–0.4	1.6	1.9	2.5
4,096	–0.4	1.1	1.4	1.8
8,192	–0.4	0.8	0.9	1.3
16,384	–0.4	0.6	0.7	0.9

Figure 11 (200% of the 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	50.7	0.0	49.0	49.3	–98.8
5–9	1.0	50.0	0.0	49.0	50.0	–96.3
10–14	2.1	48.8	0.0	49.0	51.1	–91.6
15–19	4.3	46.6	0.1	49.0	53.3	–82.9
20–24	7.6	43.4	0.3	48.7	56.3	–69.5
25–29	13.0	38.0	0.8	48.2	61.2	–47.5
30–34	20.4	30.6	2.0	47.1	67.4	–16.2
35–39	28.1	22.9	4.1	44.9	72.9	+18.2
40–44	35.9	15.1	7.9	41.1	77.0	+56.2
45–49	41.8	9.2	12.6	36.4	78.2	+75.2
50–54	46.3	4.7	18.8	30.2	76.5	+63.1
55–59	48.7	2.3	25.3	23.7	72.4	+50.4
60–64	50.1	0.9	31.2	17.9	67.9	+38.9
65–69	50.6	0.4	36.4	12.6	63.2	+28.5
70–74	50.9	0.1	41.0	8.0	58.8	+19.5
75–79	51.0	0.0	45.3	3.7	54.7	+11.2
80–84	51.0	0.0	47.7	1.3	52.3	+6.4
85–89	51.0	0.0	48.7	0.4	51.3	+4.5
90–94	51.0	0.0	48.9	0.1	51.1	+4.1
95–100	51.0	0.0	49.0	0.0	51.0	+3.8

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

Figure 12 (200% of the national line): For a given score cut-off, the percentage of all households who are targeted (that is, have a score equal to or less than the cut-off), the percentage of targeted households who are poor (that is, below the poverty line), the percentage of poor households who are targeted, and the number of poor households who are successful targeted (coverage) per non-poor household mistakenly targeted (leakage), 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	0.6	Only poor targeted
5–9	1.0	100.0	1.9	Only poor targeted
10–14	2.1	99.2	4.2	129.6:1
15–19	4.4	98.4	8.5	60.9:1
20–24	7.9	96.0	14.9	24.1:1
25–29	13.8	94.0	25.5	15.8:1
30–34	22.3	91.2	40.0	10.4:1
35–39	32.2	87.2	55.0	6.8:1
40–44	43.8	81.9	70.3	4.5:1
45–49	54.4	76.8	82.0	3.3:1
50–54	65.1	71.1	90.9	2.5:1
55–59	74.0	65.8	95.5	1.9:1
60–64	81.2	61.6	98.2	1.6:1
65–69	87.0	58.1	99.2	1.4:1
70–74	91.9	55.3	99.8	1.2:1
75–79	96.2	52.9	100.0	1.1:1
80–84	98.7	51.7	100.0	1.1:1
85–89	99.6	51.2	100.0	1.0:1
90–94	99.9	51.0	100.0	1.0:1
95–100	100.0	51.0	100.0	1.0:1

USAID "Extreme" Poverty Line Tables

Figure 4 (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	77.9
5–9	57.3
10–14	45.3
15–19	35.2
20–24	18.1
25–29	10.3
30–34	4.7
35–39	1.9
40–44	0.7
45–49	0.4
50–54	0.2
55–59	0.0
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 7 (USAID "extreme" line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) with confidence intervals, 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	–20.4	11.1	11.1	11.1
5–9	–18.0	12.5	13.2	14.8
10–14	–15.6	10.8	11.4	12.4
15–19	+3.1	4.7	5.5	7.3
20–24	–0.1	3.1	3.7	4.7
25–29	–3.6	2.9	3.0	3.3
30–34	–2.2	1.8	1.9	2.1
35–39	–0.0	0.6	0.7	0.9
40–44	–0.5	0.5	0.6	0.8
45–49	+0.3	0.0	0.1	0.1
50–54	+0.1	0.1	0.1	0.2
55–59	+0.0	0.0	0.0	0.0
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 8 (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	–0.3	5.2	50.0	70.3
4	–0.2	17.7	24.0	34.8
8	–0.6	13.1	16.2	21.9
16	–0.6	9.2	11.2	15.5
32	–0.8	6.3	7.4	10.1
64	–0.8	4.3	5.3	7.4
128	–0.8	3.2	3.9	5.1
256	–0.8	2.2	2.7	3.6
512	–0.8	1.6	2.0	2.7
1,024	–0.8	1.2	1.4	1.8
2,048	–0.8	0.8	1.0	1.3
4,096	–0.8	0.6	0.7	0.9
8,192	–0.8	0.4	0.5	0.6
16,384	–0.8	0.3	0.3	0.4

Figure 11 (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	0.3	4.0	0.0	95.7	96.0	–85.7
5–9	0.8	3.5	0.2	95.6	96.3	–59.5
10–14	1.5	2.8	0.7	95.1	96.6	–14.8
15–19	2.1	2.1	2.3	93.5	95.6	+46.9
20–24	2.7	1.5	5.2	90.6	93.3	–21.1
25–29	3.5	0.8	10.3	85.4	88.8	–142.2
30–34	3.9	0.3	18.4	77.3	81.3	–331.1
35–39	4.1	0.1	28.1	67.7	71.8	–557.3
40–44	4.2	0.0	39.5	56.2	60.4	–825.8
45–49	4.2	0.0	50.2	45.6	49.8	–1,075.1
50–54	4.3	0.0	60.9	34.9	39.1	–1,325.5
55–59	4.3	0.0	69.7	26.0	30.3	–1,532.6
60–64	4.3	0.0	76.9	18.8	23.1	–1,702.0
65–69	4.3	0.0	82.7	13.0	17.3	–1,837.9
70–74	4.3	0.0	87.6	8.1	12.4	–1,952.7
75–79	4.3	0.0	92.0	3.8	8.0	–2,054.1
80–84	4.3	0.0	94.4	1.3	5.6	–2,110.9
85–89	4.3	0.0	95.4	0.4	4.6	–2,133.7
90–94	4.3	0.0	95.6	0.1	4.4	–2,139.1
95–100	4.3	0.0	95.7	0.0	4.3	–2,142.2

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

Figure 12 (USAID "extreme" line): For a given score cut-off, the percentage of all households who are targeted (that is, have a score equal to or less than the cut-off), the percentage of targeted households who are poor (that is, below the poverty line), the percentage of poor households who are targeted, and the number of poor households who are successful targeted (coverage) per non-poor household mistakenly targeted (leakage), 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	96.8	7.0	29.9:1
5–9	1.0	81.3	18.1	4.4:1
10–14	2.1	69.4	34.9	2.3:1
15–19	4.4	48.4	49.9	0.9:1
20–24	7.9	34.7	64.4	0.5:1
25–29	13.8	25.1	81.1	0.3:1
30–34	22.3	17.6	92.1	0.2:1
35–39	32.2	12.8	96.7	0.1:1
40–44	43.8	9.7	99.3	0.1:1
45–49	54.4	7.8	99.5	0.1:1
50–54	65.1	6.6	100.0	0.1:1
55–59	74.0	5.8	100.0	0.1:1
60–64	81.2	5.3	100.0	0.1:1
65–69	87.0	4.9	100.0	0.1:1
70–74	91.9	4.6	100.0	0.0:1
75–79	96.2	4.4	100.0	0.0:1
80–84	98.7	4.3	100.0	0.0:1
85–89	99.6	4.3	100.0	0.0:1
90–94	99.9	4.3	100.0	0.0:1
95–100	100.0	4.3	100.0	0.0:1

$2.50/day 2005 PPP Poverty Line Tables

Figure 4 ($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	77.9
5–9	55.6
10–14	45.3
15–19	34.9
20–24	16.0
25–29	10.1
30–34	4.3
35–39	1.6
40–44	0.7
45–49	0.4
50–54	0.2
55–59	0.0
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 7 ($2.50/day line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) with confidence intervals, 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	–20.4	11.1	11.1	11.1
5–9	–13.7	10.6	11.3	12.9
10–14	–15.4	10.7	11.3	12.3
15–19	+3.8	4.5	5.6	7.7
20–24	–2.0	3.1	3.7	4.7
25–29	–2.1	2.1	2.4	3.0
30–34	–0.8	1.2	1.3	1.7
35–39	–0.0	0.6	0.7	0.9
40–44	–0.5	0.5	0.6	0.8
45–49	+0.3	0.0	0.1	0.1
50–54	+0.1	0.1	0.1	0.2
55–59	+0.0	0.0	0.0	0.0
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 8 ($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	–0.3	5.1	50.0	70.5
4	–0.2	17.3	23.5	34.6
8	–0.5	12.6	15.9	21.7
16	–0.5	8.8	10.8	14.8
32	–0.6	6.0	7.0	9.5
64	–0.6	4.2	4.8	7.1
128	–0.6	3.1	3.7	5.1
256	–0.6	2.2	2.6	3.6
512	–0.6	1.5	1.8	2.6
1,024	–0.6	1.2	1.4	1.8
2,048	–0.6	0.8	1.0	1.1
4,096	–0.6	0.5	0.6	0.8
8,192	–0.6	0.4	0.5	0.6
16,384	–0.6	0.3	0.3	0.4

Figure 11 ($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	0.3	3.7	0.0	96.0	96.3	–84.6
5–9	0.7	3.2	0.2	95.8	96.6	–57.3
10–14	1.4	2.5	0.7	95.3	96.8	–9.4
15–19	2.1	1.9	2.3	93.7	95.8	+41.4
20–24	2.7	1.3	5.2	90.8	93.5	–32.2
25–29	3.3	0.7	10.5	85.5	88.8	–164.7
30–34	3.7	0.3	18.7	77.4	81.0	–370.4
35–39	3.8	0.1	28.4	67.7	71.5	–614.7
40–44	3.9	0.0	39.8	56.2	60.1	–903.8
45–49	3.9	0.0	50.5	45.6	49.5	–1,172.0
50–54	4.0	0.0	61.2	34.9	38.8	–1,441.4
55–59	4.0	0.0	70.0	26.0	30.0	–1,664.3
60–64	4.0	0.0	77.2	18.8	22.8	–1,846.5
65–69	4.0	0.0	83.0	13.0	17.0	–1,992.8
70–74	4.0	0.0	87.9	8.1	12.1	–2,116.2
75–79	4.0	0.0	92.3	3.8	7.7	–2,225.4
80–84	4.0	0.0	94.7	1.3	5.3	–2,286.4
85–89	4.0	0.0	95.7	0.4	4.3	–2,311.0
90–94	4.0	0.0	95.9	0.1	4.1	–2,316.8
95–100	4.0	0.0	96.0	0.0	4.0	–2,320.1

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

Figure 12 ($2.50/day line): For a given score cut-off, the percentage of all households who are targeted (that is, have a score equal to or less than the cut-off), the percentage of targeted households who are poor (that is, below the poverty line), the percentage of poor households who are targeted, and the number of poor households who are successful targeted (coverage) per non-poor household mistakenly targeted (leakage), 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	96.8	7.6	29.9:1
5–9	1.0	77.8	18.7	3.5:1
10–14	2.1	67.5	36.5	2.1:1
15–19	4.4	47.1	52.2	0.9:1
20–24	7.9	33.8	67.4	0.5:1
25–29	13.8	23.9	83.2	0.3:1
30–34	22.3	16.4	92.6	0.2:1
35–39	32.2	11.9	96.6	0.1:1
40–44	43.8	9.0	99.2	0.1:1
45–49	54.4	7.3	99.5	0.1:1
50–54	65.1	6.1	100.0	0.1:1
55–59	74.0	5.4	100.0	0.1:1
60–64	81.2	4.9	100.0	0.1:1
65–69	87.0	4.6	100.0	0.0:1
70–74	91.9	4.3	100.0	0.0:1
75–79	96.2	4.1	100.0	0.0:1
80–84	98.7	4.0	100.0	0.0:1
85–89	99.6	4.0	100.0	0.0:1
90–94	99.9	4.0	100.0	0.0:1
95–100	100.0	4.0	100.0	0.0:1

$3.75/day 2005 PPP Poverty Line Tables

Figure 4 ($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	92.5
10–14	72.9
15–19	74.3
20–24	59.6
25–29	45.1
30–34	28.8
35–39	16.9
40–44	9.4
45–49	5.5
50–54	1.4
55–59	0.7
60–64	0.1
65–69	0.0
70–74	0.6
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Figure 7 ($3.75/day line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) with confidence intervals, 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	+1.8	2.2	2.5	3.4
5–9	–3.7	3.2	3.6	4.7
10–14	–12.3	8.4	8.7	9.5
15–19	+0.5	4.1	4.7	6.0
20–24	+7.3	3.9	4.8	6.5
25–29	–1.3	2.9	3.4	4.4
30–34	–5.2	3.8	4.1	4.5
35–39	+1.3	1.8	2.2	2.7
40–44	+0.4	1.2	1.5	2.0
45–49	+1.4	0.9	1.1	1.5
50–54	–0.2	0.5	0.6	0.9
55–59	+0.2	0.3	0.4	0.4
60–64	+0.1	0.1	0.1	0.1
65–69	+0.0	0.0	0.0	0.0
70–74	+0.6	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 8 ($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	+1.0	58.2	64.1	82.5
4	+0.5	29.2	35.0	44.8
8	+0.1	19.4	22.8	34.1
16	–0.1	13.1	15.8	22.1
32	–0.0	9.3	11.3	15.6
64	–0.0	6.9	8.2	10.2
128	–0.0	4.8	5.7	7.7
256	–0.1	3.4	4.1	5.5
512	–0.1	2.3	2.7	3.4
1,024	–0.1	1.6	1.9	2.6
2,048	–0.0	1.2	1.4	2.0
4,096	–0.0	0.8	1.0	1.3
8,192	–0.0	0.6	0.7	1.0
16,384	–0.0	0.4	0.5	0.7

Figure 11 ($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	13.4	0.0	86.3	86.6	–95.5
5–9	0.9	12.8	0.0	86.3	87.2	–86.3
10–14	1.9	11.7	0.2	86.1	88.1	–70.0
15–19	3.6	10.1	0.8	85.5	89.1	–41.5
20–24	5.5	8.2	2.4	83.9	89.4	–2.0
25–29	8.1	5.6	5.7	80.6	88.7	+58.3
30–34	10.7	2.9	11.6	74.7	85.5	+15.1
35–39	12.1	1.5	20.1	66.3	78.4	–46.7
40–44	13.0	0.6	30.7	55.6	68.7	–124.8
45–49	13.4	0.2	41.0	45.4	58.8	–199.9
50–54	13.6	0.0	51.5	34.8	48.4	–277.0
55–59	13.7	0.0	60.3	26.0	39.7	–341.4
60–64	13.7	0.0	67.5	18.8	32.5	–394.3
65–69	13.7	0.0	73.3	13.0	26.7	–436.7
70–74	13.7	0.0	78.2	8.1	21.8	–472.6
75–79	13.7	0.0	82.6	3.8	17.4	–504.3
80–84	13.7	0.0	85.0	1.3	15.0	–522.0
85–89	13.7	0.0	86.0	0.4	14.0	–529.1
90–94	13.7	0.0	86.2	0.1	13.8	–530.8
95–100	13.7	0.0	86.3	0.0	13.7	–531.8

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

Figure 12 ($3.75/day line): For a given score cut-off, the percentage of all households who are targeted (that is, have a score equal to or less than the cut-off), the percentage of targeted households who are poor (that is, below the poverty line), the percentage of poor households who are targeted, and the number of poor households who are successful targeted (coverage) per non-poor household mistakenly targeted (leakage), 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	96.8	2.2	29.9:1
5–9	1.0	95.9	6.7	23.3:1
10–14	2.1	90.8	14.3	9.8:1
15–19	4.4	81.8	26.3	4.5:1
20–24	7.9	69.1	40.1	2.2:1
25–29	13.8	58.7	59.3	1.4:1
30–34	22.3	48.1	78.6	0.9:1
35–39	32.2	37.7	88.8	0.6:1
40–44	43.8	29.8	95.5	0.4:1
45–49	54.4	24.7	98.3	0.3:1
50–54	65.1	20.9	99.7	0.3:1
55–59	74.0	18.5	100.0	0.2:1
60–64	81.2	16.8	100.0	0.2:1
65–69	87.0	15.7	100.0	0.2:1
70–74	91.9	14.9	100.0	0.2:1
75–79	96.2	14.2	100.0	0.2:1
80–84	98.7	13.9	100.0	0.2:1
85–89	99.6	13.7	100.0	0.2:1
90–94	99.9	13.7	100.0	0.2:1
95–100	100.0	13.7	100.0	0.2:1

$5.00/day 2005 PPP Poverty Line Tables

Figure 4 ($5.00/day line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	93.0
15–19	92.3
20–24	89.4
25–29	76.7
30–34	56.9
35–39	45.6
40–44	33.1
45–49	20.4
50–54	13.1
55–59	6.7
60–64	4.6
65–69	1.1
70–74	0.8
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Figure 7 ($5.00/day line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) with confidence intervals, 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	+3.4	2.9	3.7	4.5
10–14	–3.8	2.9	3.2	3.2
15–19	–1.2	2.1	2.4	3.2
20–24	+4.4	2.6	3.1	4.1
25–29	+0.2	2.6	3.1	4.4
30–34	–7.2	4.8	5.0	5.4
35–39	–3.3	2.8	3.0	3.5
40–44	–1.3	2.1	2.5	3.1
45–49	–5.3	3.7	3.9	4.2
50–54	–2.4	2.1	2.2	2.5
55–59	–0.1	1.3	1.6	2.0
60–64	–1.9	1.7	1.8	2.4
65–69	+0.2	0.5	0.6	0.8
70–74	–0.2	0.6	0.7	0.9
75–79	–0.1	0.1	0.1	0.1
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 8 ($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	+0.3	61.9	78.1	91.4
4	–0.9	32.5	39.3	54.0
8	–1.6	23.8	28.5	35.7
16	–1.7	16.4	19.2	26.0
32	–1.8	11.4	13.5	18.3
64	–2.0	8.6	10.2	14.3
128	–1.8	5.8	7.0	9.8
256	–1.9	4.3	5.1	6.9
512	–1.8	3.0	3.6	4.7
1,024	–1.8	2.1	2.4	3.3
2,048	–1.8	1.5	1.8	2.5
4,096	–1.8	1.1	1.4	1.7
8,192	–1.8	0.8	1.0	1.3
16,384	–1.8	0.6	0.7	0.8

Figure 11 ($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	29.1	0.0	70.6	70.9	–97.9
5–9	0.9	28.5	0.0	70.6	71.5	–93.6
10–14	2.1	27.3	0.1	70.6	72.7	–85.6
15–19	4.2	25.2	0.2	70.4	74.6	–70.9
20–24	7.1	22.3	0.8	69.8	76.9	–48.9
25–29	11.5	17.9	2.3	68.3	79.9	–13.8
30–34	16.7	12.6	5.6	65.0	81.8	+33.0
35–39	21.1	8.3	11.1	59.5	80.6	+62.2
40–44	24.6	4.8	19.1	51.5	76.1	+34.9
45–49	27.0	2.4	27.4	43.2	70.1	+6.6
50–54	28.4	1.0	36.7	33.9	62.3	–25.0
55–59	28.9	0.5	45.0	25.6	54.5	–53.3
60–64	29.3	0.1	51.9	18.7	48.0	–76.7
65–69	29.3	0.1	57.7	12.9	42.3	–96.3
70–74	29.4	0.0	62.5	8.1	37.5	–112.8
75–79	29.4	0.0	66.9	3.8	33.1	–127.5
80–84	29.4	0.0	69.3	1.3	30.7	–135.8
85–89	29.4	0.0	70.3	0.4	29.7	–139.1
90–94	29.4	0.0	70.5	0.1	29.5	–139.9
95–100	29.4	0.0	70.6	0.0	29.4	–140.3

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

Figure 12 ($5.00/day line): For a given score cut-off, the percentage of all households who are targeted (that is, have a score equal to or less than the cut-off), the percentage of targeted households who are poor (that is, below the poverty line), the percentage of poor households who are targeted, and the number of poor households who are successful targeted (coverage) per non-poor household mistakenly targeted (leakage), 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	1.0	97.9	3.2	47.6:1
10–14	2.1	97.4	7.1	37.6:1
15–19	4.4	94.8	14.2	18.1:1
20–24	7.9	89.6	24.2	8.7:1
25–29	13.8	83.6	39.3	5.1:1
30–34	22.3	75.0	57.0	3.0:1
35–39	32.2	65.5	71.8	1.9:1
40–44	43.8	56.3	83.8	1.3:1
45–49	54.4	49.6	91.8	1.0:1
50–54	65.1	43.6	96.7	0.8:1
55–59	74.0	39.1	98.5	0.6:1
60–64	81.2	36.0	99.6	0.6:1
65–69	87.0	33.7	99.8	0.5:1
70–74	91.9	32.0	100.0	0.5:1
75–79	96.2	30.5	100.0	0.4:1
80–84	98.7	29.8	100.0	0.4:1
85–89	99.6	29.5	100.0	0.4:1
90–94	99.9	29.4	100.0	0.4:1
95–100	100.0	29.4	100.0	0.4:1

Laeken Poverty Line Tables

Figure 4 (Laeken 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	87.6
10–14	78.5
15–19	79.9
20–24	62.6
25–29	56.6
30–34	49.1
35–39	33.1
40–44	29.3
45–49	14.3
50–54	10.0
55–59	5.9
60–64	4.9
65–69	2.2
70–74	3.2
75–79	0.3
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Figure 7 (Laeken line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) with confidence intervals, 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	+3.0	3.4	3.8	4.9
5–9	–4.1	4.2	5.0	6.9
10–14	–7.6	5.8	6.2	6.8
15–19	+4.1	4.1	4.9	6.4
20–24	+6.7	3.7	4.4	6.1
25–29	+1.1	3.0	3.5	4.6
30–34	+4.0	2.6	3.1	3.8
35–39	+6.4	2.0	2.3	3.1
40–44	+8.4	1.7	1.9	2.6
45–49	+0.6	1.5	1.8	2.4
50–54	+2.9	1.1	1.2	1.6
55–59	+1.8	0.9	1.1	1.4
60–64	+2.0	0.8	1.0	1.3
65–69	+1.2	0.4	0.5	0.7
70–74	+2.8	0.3	0.3	0.4
75–79	–0.7	0.8	0.9	1.2
80–84	–0.6	0.6	0.7	0.8
85–89	–0.4	0.5	0.6	0.8
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 8 (Laeken 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	63.6	71.2	87.2
4	+2.8	32.4	37.7	51.9
8	+2.5	22.1	26.9	34.6
16	+2.5	15.8	18.5	25.2
32	+2.7	11.0	13.9	17.6
64	+2.8	7.6	9.1	12.9
128	+2.9	5.5	6.6	8.5
256	+2.9	3.9	4.7	5.9
512	+2.9	2.7	3.3	4.2
1,024	+2.9	1.9	2.3	2.9
2,048	+3.0	1.4	1.7	2.1
4,096	+3.0	1.0	1.2	1.5
8,192	+3.0	0.7	0.8	1.1
16,384	+3.0	0.5	0.6	0.7

Figure 11 (Laeken 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	23.4	0.0	76.3	76.6	–97.4
5–9	0.9	22.9	0.1	76.2	77.1	–92.3
10–14	1.9	21.8	0.3	76.0	77.9	–83.0
15–19	3.6	20.1	0.8	75.5	79.1	–66.3
20–24	5.7	18.0	2.2	74.1	79.8	–42.5
25–29	9.1	14.6	4.7	71.6	80.7	–3.4
30–34	13.4	10.4	9.0	67.3	80.7	+50.5
35–39	16.8	6.9	15.4	60.9	77.7	+35.1
40–44	20.1	3.7	23.7	52.6	72.6	+0.0
45–49	21.8	1.9	32.6	43.6	65.4	–37.6
50–54	22.8	0.9	42.3	33.9	56.7	–78.5
55–59	23.2	0.5	50.7	25.5	48.8	–113.9
60–64	23.5	0.2	57.7	18.6	42.1	–143.2
65–69	23.6	0.1	63.4	12.9	36.5	–167.2
70–74	23.7	0.1	68.3	8.0	31.7	–187.7
75–79	23.7	0.0	72.5	3.7	27.4	–205.8
80–84	23.7	0.0	74.9	1.3	25.0	–216.0
85–89	23.7	0.0	75.9	0.4	24.1	–220.0
90–94	23.7	0.0	76.1	0.1	23.9	–221.0
95–100	23.7	0.0	76.3	0.0	23.7	–221.6

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

Figure 12 (Laeken line): For a given score cut-off, the percentage of all households who are targeted (that is, have a score equal to or less than the cut-off), the percentage of targeted households who are poor (that is, below the poverty line), the percentage of poor households who are targeted, and the number of poor households who are successful targeted (coverage) per non-poor household mistakenly targeted (leakage), 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	95.8	1.3	23.1:1
5–9	1.0	91.2	3.7	10.4:1
10–14	2.1	88.2	8.0	7.5:1
15–19	4.4	81.8	15.2	4.5:1
20–24	7.9	72.1	24.1	2.6:1
25–29	13.8	66.0	38.4	1.9:1
30–34	22.3	59.8	56.3	1.5:1
35–39	32.2	52.2	70.8	1.1:1
40–44	43.8	45.8	84.6	0.8:1
45–49	54.4	40.0	91.8	0.7:1
50–54	65.1	35.0	96.1	0.5:1
55–59	74.0	31.4	98.0	0.5:1
60–64	81.2	29.0	99.2	0.4:1
65–69	87.0	27.1	99.6	0.4:1
70–74	91.9	25.7	99.7	0.3:1
75–79	96.2	24.6	99.9	0.3:1
80–84	98.7	24.0	100.0	0.3:1
85–89	99.6	23.8	100.0	0.3:1
90–94	99.9	23.8	100.0	0.3:1
95–100	100.0	23.7	100.0	0.3:1