Simple Poverty Scorecard® Poverty-Assessment Tool Kyrgyz Republic

Mark Schreiner

13 February 2015

Этот документ доступен на русском языке на SimplePovertyScorecard.com
This document is in English at `SimplePovertyScorecard.com`

Abstract

The Simple Poverty Scorecard poverty-assessment tool uses 10 low-cost indicators from the Kyrgyz Republic's 2012 Integrated Household Survey to estimate the likelihood that a household has consumption below a given poverty line. Field workers can collect responses in about ten minutes. The scorecard's accuracy is reported for a range of poverty lines. The scorecard is a practical way for pro-poor programs in the Kyrgyz Republic to measure poverty rates, to track changes in poverty rates over time, and to segment clients for differentiated treatment.

Acknowledgements

This paper was sponsored by Grameen Foundation (GF) with funding from the First MicroCredit Company (FMCC), fully owned by the Aga Khan Agency for Microfinance (AKAM). Data are from the Kyrgyz Republic's National Statistics Committee. Thanks go to Aibek Baibagysh uulu, Frank Ballard, Damir Esenaliev, Liliya Isambaeva, Oksana Kovalenko, Rishat Muhamedjanov, Julie Peachey, Sharada Ramanathan, and Matt Walsh. This scorecard was re-branded by Grameen Foundation (GF) as a Progress out of Poverty Index® tool. The PPI® is a performance-management tool that GF promotes to help organizations achieve their social objectives more effectively. "Progress out of Poverty Index" and "PPI" are Registered Trademarks of Innovations for Poverty Action. "Simple Poverty Scorecard" is a Registered Trademark of Microfinance Risk Management, L.L.C. for its brand of poverty-assessment tools.

Simple Poverty Scorecard® Poverty-Assessment Tool

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

Indicator	Response	Points	Score
1. In which oblast does the household reside?	A. Jalal-Abad	0	
	B. Naryn	1	
	C. Osh	2	
	D. Bishkek	5	
	E. Issykul	6	
	F. Talas	7	
	G. Chui	8	
	H. Batken	11	
2. How many household members are there?	A. Seven or more	0	
	B. Six	7	
	C. Five	13	
	D. Four	19	
	E. Three	27	
	F. Two	35	
	G. One (stop interview, score is 100)	100	
3. In the past 7 days, how many household members worked or had paid employment for at least 1 hour, or worked on a family farm or enterprise, or (if they did not work in past 7 days) had work or paid employment to which they plan to return?	A. None, or one	0	
	B. Two	2	
	C. Three or more	5	
4. In their main work or paid employment in the past 7 days, how many household members worked for a wage paid in-cash or in-kind, or for a money allowance?	A. None	0	
	B. One	3	
	C. Two or more	4	
5. What is the main source of water used by the household?	A. Public (communal) water pump, storage reservoir, river, lake, pond, *aryk*, spring, or purchased water (water cart)	0	
	B. Artesian well	2	
	C. Private water pump	3	
	D. Well, or aqueduct (running water)	8	
6. Does the household have any regular or automatic washing machines?	A. No	0	
	B. Regular (but not automatic)	4	
	C. Automatic (regardless of regular)	7	
7. Does the household have any electric heaters?	A. No	0	
	B. Yes	4	
8. How many cellular telephones does the household have?	A. None, or one	0	
	B. Two	4	
	C. Three or more	9	
9. Does the household have any bicycles or any automobiles, trucks, minivans, motorcycles, scooters, mopeds, or motorized bicycles?	A. No	0	
	B. Only bicycle	1	
	C. Motorized vehicle (regardless of bicycle)	7	
10. Does the household use any personal agricultural plots? If so, has the household in the past 12 months had any sheep, lambs, goats, kids, cows, heifers, calves, bulls, oxen, horses, donkeys, hinnies, mules, or yaks (of any age)?	A. No plot (regardless of animals)	0	
	B. Has a plot, but no animals	2	
	C. Has both a plot and animals	9	

Score:

Back-page Worksheet:
Household Members, Age, Work Status, and Salary Status

Write down the name and identification number of the client and of yourself as the enumerator, as well as the service point that the client uses and the service point from which you work. Record the date of the interview and the date when the client first participated with the organization. For the first indicator, mark the oblast where the participant resides. Then read to the respondent: *Please tell me the first name and age of all members of your household. A* household *is one or more individuals—with or without blood or marital ties—who share a residence and who together contribute resources to provide for the daily needs of the household members.* Record the first names and ages of all members. Write the total number of members in the scorecard header next to "# Household members:" and mark the second indicator. If there is only one member, you may stop the interview (if desired); the score for a one-member household is 100, regardless of all other responses.

For each member 15-years-old or older, ask: *In the past 7 days, did <name> work or have paid employment for at least 1 hour, or worked on a family farm or enterprise, or (if <name> did not work) had work or paid employment to which he/she plans to return?* Count the number who work, and mark the response for indicator 3.

For each member who works, ask: *Did <name> receive a wage paid in-cash or in-kind, or a money allowance?* Count the number of members who work for a wage or salary, and mark indicator 4.

First name	Age	If <name> is 15-years-old or older, did he/she in the past 7 days work or have paid employment for at least 1 hour, worked on a family farm or enterprise, or (if <name> did not work) have work or paid employment to which he/she plans to return?			If <name> works, does he/she receive a wage paid in-cash or in-kind, or a money allowance?		
1.		Not ≥15	No	Yes	Didn't work	No	Yes
2.		Not ≥15	No	Yes	Didn't work	No	Yes
3.		Not ≥15	No	Yes	Didn't work	No	Yes
4.		Not ≥15	No	Yes	Didn't work	No	Yes
5.		Not ≥15	No	Yes	Didn't work	No	Yes
6.		Not ≥15	No	Yes	Didn't work	No	Yes
7.		Not ≥15	No	Yes	Didn't work	No	Yes
8.		Not ≥15	No	Yes	Didn't work	No	Yes
9.		Not ≥15	No	Yes	Didn't work	No	Yes
10.		Not ≥15	No	Yes	Didn't work	No	Yes
11.		Not ≥15	No	Yes	Didn't work	No	Yes
12.		Not ≥15	No	Yes	Didn't work	No	Yes
13.		Not ≥15	No	Yes	Didn't work	No	Yes
# members:				# "Yes":		# "Yes":	

Look-up table for converting scores to poverty likelihoods

	Poverty likelihood (%)							
	National poverty lines				Intl. 2005 PPP			
Score	**100%**	**150%**	**200%**	**Median**	**$1.25**	**$2.00**	**$2.50**	**$5.00**
0–4	100.0	100.0	100.0	91.5	5.8	88.9	100.0	100.0
5–9	98.0	100.0	100.0	88.6	1.6	71.4	92.0	100.0
10–14	96.6	100.0	100.0	59.9	1.4	51.3	78.7	100.0
15–19	89.9	100.0	100.0	55.3	1.4	38.7	74.5	100.0
20–24	77.3	99.8	100.0	39.5	1.4	24.3	58.4	100.0
25–29	68.0	99.1	100.0	30.5	1.1	15.8	45.7	100.0
30–34	46.5	96.1	100.0	20.2	0.9	9.3	27.1	99.7
35–39	40.7	89.4	99.9	12.6	0.3	4.5	19.2	96.0
40–44	15.8	76.6	95.6	6.2	0.1	3.9	10.0	92.1
45–49	11.7	65.2	90.5	5.4	0.1	2.5	7.3	85.2
50–54	4.1	43.9	80.5	1.2	0.0	0.2	3.1	69.5
55–59	3.9	27.3	60.2	0.7	0.0	0.2	2.2	50.3
60–64	2.9	21.4	51.0	0.4	0.0	0.2	1.6	39.8
65–69	1.3	6.9	38.8	0.4	0.0	0.2	0.9	25.0
70–74	1.2	4.5	31.6	0.4	0.0	0.2	0.9	19.6
75–79	1.2	4.5	18.9	0.4	0.0	0.2	0.9	12.1
80–84	1.2	4.5	18.7	0.4	0.0	0.2	0.9	12.1
85–89	1.2	4.5	18.7	0.4	0.0	0.2	0.9	12.1
90–94	1.2	4.5	18.7	0.4	0.0	0.2	0.9	12.1
95–100	1.2	4.5	18.7	0.4	0.0	0.2	0.9	12.1

Simple Poverty Scorecard® Poverty-Assessment Tool Kyrgyz Republic

1. Introduction

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

The direct approach to poverty measurement via surveys is difficult and costly. As a case in point, the 2012 Kyrgyz Integrated Household Survey (KIHS) demands a great deal of time and effort from both enumerators and respondents. The hundreds of items in its 92 pages are asked of sampled households four times in a given calendar year. In addition, households complete a 3-month diary of non-food consumption items and a two-week diary of food items. The typical household is in the KIHS sample for three years.

In comparison, the indirect approach via the scorecard is simple, quick, and inexpensive. It uses ten verifiable indicators (such as "What is the main source of water used by the household?" and "Does the household have any regular or automatic washing machines?") to get a score that is highly correlated with poverty status as measured by the exhaustive KIHS survey.

The scorecard differs from "proxy-means tests" (Coady, Grosh, and Hoddinott, 2004) in that it is transparent, it is freely available,[1] and it is tailored to the capabilities and purposes not of national governments but rather of local, pro-poor organizations. The feasible poverty-measurement options for local organizations are typically blunt (such as rules based on land-ownership or housing quality) or subjective and relative (such as participatory wealth ranking facilitated by skilled field workers). Estimates from these approaches may be costly, their accuracy is unknown, and they are not comparable across places, organizations, nor time.

The scorecard can be used to measure the share of a program's participants who live in households whose per-capita consumption is below a given poverty line, for example, the Millennium Development Goals' $1.25/day per-person line at 2005 purchase-power parity (PPP). USAID microenterprise partners in the Kyrgyz Republic can use scoring with the median poverty line to report how many of their participants are "very poor".[2] Scoring can also be used to measure net movement across a poverty line over time. In all these cases, the scorecard provides a consumption-based, objective tool with known accuracy. While consumption surveys are costly even for governments,

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

[2] USAID defines a household as "very poor" if its daily per-capita consumption is below the highest of the $1.25/day 2005 PPP line (KGS32.55 in the Kyrgyz Republic in 2012, Figure 1) or the median line that divides people in households below 100% of the Kyrgyz Republic's national poverty line into two equal-size groups (KGS60.06). The scorecard is approved for use by USAID's microenterprise partners (USAID, 2013, p. 7) when re-branded as a Progress out of Poverty Index®.

some local pro-poor organizations may be able to implement an inexpensive poverty-assessment tool to help with poverty monitoring or (if desired) with segmenting participants for targeted services.

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

Beyond its simplicity and transparency, the scorecard's technical approach is innovative in how it associates scores with poverty likelihoods, in the extent of its accuracy tests, and in how it derives formulas for standard errors. Although the accuracy tests are simple and commonplace in statistical practice and in the for-profit field of credit-risk scoring, they have rarely been applied to poverty-assessment tools.

The scorecard is constructed with data from the 2012 KIHS from the Kyrgyz Republic's National Statistics Committee (NSC). Indicators are selected to be:

- Inexpensive to collect, easy to answer quickly, and simple to verify
- Strongly correlated with poverty
- Liable to change over time as poverty status changes
- Applicable across all regions in the Kyrgyz Republic

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

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

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

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

The scorecard can also be used to target services to different segments of participants. To help managers choose an appropriate targeting cut-off for their purposes, this paper reports several measures of targeting accuracy for a range of possible cut-offs.

This paper presents a single scorecard whose indicators and points are derived from household consumption data from the 2012 KIHS applied with the Kyrgyz Republic's national poverty line. Scores from this one scorecard are calibrated—again using data from the 2012 KIHS—to poverty likelihoods for eight poverty lines.

The scorecard is constructed and calibrated using half of the data from the 2012 KIHS. The other half is used to validate the scorecard's accuracy for estimating households' poverty likelihoods, for estimating groups' poverty rates at a point in time, and for targeting.

All three of scoring's estimates are *unbiased*. That is, they match the true value on average in repeated samples when constructed from (and applied to) a single, unchanging population at a point in time. Like all predictive models, the specific scorecard here is constructed from a single sample and so misses the mark to some unknown extent when applied to a different population or when applied after 2012.[3]

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

[3] Important examples include nationally representative samples at a different point in time or sub-groups that are not nationally representative (Diamond *et al.*, 2016; Tarozzi and Deaton, 2009).

When applied to the validation sample with 1,000 bootstraps of $n = 16{,}384$, the average difference between scorecard estimates of groups' poverty rates and the true rates at a point in time with the national poverty line is –2.3 percentage points. The average absolute difference across all eight poverty lines is about 1.6 percentage points, and the maximum absolute difference is 2.5 percentage points. These differences are due to sampling variation rather than bias; the average difference for a given poverty line would be zero if the whole 2012 KIHS were to be repeatedly redrawn and divided into sub-samples before repeating the entire process of constructing and validating scorecards.

The 90-percent confidence intervals for these estimates with $n = 16{,}384$ are ± 0.8 percentage points or less. For $n = 1{,}024$, the 90-percent intervals are ± 3.0 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 over time, and Section 8 covers targeting. Section 9 places the scorecard here in the context of similar exercises for the Kyrgyz Republic. The last section is a summary.

2. Data and poverty lines

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

2.1 Data

The scorecard is based on data from 4,936 households in the 2012 KIHS fielded by the Kyrgyz Republic's NSC during calendar-year 2012.[4] This is most recent national consumption survey whose data is available for the Kyrgyz Republic.

For the purposes of the scorecard, the households in the 2012 KIHS are randomly divided into two sub-samples:

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

2.2 Poverty rates

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

[4] The full 2012 KIHS has $n = 5{,}006$, but 70 cases are omitted here because they lack data on asset ownership and on the characteristics of the residence.

Suppose a program serves two households. The first household is poor (its per-capita consumption is less than a given poverty line), and it has three members, one of whom is a program participant. The second household is non-poor and has four members, two of whom are program participants.

Poverty rates are in terms of either households or people. If the program defines its *participants* as households, then the household level is relevant. The estimated household-level poverty rate is the equal-weighted average of poverty statuses (or estimated poverty likelihoods) across households with participants. In the example here, this is $\frac{1 \cdot 1 + 1 \cdot 0}{1 + 1} = \frac{1}{2} = 0.5 = 50$ percent. In the "$1 \cdot 1$" term in the numerator, the first "1" is the first household's weight, and the second "1" is the first household's poverty status (poor). In the "$1 \cdot 0$" term in the numerator, the "1" is the second household's weight, and the "0" is the second household's poverty status (non-poor). The "$1 + 1$" in the denominator is the sum of the weights of the two households. Each household has a weight of one (1) because the unit of analysis is the household.

Alternatively, a person-level rate is relevant if a program defines all people in households that benefit from its services as *participants*. In the example here, the person-level rate is the household-size-weighted average of poverty statuses for households with participants, or $\frac{3 \cdot 1 + 4 \cdot 0}{3 + 4} = \frac{3}{7} = 0.43 = 43$ percent. In the "$3 \cdot 1$" term in the numerator, the "3" is the first household's weight because it has three members, and the "1" is its poverty status (poor). In the "$4 \cdot 0$" term in the numerator, the "4" is

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

As a final example—one that pertains to what is likely the most common situation in practice—a program counts as *participants* only those household members with whom it deals with directly. For the example here, this means that some—but not all—household members are counted. The person-level rate is now the participant-weighted average of the poverty statuses of households with participants, or $\frac{1 \cdot 1 + 2 \cdot 0}{1 + 2} = \frac{1}{3} = 0.33 = 33$ percent. The first "1" in the "$1 \cdot 1$" in the numerator is the first household's weight because it has one participant, and the second "1" is its poverty status (poor). In the "$2 \cdot 0$" term in the numerator, the "2" is the second household's weight because it has two participants, and the zero is its poverty status (non-poor). The "$1 + 2$" in the denominator is the sum of the weights of the two households. Each household's weight is its number of participants because the unit of analysis is the participant.

To sum up, estimated poverty rates are weighted averages of households' poverty statuses (or estimated poverty likelihoods), where the weights are the number of relevant units in the household. When reporting, organizations should explain who is counted as a *participant* and why.

For all eight poverty lines, Figure 1 reports poverty rates in the 2012 KIHS for both households and people in the Kyrgyz Republic. Person-level poverty rates are included in Figure 1 because these are the rates reported by governments and used in most policy discussions. Household-level poverty rates are also reported because—as discussed above—household-level poverty likelihoods can be straightforwardly converted into poverty rates for other units of analysis. This is also the reason why the scorecard is constructed, calibrated, and validated with household weights.

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

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

The Kyrgyz Republic's definition of *consumption*—as outlined in Kyrgyz Republic (2011)—follows international common practice (World Bank, 2013; Deaton and Zaidi, 2002). In particular, it includes cash spent on food and non-food consumables, as well as the value of own-produced food and the use-value of consumer durables and of owner-occupied housing.

The definition of the Kyrgyz Republic's national poverty line begins with a minimum standard for food, taken to be the cost of the basket—scaled to provide 2,100 Calories—observed to be eaten by households whose food consumption is in the second to fourth deciles (Gassmann, 2013). The cost of this food standard is adjusted for price differences across 15 poverty-line regions (urban/rural in the Kyrgyz Republic's eight oblasts except for the all-urban Bishkek).

The national poverty line is defined as the minimum food standard, plus a minimum non-food standard that is taken as the average value of non-food consumption by households whose food consumption is in a range 10 percent above or below the food standard.[5] The non-food standard is not adjusted for regional price differences.

For 2012, the average value of the national poverty line (sometimes called here "100% of the national line") across all of the Kyrgyz Republic is KGS72.40 (Figure 1), giving poverty rates of 28.5 percent for households and 38.1 percent for people.[6]

[5] Gassmann (2013), citing Tsirunyan (2012).

[6] The NSC (2013) reports 38.0 percent because—unlike this paper—it does not omit the 70 households who lack data on consumer durables and on housing characteristics. The algorithm used here matches the NSC's figure when it is run on the full sample.

Because local, pro-poor programs in the Kyrgyz Republic may want to use different or various poverty lines, this paper calibrates scores from its single scorecard to poverty likelihoods for eight lines:[7]

- 100% of national
- 150% of national
- 200% of national
- Median
- $1.25/day 2005 PPP
- $2.00/day 2005 PPP
- $2.50/day 2005 PPP
- $5.00/day 2005 PPP

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

The international 2005 PPP lines are derived from:

- 2005 PPP exchange rate of KGS12.998 per $1.00 (World Bank, 2008)
- Consumer Price Index for all of the Kyrgyz Republic:[8]
  - Average in 2005: 100.000
  - Average in 2012: 200.315
- Average national line for all of the Kyrgyz Republic (Figure 1): KGS72.40
- National lines in each of the 15 poverty-line regions (Figure 2)

[7] Figure 2 reports the values of the eight lines and the corresponding poverty rates for households and people in each of the Kyrgyz Republic's 15 poverty-line regions.
[8] `stat.kg/images/stories/baza/price/10801_4.xls`, retrieved 12 February 2015.

Using Sillers (2006), the all-Kyrgyz Republic $1.25/day 2005 PPP line is:

$$\left(\text{2005 PPP exchange rate}\right) \cdot \$1.25 \cdot \left(\frac{\text{CPI}_{2012}}{\text{CPI}_{2005}}\right) =$$

$$\left(\frac{\text{KGS}12.998}{\$1.00}\right) \cdot \$1.25 \cdot \left(\frac{200.315}{100.000}\right) = \text{KGS}32.55.$$

This line applies to the Kyrgyz Republic as a whole.[9] This average $1.25/day line is adjusted for cost-of-living differences across poverty-line regions by multiplying it by the value of the national line in a given region and then dividing by the average all-Kyrgyz Republic national line. For the example of rural Issykul, this is the all-Kyrgyz Republic $1.25/day line (KGS32.55) multiplied by the value of the national line in rural Issykul (KGS68.28), divided by the average all-Kyrgyz Republic national line (KGS72.40), or 32.55 x (68.28 ÷ 72.40) = KGS30.70 (Figure 2).

USAID microenterprise partners in the Kyrgyz Republic who use the scorecard to report the share of their participants who are "very poor" to USAID should use the median poverty line. This is because USAID defines the "very poor" as those people in households whose consumption is below the highest of two lines:

- $1.25/day 2005 PPP (KGS32.55)
- Median line (KGS60.06, Figure 1).

[9] The latest person-level poverty rate for the $1.25/day line from the World Bank's PovcalNet (`iresearch.worldbank.org/PovcalNet`, retrieved 12 February 2015) is 5.1 percent, based on the 2011 KIHS. This is far from the 0.6 percent in Figure 1 for the 2012 KIHS. Schreiner (2014) argues that estimates—like this one—in the documentation of the scorecard are to be preferred because PovcalNet does not document its poverty line in KGS, how it deflates the 2005 PPP factor to 2011, nor its regional-price adjustments.

3. Scorecard construction

For the Kyrgyz Republic, about 100 candidate indicators are initially prepared in the areas of:

- Household composition (such as the number of members)
- Education (such as the highest level attained by the female head/spouse)
- Housing (such as the source of water)
- Ownership of durable assets (such as washing machines or cellular telephones)
- Employment (such as the number of household members who work)
- Agriculture (such as whether the household has a plot and large livestock)

Figure 3 lists the candidate indicators, ordered by the entropy-based "uncertainty coefficient" (Goodman and Kruskal, 1979) that measures how well a given indicator predicts poverty on its own.[10]

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

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

[10] The uncertainty coefficient is not used as a criterion when selecting scorecard indicators; it is just a way to order the candidate indicators in Figure 3.

One of these one-indicator scorecards is then selected based on several factors (Schreiner *et al.*, 2014; Zeller, 2004). These include improvement in accuracy, likelihood of acceptance by users (determined by simplicity, cost of collection, and "face validity" in terms of experience, theory, and common sense), sensitivity to changes in poverty, variety among indicators, applicability across regions, relevance for distinguishing among households at the poorer end of the consumption distribution, expected stability of the relationship between the indicator and poverty as time passes, and verifiability.

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

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

This algorithm is similar to the common R^2-based stepwise least-squares regression. It differs from naïve stepwise in that the selection of indicators considers both statistical[11] and non-statistical criteria. The non-statistical criteria can improve robustness through time and help ensure that indicators are simple, sensible, and acceptable to users.

The single scorecard here applies to all of the Kyrgyz Republic. Tests for Indonesia (World Bank, 2012), Bangladesh (Sharif, 2009), India and Mexico (Schreiner, 2006 and 2005a), Sri Lanka (Narayan and Yoshida, 2005), and Jamaica (Grosh and Baker, 1995) suggest that segmenting poverty-assessment tools by urban/rural does not improve targeting accuracy much. In general, segmentation may improve the accuracy of estimates of poverty rates (Diamond *et al.*, 2016; Tarozzi and Deaton, 2009), but segmentation may also increase the risk of overfitting (Haslett, 2012).

[11] The statistical criterion for selecting an indicator is not the p values of its coefficients but rather its contribution to the ranking of households by poverty status.

4. Practical guidelines for scorecard use

The main challenge of scorecard design is not to maximize statistical accuracy but rather to improve the chances that the scorecard is actually used (Schreiner, 2005b). When scoring projects fail, the reason is not usually statistical inaccuracy but rather the failure of an organization to decide to do what is needed to integrate scoring in its processes and to train and convince its employees to use the scorecard properly (Schreiner, 2002). After all, most reasonable scorecards have similar targeting accuracy, thanks to the empirical phenomenon known as the "flat maximum" (Caire and Schreiner, 2012; Hand, 2006; Baesens *et al.*, 2003; Lovie and Lovie, 1986; Kolesar and Showers, 1985; Stillwell, Barron, and Edwards, 1983; Dawes, 1979; Wainer, 1976; Myers and Forgy, 1963). The bottleneck is less technical and more human, not statistics but organizational-change management. Accuracy is easier to achieve than adoption.

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

To this end, the scorecard for the Kyrgyz Republic fits on one page. The construction process, indicators, and points are simple and transparent. Additional work is minimized; non-specialists can compute scores by hand in the field because the scorecard has:

- Only 10 indicators
- Only "multiple-choice" indicators
- Only simple weights (non-negative integers, and no arithmetic beyond addition)

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

- Record the names and identifiers of the participant, of the field worker, and of the relevant organizational service points
- Record the date that the participant first participated with the organization
- Record the date of the scorecard interview
- Complete the back-page worksheet with each household member's:
    - First name
    - Age
    - Work status
    - Salary status
- Record household size in the scorecard header, and mark the responses to scorecard's second, third, and fourth indicators based on the back-page worksheet
- If the household has only a single member, then stop the interview (if desired) and record the score as 100
- If the household has more than one member, then read each of the remaining six questions one-by-one from the scorecard, drawing a circle around the responses and their points, and writing each point value in the far right-hand column
- Add up the points to get a total score (which is 100 if the household has one member, regardless of all other responses)
- Implement targeting policy (if any)
- Deliver the paper scorecard to a central office for data entry and filing

Of course, field workers must be trained. The quality of outputs depends on the quality of inputs. If organizations or field workers gather their own data and believe that they have an incentive to exaggerate poverty rates (for example, if funders reward them for higher poverty rates), then it is wise to do on-going quality control via data review and random audits (Matul and Kline, 2003).[12] IRIS Center (2007a) and Toohig (2008) are useful nuts-and-bolts guides for budgeting, training field workers and supervisors, logistics, sampling, interviewing, piloting, recording data, and controlling quality.

In particular, while collecting scorecard indicators is relatively easier than alternative ways of measuring poverty, it is still absolutely difficult. Training and explicit definitions of terms and concepts in the scorecard are essential, and field workers should scrupulously study and follow the "Guidelines for the Interpretation of Scorecard Indicators" found at the end of this paper, as these guidelines—along with the "Backpage Worksheet"—are an integral part of the Simple Poverty Scorecard tool.[13]

[12] If a program does not want field workers to know the points associated with responses, then it can erase the points from the paper scorecard and then apply the points later at a central office. Schreiner (2012b) argues that hiding points in Colombia (Camacho and Conover, 2011) did little to deter cheating and that, in any case, cheating by the user's central office was more damaging than cheating by field workers and respondents. Even if points are hidden, field workers and respondents can apply common sense to guess how response options are linked with poverty.

[13] The "Guidelines" here are the only ones that organizations should give to field workers. All other issues of interpretation should be left to the judgment of field workers and respondents, as this seems to be what the NSC did in the 2012 KIHS.

For the example of Nigeria, one study (Onwujekwe, Hanson, and Fox-Rushby, 2006) found distressingly low inter-rater and test-retest correlations for indicators as seemingly simple as whether the household owns an automobile. At the same time, Grosh and Baker (1995) suggest that gross underreporting of assets does not affect targeting. For the first stage of targeting in a conditional cash-transfer program in Mexico, Martinelli and Parker (2007, pp. 24–25) find that "underreporting [of asset ownership] is widespread but not overwhelming, except for a few goods . . . [and] overreporting is common for a few goods, which implies that self-reporting may lead to the exclusion of deserving households". Still, as is done in Mexico in the second stage of its targeting process, most false self-reports can be corrected (or avoided in the first place) by field workers who make a home visit. This is the recommended procedure for local, pro-poor organizations who use scoring for targeting in the Kyrgyz Republic.

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

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

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

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

- Employees of the organization
- Third parties

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

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

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

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

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

The frequency of application can be:

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

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

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

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

5. Estimates of household poverty likelihoods

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

To get absolute units, scores must be converted to *poverty likelihoods*, that is, probabilities of being below a poverty line. This is done via simple look-up tables. For the example of 100% of the national line, scores of 30–34 have a poverty likelihood of 46.5 percent, and scores of 35–39 have a poverty likelihood of 40.7 percent (Figure 4).

The poverty likelihood associated with a score varies by poverty line. For example, scores of 30–34 are associated with a poverty likelihood of 46.5 percent for 100% of the national line but of 27.1 percent for the $2.50/day 2005 PPP line.[14]

5.1 Calibrating scores with poverty likelihoods

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

[14] Starting with Figure 4, many figures have eight versions, one for each of the eight poverty lines. To keep them straight, they are grouped by poverty line. Single tables that pertain to all lines are placed with the tables for 100% of the national line.

For the example of 100% of the national line (Figure 5), there are 14,229 (normalized) households in the calibration sub-sample with a score of 30–34. Of these, 6,618 (normalized) are below the poverty line. The estimated poverty likelihood associated with a score of 30–34 is then 46.5 percent, because 6,618 ÷ 14,229 = 0.465.

To illustrate with 100% of the national line and a score of 35–39, there are 10,481 (normalized) households in the calibration sample, of whom 4,271 (normalized) are below the line (Figure 5). The poverty likelihood for this score is then 4,271 ÷ 10,481 = 0.407.

The same method is used to calibrate scores with estimated poverty likelihoods for the other eight poverty lines.[15]

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

For all scores, Figure 6 shows the likelihood that a given household's consumption falls in a range demarcated by two adjacent poverty lines. For example, the daily per-capita consumption of a household with a score of 30–34 falls in the following ranges with probability:

- 0.9 percent below $1.25/day
- 8.4 percent between $1.25/day and $2.00/day
- 11.0 percent between $2.00/day and the median line
- 6.9 percent between the median line and $2.50/day
- 19.4 percent between $2.50/day and 100% of the national line
- 49.6 percent between 100% and 150% of the national line
- 3.6 percent between 150% of the national line and $5.00/day
- 0.3 percent between $5.00/day and 200% of the national line
- 0.0 percent above 200% of the national line

Even though the scorecard is constructed partly based on judgment related to non-statistical criteria, the calibration process produces poverty likelihoods that are objective, that is, derived from quantitative poverty lines and from survey data on consumption. The calibrated poverty likelihoods would be objective even if the process of selecting indicators and points for the scorecard did not use any data at all. In fact, objective scorecards of proven accuracy are often constructed using only expert judgment to select indicators and points (Fuller, 2006; Caire, 2004; Schreiner *et al.*, 2014). Of course, the scorecard here is constructed with both data and judgment. The fact that this paper acknowledges that some choices in scorecard construction—as in most statistical analysis—are informed by judgment in no way impugns the objectivity of the poverty likelihoods, as this depends on using data in score calibration, not on using data (and nothing else) in scorecard construction.

Although the points in the Kyrgyz Republic's scorecard are transformed coefficients from a Logit regression, (untransformed) 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. Going from scores to poverty likelihoods in this way requires no arithmetic at all, just a look-up table. This approach to calibration can also improve accuracy, especially with large samples.

5.2 Accuracy of estimates of households' poverty likelihoods

As long as the relationships between indicators and poverty do not change over time, and as long as the scorecard is applied to households that are representative of the same population from which the scorecard was originally constructed, then this calibration process produces unbiased estimates of poverty likelihoods. *Unbiased* means that in repeated samples from the same population, the average estimate matches the true value. The scorecard also produces unbiased estimates of poverty rates at a point in time and unbiased estimates of changes in poverty rates between two points in time.[16]

Of course, the relationships between indicators and poverty do change to some unknown extent over time and also across sub-groups in the population of the Kyrgyz

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

Republic. Thus, the scorecard will generally be biased when applied after December 2012 (the last month of fieldwork for the 2012 KIHS) or when applied with sub-groups that are not nationally representative.

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

- Score each household in the validation sample
- Draw a new bootstrap sample *with replacement* from the validation sample
- For each score, compute the true poverty likelihood in the bootstrap sample, that is, the share of households with the score and with consumption below a poverty line
- For each score, record the difference between the estimated poverty likelihood (Figure 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 intervals containing the central 900, 950, and 990 differences between estimated and true poverty likelihoods

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

For 100% of the national line, the average poverty likelihood across bootstrap samples for scores of 30–34 in the validation sample is too low by 21.2 percentage points. For scores of 35–39, the estimate is too high by 24.4 percentage points.[17]

For 100% of the national line, the 90-percent confidence interval for the differences for scores of 30–34 is ±11.8 percentage points (Figure 7). This means that in 900 of 1,000 bootstraps, the difference between the estimate and the true value is between –33.0 and –9.4 percentage points (because –21.2 – 11.8 = –33.0, and –21.2 + 11.8 = –9.4). In 950 of 1,000 bootstraps (95 percent), the difference is –21.2 ± 12.0 percentage points, and in 990 of 1,000 bootstraps (99 percent), the difference is –21.2 ± 12.3 percentage points.

For several scores, Figure 7 shows very large differences between estimated poverty likelihoods and true values. This is because the validation sample is a single sample that—thanks to sampling variation—differs in distribution from the construction/calibration sub-samples and from the Kyrgyz Republic'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.

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

In addition, if estimates of groups' poverty rates are to be usefully accurate, then errors for individual households' poverty likelihoods must largely balance out. As discussed in the next section, this is generally the case in samples that are representative of the population of the Kyrgyz Republic as a whole.

Another possible source of differences between estimates and true values is overfitting. Given its assumptions, the scorecard here is unbiased, but it may still be *overfit* when applied after the end of the KIHS fieldwork in December 2012. That is, it may fit the data from the 2012 KIHS 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 2012 KIHS but not in the overall population of the Kyrgyz Republic. Or the scorecard may be overfit in the sense that it is not robust when relationships between indicators and poverty change over time or when it is applied to non-nationally representative samples.

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

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

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

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

To illustrate, suppose an organization samples three households on 1 January 2015 and that they have scores of 20, 30, and 40, corresponding to poverty likelihoods of 77.3, 46.5, and 15.8 percent (100% of the national line, Figure 4). The group's estimated poverty rate is the households' average poverty likelihood of (77.3 + 46.5 + 15.8) ÷ 3 = 46.5 percent.

Be careful; the group's poverty rate is *not* the poverty likelihood associated with the average score. Here, the average score is 30, which corresponds to a poverty likelihood of 46.5 percent. For the Kyrgyz Republic, this happens—purely by chance—to be the same as the 46.5 percent found as the average of the three individual poverty likelihoods associated with each of the three scores. In general, however, the two figures are not the same. Unlike poverty likelihoods, scores are ordinal symbols, like letters in the alphabet or colors in the spectrum. Because scores are not cardinal numbers, they cannot be added up or averaged across households. Only three operations are valid for scores: conversion to poverty likelihoods, distributional analysis (Schreiner, 2012a), or comparison—if desired—with a cut-off for targeting. The safest rule to follow is: Always use poverty likelihoods, never scores.

6.1 Accuracy of estimated poverty rates at a point in time

For the Kyrgyz Republic's scorecard applied to 1,000 bootstraps of $n = 16{,}384$ from the validation sample, the average difference between the estimated poverty rate at a point in time and the true rate for 100% of the national line is –2.3 percentage points (Figure 9, summarizing Figure 8 across poverty lines). Across all eight poverty lines, the maximum absolute difference is 2.5 percentage points, and the average absolute difference is about 1.6 percentage points. At least part of these differences is due to sampling variation in the 2012 KIHS and in its division into two sub-samples.

When estimating poverty rates at a point in time, the bias reported in Figure 9 should be subtracted from the average poverty likelihood to make the estimate unbiased. For the Kyrgyz Republic's scorecard and the example of the 100% of the national line, bias is –2.3 percentage points, so the unbiased estimate in the three-household example above is 46.5 – (–2.3) = 48.8 percent.

In terms of precision, the 90-percent confidence interval for a group's estimated poverty rate at a point in time with $n = 16{,}384$ is ±0.8 percentage points or less (Figure 9). This means that in 900 of 1,000 bootstraps of this size, the average estimate (after subtracting off bias) is within 0.8 percentage points of the true value.

For example, suppose that the average poverty likelihood in a sample of $n = 16{,}384$ with the Kyrgyz Republic scorecard and 100% of the national line is 46.5 percent. Then estimates in 90 percent of such samples would be expected to fall in the range of 46.5 – (–2.3) – 0.7 = 48.1 percent to 46.5 – (–2.3) + 0.7 = 49.5 percent, with

the most likely true value being the unbiased estimate in the middle of this range (46.5 – (–2.3) = 48.8 percent). This is because the original (biased) estimate is 46.5 percent, bias is –2.3 percentage points, and the 90-percent confidence interval for 100% of the national line is ±0.7 percentage points (Figure 9).

6.2 Formula for standard errors for estimates of poverty rates

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

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

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

z is from the Normal distribution and is $\begin{cases} 1.04 \text{ for confidence levels of 70 percent} \\ 1.28 \text{ for confidence levels of 80 percent} \\ 1.64 \text{ for confidence levels of 90 percent} \end{cases}$,

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

$\hat{p}$ is the estimated proportion of households below the poverty line in the sample,

ϕ is the finite population correction factor of $\sqrt{\frac{N - n}{N - 1}}$,

N is the population size, and

n is the sample size.

For example, the direct-measure estimate of the Kyrgyz Republic's household-level poverty rate for 100% of the national line in the 2012 KIHS is $\hat{p} = 28.5$ percent

(Figure 1). If this estimate came from a sample of $n = 16{,}384$ households from a population N of 1,437,311 (an estimate of the number of households in the Kyrgyz Republic in 2012), then the finite population correction ϕ is $\sqrt{\frac{1{,}437{,}311 - 16{,}384}{1{,}437{,}311 - 1}} = 0.9943$, which can be taken as $\phi = 1$. If the desired confidence level is 90-percent ($z = 1.64$), then the confidence interval $\pm c$ is

$$\pm z \cdot \sqrt{\frac{\hat{p} \cdot (1 - \hat{p})}{n}} \cdot \sqrt{\frac{N - n}{N - 1}} = \pm 1.64 \cdot \sqrt{\frac{0.285 \cdot (1 - 0.285)}{16{,}384}} \cdot 1 = \pm 0.578 \text{ percentage points.}$$

The scorecard, however, does not measure poverty directly, so this formula is not applicable. To derive a formula for the the Kyrgyz Republic scorecard, consider Figure 8, which reports empirical confidence intervals $\pm c$ for the differences for the scorecard applied to 1,000 bootstrap samples of various sizes from the validation sample. For example, with $n = 16{,}384$ and 100% of the national line, the 90-percent confidence interval is ± 0.737 percentage points.[18]

Thus, the 90-percent confidence interval with $n = 16{,}384$ is ± 0.737 percentage points for indirect estimates via the Kyrgyz Republic's scorecard and ± 0.578 percentage points for direct measurement via the 2012 KIHS. The ratio of the two intervals is $0.737 \div 0.578 = 1.28$.

[18] Due to rounding, Figure 8 displays 0.7, not 0.737.

Now consider the same case, but with $n = 8{,}192$. The confidence interval under direct measurement is $\pm 1.64 \cdot \sqrt{\frac{0.285 \cdot (1 - 0.285)}{8{,}192}} \cdot 1 = \pm 0.818$ percentage points. The empirical confidence interval with the Kyrgyz Republic scorecard and the national line (Figure 8) is ±1.062 percentage points. Thus for $n = 8{,}192$, the ratio of the two intervals is $1.062 \div 0.818 = 1.30$.

This ratio of 1.30 for $n = 8{,}192$ is close to the ratio of 1.28 for $n = 16{,}384$. It turns out that across all sample sizes of 256 or more in Figure 8, the average ratio is 1.29, implying that confidence intervals for indirect estimates of poverty rates via the Kyrgyz Republic scorecard and 100% of the national poverty line are—for a given sample size—about 30-percent wider than confidence intervals for direct estimates via the 2012 KIHS. This 1.29 appears in Figure 9 as the "α factor for precision" because if α = 1.29, then the formula for confidence intervals c for the Kyrgyz Republic's scorecard is $\pm c = \pm z \cdot \alpha \cdot \sigma$. That is, the formula for the standard error σ for point-in-time estimates of poverty rates via scoring is $\alpha \cdot \sqrt{\frac{\hat{p} \cdot (1 - \hat{p})}{n}} \cdot \sqrt{\frac{N - n}{N - 1}}$.

In general, α can be more or less than 1.00. When α is less than 1.00, it means that the scorecard is more precise than direct measurement. This occurs for four of the eight poverty lines in Figure 9.

The formula relating confidence intervals with standard errors for the scorecard can be rearranged to give a formula for determining sample size before measurement. If $\tilde{p}$ is the expected poverty rate before measurement, then the formula for sample size n

from a population of size N that is based on the desired confidence level that corresponds to z and the desired confidence interval $\pm c$ is $n = N \cdot \left(\frac{z^2 \cdot \alpha^2 \cdot \widetilde{p} \cdot (1 - \widetilde{p})}{z^2 \cdot \alpha^2 \cdot \widetilde{p} \cdot (1 - \widetilde{p}) + c^2 \cdot (N - 1)} \right)$. If the population N is "large" relative to the sample size n, then the finite population correction factor ϕ can be taken as one (1), and the formula becomes $n = \left(\frac{\alpha \cdot z}{c} \right)^2 \cdot \widetilde{p} \cdot (1 - \widetilde{p})$.

To illustrate how to use this, suppose the population N is 1,437,311 (the number of households in the Kyrgyz Republic while the 2012 KIHS was in the field), suppose $c = 0.06114$, $z = 1.64$ (90-percent confidence), and the relevant poverty line is 100% of the national line so that the most sensible expected poverty rate $\widetilde{p}$ is the Kyrgyz Republic's overall poverty rate for that line in 2012 (28.5 percent at the household level, Figure 1). The α factor is 1.29 (Figure 9). Then the sample-size formula gives $n = 1{,}437{,}311 \cdot \left(\frac{1.64^2 \cdot 1.29^2 \cdot 0.285 \cdot (1 - 0.285)}{1.64^2 \cdot 1.29^2 \cdot 0.285 \cdot (1 - 0.285) + 0.06114^2 \cdot (1{,}437{,}311 - 1)} \right) = 244$, which is not far from the sample size of 256 observed for these parameters in Figure 8 for 100% of the national line. Taking the finite population correction factor ϕ as one (1) gives the same answer, as $n = \left(\frac{1.29 \cdot 1.64}{0.06114} \right)^2 \cdot 0.285 \cdot (1 - 0.285) = 244$.[19]

[19] Although USAID has not specified required confidence levels nor intervals, IRIS Center (2007a and 2007b) says that a sample size of $n = 300$ is sufficient for USAID reporting. USAID microenterprise partners in the Kyrgyz Republic should report using the median line. Given an α factor of 1.82 for this line (Figure 9), an expected before-measurement household-level poverty rate of 13.1 percent (the all-Kyrgyz Republic

Of course, the α factors in Figure 9 are specific to the Kyrgyz Republic, its poverty lines, and its poverty rates. The derivation of the formulas, however, is valid for any approach like the one in this paper.

In practice after the end of fieldwork for the KIHS in December 2012, a program would select a poverty line (say, 100% of the national line), note its participants' population size (say, $N = 10{,}000$ participants), select a desired confidence level (say, 90 percent, or $z = 1.64$), select a desired confidence interval (say, ±2.0 percentage points, or $c = \pm 0.02$), make an assumption about $\widetilde{p}$ (perhaps based on a previous measurement such as the household-level poverty rate for 100% of the national line for the Kyrgyz Republic of 28.5 percent in the 2012 KIHS in Figure 1), look up α (here, 1.29, Figure 9), assume that the scorecard will still work in the future and for non-nationally representative sub-groups,[20] and then compute the required sample size. In this illustration, $n = 10{,}000 \cdot \left(\frac{1.64^2 \cdot 1.29^2 \cdot 0.285 \cdot (1 - 0.285)}{1.64^2 \cdot 1.29^2 \cdot 0.285 \cdot (1 - 0.285) + 0.02^2 \cdot (10{,}000 - 1)} \right) =$ 1,857.

---

household-level rate for the median line in 2012, Figure 1), and a confidence level of 90 percent (so that $z = 1.64$), then $n = 300$ implies a confidence interval of $\pm 1.82 \cdot 1.64 \cdot \sqrt{\frac{0.131 \cdot (1 - 0.131)}{300}} = \pm 5.8$ percentage points.

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

7. Estimates of changes in poverty rates over time

The change in a group's poverty rate between two points in time is estimated as the change in the average poverty likelihood of the households in the group. With only data from the 2012 KIHS, this paper cannot test estimates of change over time for the Kyrgyz Republic, so it can only suggest approximate formulas for standard errors. Nevertheless, the relevant concepts are presented here because, in practice, pro-poor organizations in the Kyrgyz Republic can apply the scorecard to collect their own data and measure change through time.

7.1 Warning: Change is not impact

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

7.2 Calculating estimated changes in poverty rates over time

Consider the illustration begun in the previous section. On 1 January 2015, an organization samples three households who score 20, 30, and 40 and so have poverty likelihoods of 77.3, 46.5, and 15.8 percent (100% of the national line, Figure 4). Adjusting for the known bias of –2.3 percentage points (Figure 9), the group's baseline estimated poverty rate is the households' average poverty likelihood of [(77.3 + 46.5 + 15.8) ÷ 3] – (–2.3)= 48.8 percent.

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

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

By way of illustration, suppose that two years later on 1 January 2017, the organization samples three additional households who are in the same population as the three original households (or suppose that the same three original households are scored a second time) and finds that their scores are now 25, 35, and 45 (poverty likelihoods of 68.0, 40.7, and 11.7 percent, 100% of the national line, Figure 4). Adjusting for the known bias, the average poverty likelihood at follow-up is [(68.0 + 40.7 + 11.7) ÷ 3] – (–2.3) = 42.4 percent, an improvement of 48.8 – 42.4 = 6.4 percentage points.[21]

[21] Of course, such a huge reduction in poverty in two years is highly unlikely, but this is just an example to show how the scorecard can be used to estimate change.

Thus, about one in 16 participants in this hypothetical example crossed the poverty line in 2015/7.[22] Among those who started below the line, about one in eight (6.4 ÷ 48.8 = 13.1 percent) on net ended up above the line.[23]

7.3 Accuracy for estimated change in two independent samples

With only the 2012 KIHS, 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 use the Kyrgyz Republic's scorecard to estimate change. The rest of this section suggests approximate formulas for standard errors that may be used until there is additional data.

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

$$\pm c = \pm z \cdot \sigma = \pm z \cdot \alpha \cdot \sqrt{\frac{2 \cdot \hat{p} \cdot (1 - \hat{p})}{n}} \cdot \sqrt{\frac{N - n}{N - 1}} .$$

Here, z, c, $\hat{p}$ and N are defined as above, n is the sample size at both baseline and follow-up,[24] and α is the average (across a range of bootstrapped sample sizes) of

[22] This is a net figure; some people start above the line and end below it, and vice versa.
[23] The scorecard does not reveal the reasons for this change.
[24] This means that—for a given level of precision—estimating the change in a poverty rate between two points in time requires four times as many measurements (not twice as many) as does estimating a poverty rate at a point in time.

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

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

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

To illustrate how to determine sample size for estimating changes in poverty rates across two independent samples, suppose the desired confidence level is 90 percent ($z = 1.64$), the desired confidence interval is ± 2 percentage points ($\pm c = \pm 0.02$), the poverty line is 100% of the national line, $\alpha = 1.15$, $\hat{p} = 0.285$ (the household-level poverty rate in 2012 for 100% of the national line in Figure 1), and the population N is large enough relative to the expected sample size n that the finite population correction ϕ can be taken as one. Then the baseline sample size is

$n = 2 \cdot \left(\frac{1.15 \cdot 1.64}{0.02} \right)^2 \cdot 0.285 \cdot (1 - 0.285) \cdot 1 = 3{,}625$, and the follow-up sample size is also 3,625.

7.4 Precision for estimated change for one sample, scored twice

Analogous to previous derivations, the general formula relating the confidence interval $\pm c$ to the standard error σ when using a scorecard to estimate change for a single group of households, all of whom are scored at two points in time, is:[25]

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

where z, c, α, N, and n are defined as usual, $\hat{p}_{12}$ is the share of all sampled households that move from below the poverty line to above it, and $\hat{p}_{21}$ is the share of all sampled households that move from above the line to below it.

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

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

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

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

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

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

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

In Peru (the only source of a data-based estimate, Schreiner, 2009e), the average α across years and poverty lines is about 1.30.

To illustrate the use of this formula, suppose the desired confidence level is 90 percent ($z = 1.64$), the desired confidence interval is ±2.0 percentage points ($\pm c = \pm 0.02$), the poverty line is 100% of the national line, the sample will first be scored in 2015 and then again in 2018 ($y = 3$), and the population N is so large relative to the

expected sample size n that the finite population correction ϕ can be taken as one. The pre-baseline poverty rate p_{2012} is taken as 28.5 percent (Figure 1), and α is assumed to be 1.30. Then the baseline sample size is

$n = 2 \cdot \left(\frac{1.30 \cdot 1.64}{0.02}\right)^2 \cdot \{-0.02 + 0.016 \cdot 3 + 0.47 \cdot [0.285 \cdot (1 - 0.285)]\} \cdot 1 = 2{,}814$. The same group of 2,814 households is scored at follow-up as well.

8. Targeting

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

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

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

Figure 10 depicts these four possible targeting outcomes. Targeting accuracy varies by the cut-off score; a higher cut-off has better inclusion (but worse leakage), while a lower cut-off has better exclusion (but worse undercoverage).

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

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

Figure 11 shows the distribution of households by targeting outcome for the Kyrgyz Republic. For an example cut-off of 34 or less, outcomes for 100% of the national line in the validation sample are:

- Inclusion: 19.5 percent are below the line and correctly targeted
- Undercoverage: 9.0 percent are below the line and mistakenly not targeted
- Leakage: 11.2 percent are above the line and mistakenly targeted
- Exclusion: 60.3 percent are above the line and correctly not targeted

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

- Inclusion: 22.4 percent are below the line and correctly targeted
- Undercoverage: 6.0 percent are below the line and mistakenly not targeted
- Leakage: 18.7 percent are above the line and mistakenly targeted
- Exclusion: 52.8 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. A program that uses targeting—with or without scoring—should thoughtfully consider

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

A common choice of benefits and costs is the "hit rate", where total net benefit is the number of households correctly included or correctly excluded:

Hit rate =	1	x	Households correctly included	–
	0	x	Households mistakenly undercovered	–
	0	x	Households mistakenly leaked	+
	1	x	Households correctly excluded.	

Figure 11 shows the hit rate for all cut-offs for the Kyrgyz Republic scorecard. For 100% of the national line in the validation sample, total net benefit is greatest (79.8) for a cut-off of 34 or less, with four in five households in the Kyrgyz Republic correctly classified.

The hit rate weighs successful inclusion of households below the line the same as successful exclusion of households above the line. If a program values inclusion more (say, twice as much) than exclusion, then it can reflect this by setting the benefit for inclusion to 2 and the benefit for exclusion to 1. Then the chosen cut-off will maximize (2 x Households correctly included) + (1 x Households correctly excluded).[26]

[26] Figure 11 also reports BPAC, the Balanced Poverty Accuracy Criteria adopted by USAID for certifying poverty-assessment tools. IRIS Center (2005) made BPAC to consider accuracy in terms of the bias of estimated poverty rates and in terms of targeting inclusion. BPAC = (Inclusion – |Undercoverage – Leakage|) x [100 ÷ (Inclusion + Undercoverage)]. Schreiner (2014) explains why BPAC does not add any useful information over-and-above that provided by the other, more-standard measures here.

As an alternative to assigning benefits and costs to targeting outcomes and then choosing a cut-off to maximize total net benefits, a program could set a cut-off to achieve a desired poverty rate among targeted households. The third column of Figure 12 ("% targeted HHs who are poor") shows, for the Kyrgyz Republic scorecard applied to the validation sample, the expected poverty rate among households who score at or below a given cut-off. For the example of 100% of the national line, targeting households in the validation sample who score 34 or less would target 30.7 percent of all households (second column) and be associated with a poverty rate among those targeted of 63.4 percent (third column).

Figure 12 also reports two other measures of targeting accuracy. The first is a version of coverage ("% poor HHs who are targeted"). For the example of 100% of the national line with the validation sample and a cut-off of 34 or less, 68.4 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 100% of the national line with the validation sample and a cut-off of 34 or less, covering 1.7 poor households means leaking to 1 non-poor household.

9. Context of poverty-assessment tools in the Kyrgyz Republic

This section discusses three existing poverty-assessment tools for the Kyrgyz Republic in terms of their goals, methods, definitions of *poverty*, data, indicators, cost, bias, and precision. In general, the advantages of the scorecard here are its:

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

9.1 Gwatkin *et al.*

Gwatkin *et al.* (2007) construct a poverty-assessment tool for the Kyrgyz Republic with an approach that they use in 56 countries with Demographic and Health Surveys (Rutstein and Johnson, 2004). They use Principal Components Analysis to make an asset index from simple, low-cost indicators available for the 3,672 households in the Kyrgyz Republic's 1997 DHS.[27] The PCA index is like the scorecard except that, because the DHS does not collect data on consumption, the index is based on a different conception of *poverty*, its accuracy vis-à-vis consumption-based poverty is unknown, and it can only be assumed to be a proxy for long-term wealth/economic status.[28] Well-

[27] All DHS datasets for the Kyrgyz Republic since 1997 include households' scores on an asset index (`dhsprogram.com/topics/wealth-index/`, retrieved 12 February 2015).
[28] Nevertheless, the indicators are similar and the "flat maximum" is important, so carefully built PCA indexes and consumption-based poverty-assessment tools may pick up the same underlying construct (perhaps "permanent income", see Bollen, Glanville,

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

The 13 indicators in Gwatkin *et al.* are similar to those in the scorecard here in terms of their simplicity, low cost, and verifiability:

- Characteristics of the residence:
  - Presence of electricity
  - Type of floor
  - Source of drinking water
  - Type of toilet arrangement
- Ownership of consumer durables:
  - Radios
  - Televisions
  - Telephones
  - Refrigerators
  - Bicycles
  - Motorcycles
  - Cars
- Number of household members per sleeping room
- Whether members of the household work their own or family's agricultural land

---

and Stecklov, 2007), and they may rank households much the same. Comparisons of rankings by PCA indexes and consumption-based poverty-assessment tools include Filmer and Scott (2012), Lindelow (2006), Sahn and Stifel (2003), Wagstaff and Watanabe (2003), and Montgomery *et al.* (2000).

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

- Segmenting households by the quintile of their index to see how health varies with socio-economic status
- Monitoring (via exit surveys) how well local health-service posts reach the poor
- Measuring local coverage of health services via small-scale surveys

The first goal is akin to targeting, and the last two goals deal with performance monitoring, so the asset index would be used much like the scorecard.

Still, the Gwatkin *et al.* index is more difficult and costly than the scorecard. While the scorecard requires adding up 10 integers (some of which are often zeroes), Gwatkin *et al.*'s asset index requires adding up 55 numbers, each with five decimal places and half with negative signs.

Unlike the asset index, the scorecard here is linked directly to a consumption-based poverty line. Thus, while both approaches can rank households, only the scorecard estimates consumption-based poverty likelihood.

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

The asset-index approach defines people as *poor* if their assets (physical, human, financial, and social) fall below a threshold. Arguments for an asset-based view of development include Carter and Barrett (2006), Schreiner and Sherraden (2006), Sahn and Stifel (2003), and Sherraden (1991). The main advantages of the asset-based view are that:

- Asset ownership is easier to measure accurately than consumption
- Access to resources in the long term—and thus capacity to produce income and to consume—depends on the control of assets
- Assets reflect realizations rather than capacity, the difference between, say, "Does the toilet drain to a septic tank?" versus "Does income permit adequate sanitation?"

While the asset view and the income/consumption view are distinct, they are also tightly linked. After all, income and consumption are flows of resources received/consumed from the use of stocks of assets. Both views are low-dimensional simplifications—due to practical limits on definitions and measurement—of a higher-dimensional and more complete conception of the production of human well-being.

9.2 Mahadevan, Yoshida, and Praslova

Mahadevan, Yoshida, and Praslova (2013) use "poverty mapping" (Elbers, Lanjouw, and Lanjouw, 2003) to estimate poverty rates for the Kyrgyz Republic's eight oblasts and 56 rayons. The goal is to "help policy makers provide resources to the areas that need it the most in an objective and transparent manner" (p. 3).

To this end, Mahadevan, Yoshida, and Praslova construct five poverty-assessment tools with stepwise least-squares regression of the logarithm of per-capita

consumption against indicators found both in the 2009 KIHS and in the 2009 Census. Some indicators are village-level census means, as the distributions of household-level indicators in the 2009 KIHS often do not match the distributions of the corresponding indicators in the 2009 Census. One tool is constructed for each of the four oblasts of Issykul, Jalal-Abad, Chui, and Bishkek, and a fifth tool is made for the four oblasts of Naryn, Batken, Osh, and Talas.

The five tools are applied to the census data with a poverty line[29] to estimate poverty rates for smaller areas (rayons) than would be possible with only the 2009 KIHS. Finally, the results are displayed via "poverty maps" that quickly show how estimated poverty rates vary across rayons in a way that makes sense to lay people.

Poverty mapping in Mahadevan, Yoshida, and Praslova has much in common with the scorecard here in that they both:

- Build poverty-assessment tools with data that is representative of a given population (all-Kyrgyz Republic for the scorecard, and oblasts for Mahadevan, Yoshida, and Praslova) and then apply them to other data on groups that are not, in general, representative of the same populations
- Estimate poverty rates for groups
- Test accuracy empirically
- Report bias and standard errors
- Seek to be useful in practice and so aim to be understood by non-specialists

[29] Mahadevan, Yoshida, and Praslova do not report the poverty line nor its associated all-Kyrgyz Republic poverty rate, but they probably use the national line of KGS53.20 per person per day in 2009 prices (Kyrgyz Republic, 2011), giving a person-level poverty rate of 31.7 percent (NSC, 2013).

Strengths of poverty mapping include that it:

- Has formally established theoretical properties
- Can be applied straightforwardly to distributional measures of well-being (such as the poverty gap or the Gini coefficient) that go beyond head-count poverty rates
- Accounts for uncertainty in the estimation of a tool's points when estimating standard errors
- Requires data on fewer households for construction and calibration
- Includes village-level indicators, increasing accuracy and precision
- Uses only indicators that appear in a census

Strengths of the scorecard include that it:

- Uses simple, verifiable indicators that are quick and inexpensive to collect
- Is simpler in terms of both construction and application
- Links poverty likelihoods with scores non-parametrically
- Surfaces estimates of poverty likelihoods for individual households
- Provides unbiased estimates when its assumptions hold
- Reduces overfitting by selecting indicators with statistical and non-statistical criteria
- Reports confidence intervals and simple formulas for standard errors
- Aims to be transparent to non-specialists

The basic difference between the two approaches is that poverty mapping seeks to help governments to target pro-poor policies, while the scorecard seeks to help local pro-poor organizations to manage their social performance.[30] On a technical level,

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

Mahadevan, Yoshida, and Praslova estimate consumption, whereas the scorecard estimates poverty likelihoods.[31]

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

The five tools in Mahadevan, Yoshida, and Praslova use an average of about 11 indicators from among the following 23:

- Household-level indicators:
  - Demographics:
    - Number of members
    - Share of members who are children
    - Share of members who are 60-years-old or older
    - Age (presumably of head)
    - Sex (presumably of head)
  - Share of household members with a given level of education
  - Employment: Share of household members who are:
    - Employed
    - Unemployed
    - Of working age
    - Employed in a given sector
  - Characteristics of the residence:
    - Type
    - Area
    - Type of toilet arrangement
  - Location:
    - Rayon
    - Interaction of a given rayon with:
      - Number of household members
      - Education level (presumably of head)
      - Occupation (presumably of head)
      - Type of toilet arrangement
- Village-level census means:
  - Household size
  - Share of household members with a given education level
  - Share of household members employed in a given sector
  - Sex (presumably of head)
  - Employment status (presumably of head)

The poverty map of Mahadevan, Yoshida, and Praslova is not designed for field use by local, pro-poor organizations. Having five tools complicates administration if an organization works in multiple oblasts. Also, field workers cannot compute scores, and

an organization's back-office must match up a household and its location with average census values for its village.

Although Mahadevan, Yoshida, and Praslova use five tools, they note (pp. 6–7) that "the latest recommendation from the World Bank Research Department is not to use multiple [poverty-assessment tools] to predict household consumption" as multiple tools can be "problematic since the number of observations for each area becomes small and, as a result, the regression coefficients become less stable."[32]

Mahadevan, Yoshida, and Praslova do not use a single, all-Kyrgyz Republic tool because—compared with a single tool—the five-tool approach leads to estimates of oblast-level poverty rates that are closer to those derived from the direct measures of consumption in the 2009 KIHS.

It is not possible to compare accuracy across poverty-rate estimates for the poverty map versus the scorecard. While Mahadevan, Yoshida, and Praslova report their estimates for the Kyrgyz Republic's eight oblasts and 46 rayons, they do not report the corresponding 2009 KIHS estimates (and those do not seem to be available from other sources). And of course, the poverty map uses the 2009 KIHS, while the scorecard here uses the 2012 KIHS.

[32] This is consistent with Haslett's (2012) recommendation of a single, all-country scorecard for poverty mapping to avoid overfitting.

9.3 Gassmann

Gassmann (2013) uses stepwise least-squares regression to make two poverty-assessment tools (urban and rural) relating the logarithm of per-capita consumption to indicators from the 2010 KIHS. The goal is "to analyze the potential of improving targeting performance . . . of [the Kyrgyz Republic's] Monthly Benefit for Poor Families" (p. 3). The MBPF gives cash transfers to 11 percent of people in the Kyrgyz Republic so as to bring their income up to a minimum standard. Only households with children are eligible. Under the current income-based targeting system, 30 percent of people in the poorest quintile receive benefits (compared with a possible 55 percent).

Gassmann compares her tool's targeting against the MBPF's current system and against a 1-indicator tool based on the number of children under 16-years-old. She finds (p. 4) that for a given budget, the "[tools] would achieve better outcomes in terms of targeting performance."

Gassmann's two poverty-assessment tools each use about 20 of the following 30 indicators:

- Logarithm of per-capita income
- Demographics:
  - Number of household members:
    - Total (linear and non-linear)
    - Pensioners
    - Old-age pensioners
  - Characteristics of the head of the household:
    - Sex
    - Marital status
- Education: Number of members who have completed:
  - Less than secondary school
  - Secondary professional education
  - Higher education
- Employment:
  - Whether the head is unemployed
  - Number of household members who are self-employed
- Caracteristics of the residence:
  - Oblast
  - Rooms per person
  - Type of walls
  - Type of cooking fuel
  - Source of drinking water
  - Type of toilet arrangement
  - Type of heating arrangement
  - Presence of garage
  - Ownership of second house
- Ownership of consumer durables:
  - Washing machines
  - Mobile telephones
  - Satellite dishes
  - Computers
  - Cars

- Agriculture:
  - Area of land owned
  - Number of livestock owned:
    - Poultry
    - Goats
    - Cows
    - Horses

Gassman's tools are meant for use by the government of the Kyrgyz Republic. They require an estimate of a household's income, so while their application might be feasible with motivated respondents seeking to qualify for cash assistance from the MBFP, they would be difficult for local, pro-poor organizations.

How does targeting accuracy compare between Gassman and the scorecard? The comparison is not apples-to-apples for three reasons. First, Gassmann uses 2010 data, while this paper uses 2012. Second, Gassmann constructs and tests her tools with person-level weights, while the scorecard is constructed with household-level weights. Third, Gassmann tests *in-sample* (with data that was also used in construction), but the scorecard tests *out-of-sample* (with data that was not used in construction). In-sample tests tend to overstate accuracy.

Gassmann presents in-sample accuracy results for several scenarios, but the best one for a comparison here is that which targets the lowest-scoring 20 percent of people. With a person-level poverty rate of 20 percent, Gassmann's tools have inclusion of 13.3 percent, exclusion of 73.4 percent, and a hit rate of 86.7 percent.

When the scorecard is re-constructed using the full 2012 KIHS and then tested by targeting the lowest-scoring 20 percent of people in-sample with person-level weights and a poverty line set to give a poverty rate of 20 percent, inclusion is 10.2 percent, exclusion is 70.4 percent, and the hit rate is 80.6 percent.

Thus, Gassmann's tools target about six more people correctly per 100 than the scorecard. This is likely due to its greater number of indicators and, especially, the use of an indicator for income.

10. Conclusion

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

The scorecard is constructed with half of the data from the Kyrgyz Republic's 2012 KIHS, calibrated to eight poverty lines, and tested on the other half of the 2012 data.

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

When the scorecard is applied to the validation sample, the maximum absolute difference between estimates versus true poverty rates for groups of households at a point in time across the eight poverty lines is 2.5 percentage points. The average absolute difference is about 1.6 percentage points. Unbiased estimates may be had by subtracting the known bias from the original estimates. For $n = 16{,}384$ and 90-percent confidence, the precision of these differences is ± 0.8 percentage points or better.

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

Although the statistical technique is innovative, and although technical accuracy is important, the design of the scorecard focuses on transparency and ease-of-use. After all, accuracy is irrelevant if an organization feels so daunted by a scorecard's complexity or its cost that it does not even try to use it. For this reason, the scorecard is kept simple, using ten indicators that are straightforward, low-cost, and verifiable. Points are all zeros or positive integers, and scores range from 0 (most likely below a poverty line) to 100 (least likely below a poverty line). Scores are converted to poverty likelihoods via simple look-up tables, and targeting cut-offs are likewise straightforward to apply. The design attempts to facilitate adoption by helping managers understand and trust scoring and by allowing non-specialists to generate scores quickly in the field.

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

References

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

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

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

Caire, Dean. (2004) "Building Credit Scorecards for Small-Business Lending in Developing Markets", `microfinance.com/English/Papers/Scoring_SMEs_Hybrid.pdf`, retrieved 12 February 2015.

_____; and Mark Schreiner. (2012) "Cross-Tab Weighting for Credit Scorecards in Developing Markets", `dean_caire@hotmail.com`.

Camacho, Adriana; and Emily Conover. (2011) "Manipulation of Social-Program Eligibility", *American Economic Journal: Economic Policy*, Vol. 3, No. 2, pp. 41–65.

Carter, Michael R.; and Christopher B. Barrett. (2006) "The Economics of Poverty Traps and Persistent Poverty: An Asset-Based Approach", *Journal of Development Studies,* Vol. 42, No. 2, pp. 178–199.

Chen, Shiyuan; and Mark Schreiner. (2009) "Simple Poverty Scorecard Poverty-Assessment Tool: Vietnam", `SimplePovertyScorecard.com/VNM_2006_ENG.pdf`, retrieved 12 February 2015.

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

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

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

Deaton, Angus; and Salman Zaidi. (2002) "Guidelines for Constructing Consumption Aggregates for Welfare Analysis", World Bank LSMS Working Paper No. 135, go.worldbank.org/8YRCR9ERJ0, retrieved 2 February 2015.

Demombynes, Gabriel; Elbers, Chris; Lanjouw, Jean O.; Lanjouw, Peter; Mistiaen, Johan; and Berk Özler. (2004) "Producing an Improved Geographic Profile of Poverty: Methodology and Evidence from Three Developing Countries", pp. 154–176 in Anthony Shorrocks and Rolph van der Hoeven (eds.) *Growth, Inequality, and Poverty*.

Diamond, Alexis; Gill, Michael; Rebolledo Dellepiane, Miguel Angel; Skoufias, Emmanuel; Vinha, Katja; and Yiqing Xu. (2016) "Estimating Poverty Rates in Target Populations: An Assessment of the Simple Poverty Scorecard and Alternative Approaches", World Bank Policy Research Working Paper No. 7793, hdl.handle.net/10986/25038, retrieved 11 January 2017.

Elbers, Chris; Fujii, Tomoki; Lanjouw, Peter; Özler, Berk; and Wesley Yin. (2007) "Poverty Alleviation through Geographic Targeting: How Much Does Disaggregation Help?", *Journal of Development Economics*, Vol. 83, pp. 198–213.

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

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

_____; and Kinnon Scott. (2012) "Assessing Asset Indices", *Demography*, Vol. 49, pp. 359–392.

Friedman, Jerome H. (1997) "On Bias, Variance, 0–1 Loss, and the Curse-of-Dimensionality", *Data Mining and Knowledge Discovery*, Vol. 1, pp. 55–77.

Fuller, Rob. (2006) "Measuring the Poverty of Microfinance Clients in Haiti", microfinance.com/English/Papers/Scoring_Poverty_Haiti_Fuller.pdf, retrieved 12 February 2015.

Gassmann, Franziska. (2013) "Kyrgyz Republic: Minimum Living Standards and Alternative Targeting Methods for Social Transfers", World Bank Report No. 78168, documents.worldbank.org/curated/en/2013/06/17796512/kyrgyz-republic-minimum-living-standards-alternative-targeting-methods-social-transfers-policy-note, retrieved 12 February 2015.

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

Grosh, Margaret; and Judy L. Baker. (1995) "Proxy-Means Tests for Targeting Social Programs: Simulations and Speculation", World Bank LSMS Working Paper No. 118, `go.worldbank.org/W9OWN57PD0`, retrieved 12 February 2015.

Gwatkin, Davidson R.; Rutstein, Shea; Johnson, Kiersten; Suliman, Eldaw; Wagstaff, Adam; and Agbessi Amouzou. (2007) "Socio-Economic Differences in Health, Nutrition, and Population: Kyrgyz Republic", `go.worldbank.org/T6LCN5A340`, retrieved 12 February 2015.

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

Haslett, Stephen. (2012) "Practical Guidelines for the Design and Analysis of Sample Surveys for Small-Area Estimation", *Journal of the Indian Society of Agricultural Statistics*, Vol. 66, No. 1, pp. 203–212.

_____; and Geoffrey Jones. (2006) "Small-Area Estimation of Poverty, Caloric Intake, and Malnutrition in Nepal", `un.org.np/node/10501`, retrieved 12 February 2015.

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

IRIS Center. (2007a) "Manual for the Implementation of USAID Poverty Assessment Tools", `povertytools.org/training_documents/Manuals/USAID_PAT_Manual_Eng.pdf`, retrieved 12 February 2015.

_____. (2007b) "Introduction to Sampling for the Implementation of PATs", `povertytools.org/training_documents/Sampling/Introduction_Sampling.ppt`, retrieved 12 February 2015.

_____. (2005) "Notes on Assessment and Improvement of Tool Accuracy", `povertytools.org/other_documents/AssessingImproving_Accuracy.pdf`, retrieved 12 February 2015.

Johnson, Glenn. (2007) "Lesson 3: Two-Way Tables—Dependent Samples", `onlinecourses.science.psu.edu/stat504/node/96`, retrieved 12 February 2015.

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

Kyrgyz Republic (2011) "On the Methodology for Determining the Poverty Line", Decision 25.03.2011g, Number 115.

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

Lovie, Alexander D.; and Patricia Lovie. (1986) "The Flat-Maximum Effect and Linear Scoring Models for Prediction", *Journal of Forecasting*, Vol. 5, No. 3, pp. 159–168.

Mahadevan, Meera; Yoshida, Nobou; and Larisa Praslova. (2013) "Poverty Mapping in the Kyrgyz Republic: Methodology and Key Findings", World Bank Report No. 76690, `documents.worldbank.org/curated/en/2013/04/17584758/kyrgyz-republic-poverty-mapping-methodology-key-findings`, retrieved 12 February 2015.

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

Matul, Michal; and Sean Kline. (2003) "Scoring Change: Prizma's Approach to Assessing Poverty", Microfinance Centre for Central and Eastern Europe and the New Independent States Spotlight Note No. 4, `mfc.org.pl/sites/mfc.org.pl/files/spotlight4.PDF`, retrieved 12 February 2015.

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

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

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

Narayan, Ambar; and Nobuo Yoshida. (2005) "Proxy-Means Tests for Targeting Welfare Benefits in Sri Lanka", World Bank Report No. SASPR–7, documents.worldbank.org/curated/en/2005/07/6209268/proxy-means-test-targeting-welfare-benefits-sri-lanka, retrieved 12 February 2015.

National Statistics Committee. (2013) *Living Standards in the Kyrgyz Republic, 2008 to 2012*.

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

Rutstein, Shea Oscar; and Kiersten Johnson. (2004) "The DHS Wealth Index", DHS Comparative Reports No. 6, measuredhs.com/pubs/pdf/CR6/CR6.pdf, retrieved 12 February 2015.

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

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

SAS Institute Inc. (2004) "The LOGISTIC Procedure: Rank Correlation of Observed Responses and Predicted Probabilities", in *SAS/STAT User's Guide, Version 9*, support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_logistic_sect035.htm, retrieved 12 February 2015.

Schreiner, Mark. (2014) "How Do the Simple Poverty Scorecard and the PAT Differ?", microfinance.com/English/Papers/Scorecard_versus_PAT.pdf, retrieved 12 February 2015.

_____. (2013a) "Simple Poverty Scorecard Poverty-Assessment Tool: Bangladesh", SimplePovertyScorecard.com/BGD_2010_ENG.pdf, retrieved 12 February 2014.

_____. (2013b) "Simple Poverty Scorecard Poverty-Assessment Tool: Nicaragua", SimplePovertyScorecard.com/NIC_2009_ENG.pdf, retrieved 12 February 2015.

_____. (2012a) "An Expert-Based Poverty Scorecard for Rural China", microfinance.com/English/Papers/Scoring_Poverty_China_EN.pdf, retrieved 12 February 2015.

_____. (2012b) "Simple Poverty Scorecard Poverty-Assessment Tool: Colombia", SimplePovertyScorecard.com/COL_2009_ENG.pdf, retrieved 12 February 2015.

_____. (2012c) "Simple Poverty Scorecard Poverty-Assessment Tool: Peru", SimplePovertyScorecard.com/PER_2010_ENG.pdf, retrieved 12 February 2015.

_____. (2010) "Simple Poverty Scorecard Poverty-Assessment Tool: Honduras", SimplePovertyScorecard.com/HND_2007_ENG.pdf, retrieved 12 Feburary 2015.

_____. (2009a) "Simple Poverty Scorecard Poverty-Assessment Tool: Philippines", SimplePovertyScorecard.com/PHL_2002_ENG.pdf, retrieved 12 February 2015.

_____. (2009b) "Simple Poverty Scorecard Poverty-Assessment Tool: Pakistan", SimplePovertyScorecard.com/PAK_2005_ENG.pdf, retrieved 12 February 2015.

_____. (2009c) "Simple Poverty Scorecard Poverty-Assessment Tool: Bolivia", SimplePovertyScorecard.com/BOL_2007_ENG.pdf, retrieved 12 Feburary 2015.

_____. (2009d) "Simple Poverty Scorecard Poverty-Assessment Tool: Mexico", SimplePovertyScorecard.com/MEX_2008_ENG.pdf, retrieved 12 Feburary 2015.

_____. (2009e) "Simple Poverty Scorecard Poverty-Assessment Tool: Peru", SimplePovertyScorecard.com/PER_2007_ENG.pdf, retrieved 12 February 2015.

_____. (2008a) "Simple Poverty Scorecard Poverty-Assessment Tool: Peru", SimplePovertyScorecard.com/PER_2003_ENG.pdf, retrieved 12 February 2015.

_____. (2008b) "Simple Poverty Scorecard Poverty-Assessment Tool: Ecuador", SimplePovertyScorecard.com/ECU_2005_ENG.pdf, retrieved 12 February 2015.

_____. (2006) "Is One Simple Poverty Scorecard Poverty-Assessment Tool Enough for India?", microfinance.com/English/Papers/coring_Poverty_India_Segments.pdf, retrieved 12 Feburary 2015.

_____. (2005a) "La Herramienta del Índice de Calificacíon de la Pobreza™: México", SimplePovertyScorecard.com/MEX_2002_SPA.pdf, retrieved 12 February 2015.

_____. (2005b) "IRIS Questions on the Simple Poverty Scorecard Poverty-Assessment Tool ", microfinance.com/English/Papers/Scoring_Poverty_Response_to_IRIS.pdf, retrieved 12 February 2015.

_____. (2002) *Scoring: The Next Breakthrough in Microfinance?* CGAP Occasional Paper No. 7, microfinance.com/English/Papers/Scoring_Breakthrough_CGAP.pdf, retrieved 12 Feburary 2015.

_____; Matul, Michal; Pawlak, Ewa; and Sean Kline. (2014) "Poverty Scoring: Lessons from a Microlender in Bosnia-Herzegovina", *Poverty and Public Policy*, Vol. 6, No. 4, pp. 407–428.

_____; and Michael Sherraden. (2006) *Can the Poor Save? Saving and Asset Accumulation in Individual Development Accounts.*

_____; and Gary Woller. (2010a) "Simple Poverty Scorecard Poverty-Assessment Tool: Ghana", SimplePovertyScorecard.com/GHA_2005_ENG.pdf, retrieved 12 February 2015.

_____; and Gary Woller. (2010b) "Simple Poverty Scorecard Poverty-Assessment Tool: Guatemala", SimplePovertyScorecard.com/GTM_2006_ENG.pdf, retrieved 12 February 2015.

Sharif, Iffath Anwar. (2009) "Building a Targeting System for Bangladesh Based on Proxy-Means Testing", World Bank Social Protection Discussion Paper No. 0914, siteresources.worldbank.org/SOCIALPROTECTION/Resources/SP-Discussion-papers/Safety-Nets-DP/0914.pdf, retrieved 12 February 2015.

Sherraden, Michael. (1991) *Assets and the Poor: A New American Welfare Policy.*

Sillers, Don. (2006) "National and International Poverty Lines: An Overview", pdf.usaid.gov/pdf_docs/Pnadh069.pdf, retrieved 12 February 2015.

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

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

Tarozzi, Alessandro; and Angus Deaton. (2009) "Using Census and Survey Data to Estimate Poverty and Inequality for Small Areas", *Review of Economics and Statistics*, Vol. 91, No. 4, pp. 773–792.

Toohig, Jeff. (2008) "PPI Pilot Training Guide", microfinancegateway.org/sites/default/files/mfg-en-paper-progress-out-of-poverty-index-ppi-pilot-training-mar-2008.pdf, retrieved 12 February 2015.

Tsirunyan, Sasun. (2012) "Methodology of Calculation of Poverty Line Based on Kyrgyz Integrated Household Surveys (KIHS) 2011".

USAID (2013) *Microenterprise Results Reporting, with Methodology and Statistical Annexes, Fiscal Year 2012*, pdf.usaid.gov/pdf_docs/pdacx521.pdf, retrieved 12 February 2015.

United States Congress. (2004) "Microenterprise Results and Accountability Act of 2004 (HR 3818 RDS)", November 20, smith4nj.com/laws/108-484.pdf, retrieved 12 February 2015.

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

Wainer, Howard. (1978) "On the Sensitivity of Regression and Regressors", *Psychological Bulletin*, Vol. 85, pp. 267–73.

World Bank. (2013) *The Kyrgyz Republic: Poverty Update, 2011*, Report No. 78212–KG, documents.worldbank.org/curated/en/2013/06/17924482/kyrgyz-republic-poverty-update-2011, retrieved 12 February 2015.

_____. (2012) *Targeting Poor and Vulnerable Households in Indonesia*, documents.worldbank.org/curated/en/2012/01/15879773/targeting-poor-vulnerable-households-indonesia, retrieved 12 February 2015.

_____. (2008) "International Comparison Project: Tables of Results", siteresources.worldbank.org/ICPINT/Resources/icp-final-tables.pdf, retrieved 12 February 2015.

Zeller, Manfred. (2004) "Review of Poverty Assessment Tools", pdf.usaid.gov/pdf_docs/PNADH120.pdf, retrieved 12 February 2015.

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

Guidelines for the Interpretation of Scorecard Indicators

The following is based on a translation from:

National Statistics Committee. (2010) *Enumerator Manual*, Bishkek [the *Manual*]

Some guidelines are gleaned from the 2012 KIHS survey instrument.

When an issue arises that is not addressed here, its resolution should be left to the unaided judgment of the enumerator, since that apparently is what was done by the NSC in the Kyrgyz Republic's 2012 KIHS. That is, an organization using the scorecard should not promulgate a standardized response to an issue to be used by all its field agents. Anything not explicitly addressed in the guidelines here is to be left to the unaided judgment of the individual enumerator.

In general, the enumerator should accept the responses given by the respondent. Nevertheless, if the respondent says something—or the enumerator sees or senses something—that suggests that the response may not reflect reality or that the respondent is uncertain about his/her response or that the respondent desires assistance in figuring out how to respond, then the enumerator should read the question again and provide whatever assistance he/she deems appropriate based on these "Guidelines". The enumerator does not need to verify responses unless something suggests to the enumerator that the response may not reflect reality.

In general, the application of scorecard should mimic as closely as possible the application of the KIHS 2012. For example, the poverty-scoring interview should be conducted in the respondent's homestead because the 2012 KIHS was conducted in respondents' homesteads.

General Guidelines

Do not read the response options to the respondent. Unless instructed otherwise here, read the question, and then stop; wait for a response. If the respondent asks for clarification or otherwise hesitates or seems confused, then read the question again or provide additional assistance based on these guidelines or as you, the enumerator, deem appropriate.

Guidelines for Scorecard Indicators

1. In which oblast does the household reside?
    A. Jalal-Abad
    B. Naryn
    C. Osh
    D. Bishkek
    E. Issykul
    F. Talas
    G. Chui
    H. Batken

If the enumerator knows the oblast of residence with complete certainty, then the response to this indicator can be recorded without asking the question of the respondent. If there is any doubt at all, then the question should be asked of the respondent.

No additional information is available for this indicator.

2. How many household members are there?
    A. Seven or more
    B. Six
    C. Five
    D. Four
    E. Three
    F. Two
    G. One (stop interview, score is 100)

The response to this question is based on the responses of the household to the battery of questions in the "Back-page Worksheet". In particular, do *not* directly ask the respondent "How many household members are there?" Instead, count the number of household members listed on the "Back-page Worksheet".

In addition, record the *total* number of household members in the box at the top of the scorecard sheet labelled "# Household members:". It is not sufficient to merely mark the response to the second indicator because households with seven or more members all fall into a single response, "Seven or more". The precise number of total household members is needed to enable conversion from household-level weights to person-level weights, should that be necessary, as well as for other possible analyses.

This is a *gateway* question in that the interview may be stopped with this question if the household has only one member. All one-member households are assigned a score of 100, and their poverty likelihood for a given poverty line is the poverty likelihood associated with a score of 100.

Of course, an organization may choose to instruct its enumerators to go ahead and record responses for all the other scorecard items even for one-person households. Asking all the questions will allow greater breadth of possible analysis. In any case, the score for a one-person household will always be 100, regardless of the responses to the other scorecard indicators.

According to the *Enumerator Manual*, a *household* is one or more individuals—with or without blood or marital ties—who share a residence and who together contribute resources to provide for the daily needs of the household members.

A household may consist of a single person who lives in his/her separate residence and who is solely responsible for meeting his/her own daily needs.

Members of a household include:

- All people who live together and meet their daily needs by pooling their resources
- The head of the household
- Newborn babies of members of the household (even though the babies do not contribute to meeting the household's needs)
- Spouses of household members
- All other people—regardless of whether they have a blood or marital relationship with the head of the household—who live together with the household and who help to meet the daily needs of the members of the household
- Persons who usually reside with the household, but who are, at the time of the interview, temporarily absent (for example, working elsewhere, recovering in a hospital, attending school, serving in the military, or being held temporarily in jail)
- Students who do not reside with the household but who depend on it to meet their needs and who visit the household at least once every three months

The following are not to be counted as *members of the household*:

- Nannies or other hired caregivers
- Anyone with only tangential, distant involvement the household's day-to-day life or who is permanently absent
- Servants, guests, friends, other relatives, or lodgers (that is, persons who rent quarters from the household and perhaps also pay the household for food)
- All other persons who are not kin of the head of the household and who do not contribute resources to meet the daily needs of the household members, even if they do share a residence with the household, are not household members. For example, a tenant who lives with a landlord but who makes his/her own decisions and who is responsible for meeting his/her own daily needs is not to be counted as a *member of the household*

If a given case is ambiguous, unclear, or confusing, then the interviewer should ask, "Does <name> eat together with the household and contribute to its functioning?"

3. In the past 7 days, how many household members worked or had paid employment for at least 1 hour, or worked on a family farm or enterprise, or (if they did not work in the past 7 days) had work or paid employment to which they plan to return?
    A. None, or one
    B. Two
    C. Three or more

According to the "Employment and Labor" module of the 2012 KIHS, this question concerns only household members who are 15-years-old or older.

Age is taken as the number of completed years as of a person's most recent birthday.

According to the NSC, someone is counted as *working* if he/she performs an income-generating activity. For example, if a person works on a personal subsidiary plot and sells the produce, then he/she is counted as *working*. If the produce is not sold, however, then the person is not counted as *working*.

The response to this question should be marked based on the responses of the household to the battery of questions on the "Back-page Worksheet". In particular, the enumerator should *not* directly ask the respondent "In the past 7 days, how many household members worked or had paid employment for at least 1 hour, or worked on a family farm or enterprise, or (if they did not work in the past 7 days) had work or paid employment to which they plan to return?" Instead, the enumerator should count the number of household members who answered "Yes" to the person-by-person application of this question on the "Back-page Worksheet". Then the response to this scorecard indicator is marked accordingly.

No additional information is available for this indicator.

4. In their main work or paid employment in the past 7 days, how many household members worked for a wage paid in-cash or in-kind, or for a money allowance?
    A. None
    B. One
    C. Two or more

According to the "Employment and Labor" module of the 2012 KIHS, this question concerns only household members who are 15-years-old or older and who are counted as working (see the guidelines for the previous indicator).

The response to this question should be marked based on the responses of the household to the battery of questions on the "Back-page Worksheet". In particular, the enumerator should *not* directly ask the respondent "In their main work or paid employment in the past 7 days, how many household members worked for a wage paid in-cash or in-kind, or for a money allowance?" Instead, the enumerator should count the number of household members who answered "Yes" to the person-by-person application of this question on the "Back-page Worksheet". Then the response to this scorecard indicator is marked accordingly.

According to the 2012 KIHS questionnaire, self-employed people (whether in agricultural or non-agricultural activities) are not to be counted as receiving wages or monetary allowances for the purposes of this question.

According to the *Enumerator Manual*, the concept of *wages in cash* includes:

- Wages accrued for work performed (time worked) at piece-rates, tariff rates, a basic salary or a percentage of revenue, or as a proportion of income, regardless of the form and pay systems adopted in the enterprise
- Incentive payments (bonuses based on performance, year-end bonuses, long-service benefits, seniority benefits, and performance-based supplements to tariff rates or salaries)
- Remuneration related to the mode of operation and working conditions (for example, payment due to regional regulation of wages, payment for work in special circumstances, allowances for travel or for the itinerant nature of work for employees engaged in construction or hauling, field allowances, night work, overtime, or work on weekends and holidays)
- Payment for time worked in accordance with the law (for example, payment of annual and additional holidays, monetary compensation for unused vacation, overtime pay, payment for delays that are not the fault of the workers, or wages for training)

- Payment for the production of goods (products or services) that are damaged through no fault of the employee
- Additional regular allowances (for example, compensation for rising prices in excess of government indexation, compensation to increase purchasing power in the employee cafeteria, coffee shops, dispensaries, allowances for housing costs, travel costs to and from work, or per diem)
- Remuneration of part-time workers
- Pay to the students of higher educational institutions and specialized secondary and vocational schools who are doing practical training/apprenticeships/internships in the enterprise, as well as pay for econdary school students who are receiving vocational training
- Amounts charged to work at a factory under special labor contracts with government agencies
- Commissions, cash gratuities, or gifts received by employees
- Royalty fees for writers in newspapers, magazines and other media, and the fees paid to freelance workers (for example, for lectures, performances on radio/television)
- Wages of domestic workers.

The concept of *wages in kind* includes:

- Value of goods and services given by enterprises for their employees in the form of the products produced (for example, corn given to employees of agricultural firms)
- Value provided to employees free of charge (in accordance with the legislation) for utilities, food, uniforms and uniforms that employees can wear home, or the benefit derived from their sale at reduced prices
- Costs of providing employees with free housing or cash compensation for not providing free housing (in accordance with the law)
- Travel expenses of employees to and from work by public transport
- Insurance premiums paid by companies to benefit individual employees
- School fees for the children of employees that are paid by an employer
- Provision of free or reduced-price food to workers in cafeterias, coffee shops, or dispensaries
- Provision of vouchers to employees and their families for medical treatment, rest, excursions, and travel at the expense of the enterprise
- Costs of reimbursement of the price difference for products or services provided to employees
- Costs of payment provided to employees of housing, rent, or lodging in a hotel
- Cost of equity shares (or discounts on shares) issued to employees as incentives
- Other in-kind benefits (for example, payment of premiums for health-care insurance, lessons, clubs, subscriptions to newspapers and magazines, and prosthetics)

5. What is the main source of water used by the household?
    A. Public (communal) water pump, storage reservoir, river, lake, pond, *aryk*, spring, or purchased water (water cart)
    B. Artesian well
    C. Private water pump
    D. Well, or aqueduct (running water)

No additional information is available for this indicator.

6. Does the household have any regular or automatic washing machines?
    A. No
    B. Regular (but not automatic)
    C. Automatic (regardless of regular)

According to the *Manual*, count only washing machines that are in working condition and that are available for use in the household or that have been temporarily loaned to family or friends. Also count washing machines that are currently being repaired. Do not count washing machines that the household has borrowed or rented. Likewise, do not count broken washing machines that cannot be repaired.

The relationship between possible combinations of responses and the response-option to be marked is:

Has regular washing machine?	Has automatic washing machine?	Response-option to mark
No	No	A. No
No	Yes	C. Automatic (regardless of regular)
Yes	No	B. Regular (but not automatic)
Yes	Yes	C. Automatic (regardless of regular)

7. Does the household have any electric heaters?
    A. No
    B. Yes

According to the *Manual*, count only electric heaters that are in working condition and that are available for use in the household or that have been temporarily loaned to family or friends. Also count electric heaters that are currently being repaired. Do not count electric heaters that the household has borrowed or rented. Likewise, do not count broken electric heaters that cannot be repaired.

8. How many cellular telephones does the household have?
    A. None, or one
    B. Two
    C. Three or more

According to the *Manual*, count only cellular telephones that are in working condition and that are available for use in the household or that have been temporarily loaned to family or friends. Also count cellular telephones that are currently being repaired. Do not count cellular telephones that the household has borrowed or rented. Likewise, do not count broken cellular telephones that cannot be repaired.

9. Does the household have any bicycles or any automobiles, trucks, minivans, motorcycles, scooters, mopeds, or motorized bicycles?
    A. No
    B. Only bicycle
    C. Motorized vehicle (regardless of bicycle)

According to the *Manual*, count only bicycles, automobiles, trucks, minivans, motorcycles, scooters, mopeds, or motorized bicycles that are in working condition and that are available for use in the household or that have been temporarily loaned to family or friends. Also count bicycles, automobiles, trucks, minivans, motorcycles, scooters, mopeds, or motorized bicycles that are currently being repaired. Do not count bicycles, automobiles, trucks, minivans, motorcycles, scooters, mopeds, or motorized bicycles that the household has borrowed or rented. Likewise, do not count broken bicycles, automobiles, trucks, minivans, motorcycles, scooters, mopeds, or motorized bicycles that cannot be repaired.

The relationship between the possible combinations of responses and the response-option to be marked is:

Has bicycle?	Has automobile, truck, minivan, motorcycle, scooter, moped, or motorized bicycle?	Response-option to mark
No	No	A. No
No	Yes	C. Motorized vehicle (regardless of bicycle)
Yes	No	B. Only bicycle
Yes	Yes	C. Motorized vehicle (regardless of bicycle)

10. Does the household use any personal agricultural plots? If so, has the household in the past 12 months had any sheep, lambs, goats, kids, cows, heifers, calves, bulls, oxen, horses, donkeys, hinnies, mules, or yaks (of any age)?
    A. No plot (regardless of animals)
    B. Has a plot, but no animals
    C. Has both a plot and animals

The relationship between possible combinations of responses and the response-option to be marked is:

Use of personal agricultural plot?	Had, in the past 12 months, sheep, lambs, goats, kids, cows, heifers, calves, bulls, oxen, horses, donkeys, hinnies, mules, or yaks (of any age)?	Response-option to mark
No	No	A. No plot (regardless of animals)
No	Yes	A. No plot (regardless of animals)
Yes	No	B. Has a plot, but no animals
Yes	Yes	C. Has both a plot and animals

According to the *Manual*, citizens of the Kyrgyz Republic have the right to receive a lifetime, inheritable plot of land for subsistence farming, cash farming, for constructing a residence, for horticulture and animal husbandry, for urban development, or for traditional folk arts and crafts.

Personal agricultural plots with lifetime, inheritable rights are granted to citizens for private farming. The are to be used for residences, commercial buildings, fruit trees, and crops. Furthermore, citizens may be allocated land located outside the town limits for gardening or for growing hay. Furthermore, former collective land may now be allocated to citizens to grow crops.

Figure 1: Poverty lines and poverty rates (for households and people) for the Kyrgyz Republic for 2012 and by sub-sample

				Poverty rates (% with consumption below a given poverty line) and poverty lines (KGS/day/person)							
				National poverty lines				Intl. 2005 PPP			
Sample	**Line or rate**	**Level**	***n***	**100%**	**150%**	**200%**	**Median**	**$1.25**	**$2.00**	**$2.50**	**$5.00**
All Kyrgyzstan											
	Line	People		72.40	108.60	144.80	60.06	32.55	52.08	65.10	130.20
	Rate	HHs	4,936	28.5	64.4	81.6	13.1	0.4	7.2	18.7	76.4
		People		38.1	76.9	90.7	19.1	0.6	10.9	26.5	86.8
Construction and calibration (Selecting indicators and points, and associating scores with likelihoods)											
	Rate	HHs	2,514	28.5	64.7	81.4	12.8	0.4	7.2	18.5	76.5
Validation (measuring accuracy)											
	Rate	HHs	2,422	28.5	64.1	81.7	13.3	0.4	7.1	18.9	76.3

Source: 2012 KIHS. Poverty lines in average calendar-year prices for Kyrgyzstan as a whole.
All poverty lines are per-person.

Figure 2: Poverty lines and poverty rates (for households and people) by oblast and by urban/rural for 2012

Poverty-line region	Line or rate	Level	n	Poverty rates (% with consumption below a given poverty line) and poverty lines (KGS/day/person)							
				National poverty lines				Intl. 2005 PPP			
				100%	150%	200%	Median	$1.25	$2.00	$2.50	$ 5.00
All Kyrgyzstan	Line			72.40	108.60	144.80	60.06	32.55	52.08	65.10	130.20
	Rate	HHs	4,936	28.5	64.4	81.6	13.1	0.4	7.2	18.7	76.4
	Rate	People		38.1	76.9	90.7	19.1	0.6	10.9	26.5	86.8
Issykul, urban	Line			67.62	101.43	135.24	57.37	30.40	48.64	60.80	121.60
	Rate	HHs	392	15.7	48.9	73.6	7.0	0.0	2.7	8.8	63.8
	Rate	People		22.2	61.2	82.4	11.3	0.0	4.3	13.4	74.7
Issykul, rural	Line			68.28	102.42	136.56	56.77	30.70	49.11	61.39	122.78
	Rate	HHs	256	22.3	49.9	73.5	10.0	0.0	5.6	14.4	65.7
	Rate	People		31.0	62.6	83.4	15.4	0.0	9.2	21.2	74.7
Jalal-Abad, urban	Line			74.37	111.55	148.74	51.69	33.43	53.49	66.87	133.74
	Rate	HHs	416	50.5	79.6	92.0	22.1	2.7	24.7	41.2	88.5
	Rate	People		61.7	88.8	96.1	30.9	4.3	33.6	52.4	94.4
Jalal-Abad, rural	Line			73.80	110.69	147.59	60.10	33.18	53.08	66.35	132.70
	Rate	HHs	244	44.6	83.3	93.8	20.3	0.0	15.7	30.0	92.4
	Rate	People		53.4	89.6	97.2	26.7	0.0	21.6	38.3	96.4

Source: 2012 KIHS. Poverty lines in average calendar-year prices for Kyrgyzstan as a whole.

Figure 2 (cont.): Poverty lines and poverty rates (for households and people) by oblast and by urban/rural for 2012

Poverty-line region	Line or rate	Level	n	Poverty rates (% with consumption below a given poverty line) and poverty lines (KGS/day/person)							
				National poverty lines				Intl. 2005 PPP			
				100%	150%	200%	Median	$1.25	$2.00	$2.50	$ 5.00
Naryn, urban	Line			70.75	106.12	141.50	57.76	31.81	50.89	63.61	127.23
	Rate	HHs	262	21.1	62.3	84.3	10.8	0.0	6.1	14.0	76.8
	Rate	People		27.0	70.1	90.1	13.5	0.0	8.2	18.0	84.3
Naryn, rural	Line			71.00	106.50	142.01	59.92	31.92	51.07	63.84	127.68
	Rate	HHs	261	37.4	71.8	90.1	17.2	0.0	6.0	22.3	86.3
	Rate	People		41.6	77.9	94.0	20.5	0.0	7.2	26.4	91.2
Batken, urban	Line			74.72	112.09	149.45	63.00	33.59	53.75	67.19	134.38
	Rate	HHs	268	30.0	57.2	79.9	15.6	0.0	4.2	19.7	72.8
	Rate	People		38.9	67.7	87.0	19.2	0.0	5.9	25.4	80.9
Batken, rural	Line			70.35	105.52	140.70	57.39	31.63	50.60	63.25	126.50
	Rate	HHs	233	26.5	72.0	88.4	12.0	0.0	4.9	19.0	83.6
	Rate	People		32.5	81.4	95.0	15.9	0.0	6.1	24.1	91.4
Osh, urban	Line			73.71	110.56	147.42	64.16	33.14	53.02	66.27	132.55
	Rate	HHs	396	43.0	74.5	88.8	20.4	0.4	4.6	27.9	84.8
	Rate	People		54.8	85.6	95.7	27.4	0.5	6.4	37.1	93.6
Osh, rural	Line			73.76	110.64	147.52	62.37	33.16	53.06	66.32	132.64
	Rate	HHs	264	39.7	81.9	93.1	17.9	0.0	8.2	23.2	90.8
	Rate	People		50.0	90.2	97.6	25.0	0.0	12.5	32.1	96.2

Source: 2012 KIHS. Poverty lines in average calendar-year prices for Kyrgyzstan as a whole.

Figure 2 (cont.): Poverty lines and poverty rates (for households and people) by oblast and by urban/rural for 2012

				Poverty rates (% with consumption below a given poverty line) and poverty lines (KGS/day/person)							
				National poverty lines				Intl. 2005 PPP			
Poverty-line region	**Line or rate**	**Level**	***n***	**100%**	**150%**	**200%**	**Median**	**$1.25**	**$2.00**	**$2.50**	**$ 5.00**
Talas, urban	Line			70.74	106.12	141.49	64.78	31.80	50.89	63.61	127.22
	Rate	HHs	264	17.5	65.7	82.4	9.2	0.0	1.7	8.6	76.5
	Rate	People		23.6	77.6	90.8	11.8	0.0	2.4	11.0	86.5
Talas, rural	Line			68.76	103.14	137.52	59.24	30.91	49.46	61.83	123.65
	Rate	HHs	264	32.1	67.6	88.0	14.9	0.0	3.8	23.4	78.7
	Rate	People		42.2	78.4	93.4	21.1	0.0	6.0	31.9	87.0
Chui, urban	Line			71.66	107.49	143.32	51.41	32.22	51.54	64.43	128.86
	Rate	HHs	253	18.6	42.1	61.9	7.7	2.6	8.6	14.5	55.8
	Rate	People		24.0	52.1	70.5	12.0	5.2	12.9	19.7	65.7
Chui, rural	Line			72.37	108.55	144.74	60.89	32.53	52.06	65.07	130.14
	Rate	HHs	402	11.2	46.0	69.7	5.1	1.0	4.2	6.3	62.0
	Rate	People		14.8	58.0	81.7	7.6	1.3	6.4	9.3	74.5
Bishkek, urban	Line			72.52	108.78	145.04	61.49	32.60	52.16	65.21	130.41
	Rate	HHs	716	14.2	51.0	70.1	7.0	0.1	2.1	9.6	63.5
	Rate	People		21.5	67.2	84.0	10.8	0.0	3.5	14.6	78.6

Source: 2012 KIHS. Poverty lines in average calendar-year prices for Kyrgyzstan as a whole.

Figure 3: Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (Responses ordered starting with those most strongly linked with higher poverty likelihoods)
1,630	How many household members are there? (Seven or more; Six; Five; Four; Three; Two; One)
1,079	How many household members are 18-years-old or younger? (Four or more; Three; Two; One; None)
1,015	How many household members are 12-years-old or younger? (Three or more; Two; One; None)
990	How many household members are 16-years-old or younger? (Three or more; Two; One; None)
988	How many household members are 14-years-old or younger? (Three or more; Two; One; None)
977	How many household members are 15-years-old or younger? (Three or more; Two; One; None)
964	How many household members are 17-years-old or younger? (Three or more; Two; One; None)
950	How many household members are 13-years-old or younger? (Three or more; Two; One; None)
892	How many household members are 11-years-old or younger? (Three or more; Two; One; None)
835	Do the members of the household wash themselves in a their own bath, shower, or sauna, in a public bath, shower room, or sauna, or somewhere else? (Public bath, shower room, or sauna, but not own bath, shower, or sauna (regardless of somewhere else); None; Only somewhere else; Own bath, shower, or sauna (regardless of any others))
687	In which oblast does the household reside? (Jalal-Abad; Naryn; Osh; Bishkek; Issykul; Talas; Chui; Batken)
668	What is the main source of water used by the household? (Public (communal) water pump, storage reservoir, river, lake, pond, *aryk*, spring, or purchased water (water cart); Artesian well; Private water pump; Well, or aqueduct (running water))
602	How many landline and cellular telephones does the household have? (None; One; Two; Three; Four or more)
572	How many household members are 6-years-old or younger? (Two or more; One; None)

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

Uncertainty coefficient	Indicator (Responses ordered starting with those most strongly linked with higher poverty likelihoods)
568	Does the household cook with an oven or fireplace, electric fryer, gas stove with central gas supply, or gas stove with bottled gas? (Electric fryer, but not gas stove of any kind, (regardless of oven or fireplace); Gas stove with central gas supply, but not gas stove with bottled gas (regardless of electric fryer, oven, or fireplace); Gas stove with bottled gas (regardless of any others); Only oven or fireplace; None)
540	Does the household have any 1-chamber refrigerators, 2- or 3-chamber refrigerators, or any freezers? (No; 1-chamber, but no others; 2- or 3-chamber or freezer (regardless of 1-chamber))
538	Does the household have any landline telephones? (No; Yes)
536	Does the dwelling have a telephone? (No; Yes)
484	Is the dwelling connected to a sewer system? (No; Yes)
470	Does the dwelling have water supply? (No; Yes)
458	What is the highest level of education attained by the female head/spouse? (General secondary (complete) degree, primary professional technical (without general secondary), or incomplete higher degree; Illiterate, no elementary education, elementary education, primary professional (vocational) degree, or other; Secondary special degree, main secondary (incomplete) degree, primary professional technical (with general secondary); Higher degree; No female head/spouse)
443	Does the household have any regular or automatic washing machines? (No; Regular (but not automatic); Automatic (regardless of regular))
355	In the main work or paid employment of the female head/spouse in the past seven days, what was the principal area of activity? (On an individual basis, or individual commercial activities; Does not work; At a (peasant) farm; Wage work for private individual (individual entrepreneurs); In an enterprise, organization, collective farm, association of farms, agricultural cooperative, or other institution; No female head/spouse)

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

Uncertainty coefficient	Indicator (Responses ordered starting with those most strongly linked with higher poverty likelihoods)
335	How many household members worked on a (peasant) farm? (Two or more; One; None)
323	In their main work or paid employment in the past seven days, how many household members were laborers in elementary occupations or were skilled agricultural or fishery workers? (Two or more; One; None)
313	Does your dwelling have furnace (stove) heating, electric or electric-radiator heating, central heating, or gas heating? (None, or only furnace (stove); Electric or electric-radiator, but not central or gas (regardless of furnace (stove))l Central or gas (regardless of all others))
290	In their main work or paid employment in the past seven days, how many household members were skilled agricultural and fishery workers? (Two or more; One; None)
278	What type of dwelling does the household live in? (Separate house, lodge or a tied cottage (temporary-tenure dwelling), other non-residential premises used for residence, or other residential premises; Part of a house; Separate apartment in a multi-story house; Apartment or room in a residential hotel, or dormitory)
277	In the main work or paid employment of the female head/spouse in the past seven days, what was her occupation? (Elementary occupations; Does not work; Service workers, and shop and market sales workers, or plant and machinery operators and assemblers; Skilled agricultural and fishery workers; Craft and related trades workers; Legislators, senior officials, managers, professionals, technicians, associate professionals, or clerks; No female head/spouse)
277	What is the area (in square meters) of the living space occupied by your household? (0 to 25; 26 to 30; 31 to 35; 36 to 40; 41 to 45; 46 to 55; 56 to 65; 66 to 75; 75 to 85; 86 or more)
275	Does the dwelling have a bath or shower? (No; Yes)
271	In their main work or paid employment in the past seven days, how many household members were skilled agricultural or fishery workers, and if any were, does the household use any personal agricultural plot? (In agriculture (with or without a plot); No one in agriculture)

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

Uncertainty coefficient	Indicator (Responses ordered starting with those most strongly linked with higher poverty likelihoods)
267	What type of bathroom do the members of the household usually use? (Toilet with a cesspool, water closet with individual sewer system, other, or does not use a toilet; Water closet connected to the central sewer system)
254	How many cellular telephones does the household have? (None, or one; Two; Three)
254	Does the household have any electrical vacuum cleaners? (No; Yes)
246	How many beds does the household have? (None; One; Two; Three or more)
244	Does the dwelling have hot-water supply? (No; Yes)
241	In the main work or paid employment of the male head/spouse in the past seven days, what was the principal area of activity? (At a (peasant) farm, or in individual commercial activities; Does not work; Wage work for private individual (individual entrepreneurs); On an individual basis; In an enterprise, organization, collective farm, association of farms, agricultural cooperative, or other institution; No male head/spouse)
241	How much time does it normally take to reach the nearest bus/public transport stop? (16 minutes or more; 6 to 15 minutes; Five minutes or less)
229	Does your household dispose of garbage by burning or burying, in heaping dump, or with a garbage chute or waste-collecting containers or trucks? (None; Only burning or burying; In heaping dump, but not garbage chute or waste-collecting containers or trucks (regardless of burning or burying); With garbage chute or waste-collecting containers or trucks (regardless of any others))
227	What is the highest level of education attained by the male head/spouse? (Illiterate, no elementary education, or other; General secondary (complete) degree, primary professional technical (without general secondary), or incomplete higher degree; Primary professional technical (with general secondary); Secondary special degree, or main secondary (incomplete) degree; Elementary education, or primary professional (vocational) degree; Higher degree; No male head/spouse)

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

Uncertainty coefficient	Indicator (Responses ordered starting with those most strongly linked with higher poverty likelihoods)
219	Does the dwelling have an individual heating system or central (district) heating? (None, or individual; Central (district))
218	What is the main contruction material of the walls of the dwelling? (Earth, clay, tarpaulin, canvas, felt, slag-concrete block, or other; Wood, logs; Crude airbricks, adobe; Bricks; Concrete slabs)
201	In the main work or paid employment of the male head/spouse in the past seven days, what was his occupation? (Skilled agricultural and fishery workers; Does not work; Legislators, senior officials, managers, professionals, technicians, associate professionals, or clerks; Craft and related trades workers; Elementary occupations; Service workers, and shop and market sales workers; Plant and machinery operators and assemblers; No male head/spouse)
190	If the household uses any personal agricultural plot, then has the household, in the past 12 months, had any sheep, lambs, goats, or kids (of any age)? (Has a plot, and also has some sheep, lambs, goats, or kids (of any age); Has a plot, but does not have any sheep, lambs, goats, or kids (of any age); No plot)
189	If the household uses any personal agricultural plot, then has the household, in the past 12 months, had any cows, heifers, calves, bulls, oxen, horses, donkeys, hinnies, mules, or yaks (of any age)? (Has a plot, and also has some cows, heifers, calves, bulls, oxen, horses, donkeys, hinnies, mules, or yaks (of any age); Has a plot, but does not have any cows, heifers, calves, bulls, oxen, horses, donkeys, hinnies, mules, or yaks (of any age); No plot)
184	If the household uses any personal agricultural plot, then has the household, in the past 12 months, had any cows, heifers, calves, bulls, or oxen (of any age)? (Has a plot, and also has some cows, heifers, calves, bulls, or oxen (of any age); Has a plot, but does not have any cows, heifers, calves, bulls, or oxen (of any age); No plot)
183	Does the household use any personal agricultural plots? If so, has the household in the past 12 months had any sheep, lambs, goats, kids, cows, heifers, calves, bulls, oxen, horses, donkeys, hinnies, mules, or yaks (of any age)? (No plot (regardless of animals); Has a plot, but no animals; Has both a plot and animals)

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

Uncertainty coefficient	Indicator (Responses ordered starting with those most strongly linked with higher poverty likelihoods)
183	If the household uses any personal agricultural plot, then has the household, in the past 12 months, had any horses, donkeys, hinnies, mules, or yaks (of any age)? (Has a plot, but does not have any horses, donkeys, hinnies, mules, or yaks (of any age); No plot; Has a plot, and also has some horses, donkeys, hinnies, mules, or yaks (of any age))
182	In what altitude zone is the residence located? (High-mountainous; Flat ground; Semi-mountainous)
181	In their main work or paid employment in the past seven days, how many household members worked in an enterprise, organization, collective farm, association of farms, agricultural cooperative, or other institution? (None; One; Two or more)
180	Does the household use any personal agricultural plot? (Yes; No)
171	How many carpets, oriental carpets, or *scherdak* (rugs) does the household have? (None; One; Two; Three; Four; Five; Six; Seven; Eight; Nine or more)
171	Does the household have any automatic washing machines? (No; Yes)
170	Does the dwelling have central gas supply? (No; Yes)
162	What is the main construction material of the roof of the dwelling? (Sheet-metal plates, tin plates, cane, rushes, reeds, wood, crude airbricks, adobe, roofing tar board, or other; Roofing slates; Concrete slabs, or roof tiling)
161	Does the household have any regular washing machines? (No; Yes)
148	In their main work or paid employment in the past seven days, how many household members worked for a wage paid in-cash or in-kind, or for a money allowance? (None; One; Two or more)
147	Does the household have any (electric) irons? (No; Yes)
145	Does the household have furniture for the guest room? (No; Yes)
140	Does the household have any personal computers? (No; Yes)
136	How many living rooms does your dwelling consist of? (One; Two; Three; Four; Five or more)

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

Uncertainty coefficient	Indicator (Responses ordered starting with those most strongly linked with higher poverty likelihoods)
131	Does the household have any wireless receivers, or music centers or audio recorders? (None; Only wireless receiver; Music centers or audio recorders (regardless of wireless receivers))
119	Does the household have an armchair? (No; Yes)
115	How many household members are professionals? (None; One; Two or more)
104	In the past 7 days, how many household members worked or had paid employment for at least 1 hour, or worked on a family farm or enterprise, or (if they did not work in the past 7 days) had work or paid employment to which they plan to return? (None; One; Two; Three or more)
96	Does the household have any sewing machines or knitting machines? (No; Yes)
95	Does the household have a kitchen suite or bedroom suite? (No; Yes)
95	Does the household have any music centers or audio recorders? (No; Yes)
92	How many household members are currently employed? (None; One; Two; Three or more)
90	Does the household have a sofa or couch? (No; Yes)
76	What is the marital status of the female head/spouse? (Legally married; Married under common law (civil marriage), or widow; No female head/spouse; Divorced; Living apart, but not divorced, or never-married)
74	How many cupboards, dressers, sideboards, or wall cabinets does the household have? (None; One; Two; Three or more)
72	Does the household have any black-and-white or color televisions, and does it have any video recorders, video players, or satellite antennas? (No TVs (regardless of VCRs etc.); Black-and-white TV, but no color TV (regardless of VCRs etc.); Color TV (regardless of black-and-white TVs), but no VCRs etc.; Color TV and VCRs etc. (regardless of black-and-white TVs))
72	Does the household have any electric heaters? (No; Yes)

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

Uncertainty coefficient	Indicator (Responses ordered starting with those most strongly linked with higher poverty likelihoods)
70	Does the household have any black-and-white or color televisions? (Only black-and-white; None; One or more color (regardless of black-and-white))
68	Does your household have its own garage? (No; Yes)
63	Does the household have any color televisions? (No; Yes)
61	In the past seven days, did the male head/spouse work or have paid employment for at least one hour, or did he work on a family farm or enterprise, or (if he did not work) did he have work or paid employment to which he plans to return? (No; Yes; No male head/spouse)
57	Does the household have any bicycles or any automobiles, trucks, minivans, motorcycles, scooters, mopeds, or motorized bicycles? (No; Only bicycle; Motorized vehicle (regardless of bicycle))
57	Does the household have any wireless receivers? (No; Yes)
55	Does the household have any black-and-white televisions? (Yes; No)
53	Does the household have a microwave? (No; Yes)
51	Does the household have any automobiles, trucks, minivans, motorcycles, scooters, mopeds, or motorized bicycles? (No; Yes)
48	In the past seven days, did the female head/spouse work or have paid employment for at least one hour, or did she work on a family farm or enterprise, or (if she did not work) did she have work or paid employment to which she plans to return? (No; Yes; No female head/spouse)
42	In their main work or paid employment in the past seven days, did any household members work in individual commercial activities or on an individual basis? (Yes; No)
41	Is the male head/spouse currently employed? (No; Yes; No male head/spouse)
39	Is the male head/spouse currently employed? (No; Yes; No male head/spouse)
34	What is the marital status of the male head/spouse? (Legally married; No male head/spouse; Married under common law (civil marriage), divorced, living apart, but not divorced, widower, or never-married)

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

Uncertainty coefficient	Indicator (Responses ordered starting with those most strongly linked with higher poverty likelihoods)
33	What is the structure of household headship? (Both male and female heads/spouses; Female head/spouse only; Male head/spouse only)
31	In their main work or paid employment in the past seven days, were any household members in elementary occupations? (Yes; No)
22	Does the household have any bicycles? (No; Yes)
11	In their main work or paid employment in the past seven days, was the male or female head/spouse self-employed in non-agriculture? (Yes; No)
9	In their main work or paid employment in the past seven days, how many household members were service workers, shop-and-market sales workers, craft and related trades workers, or plant and machinery operators and assemblers? (None; One; Two or more)
4	Does the household have any video recorders, video players, or satellite antennas? (No; Yes)
2	Does the household have an electric furnace? (No; Yes)

Source: 2012 KIHS and 100% of the national poverty line

Tables for
100% of the National Poverty Line

(and Tables Pertaining
to All Eight Poverty Lines)

Figure 4 (100% of the national line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	98.0
10–14	96.6
15–19	89.9
20–24	77.3
25–29	68.0
30–34	46.5
35–39	40.7
40–44	15.8
45–49	11.7
50–54	4.1
55–59	3.9
60–64	2.9
65–69	1.3
70–74	1.2
75–79	1.2
80–84	1.2
85–89	1.2
90–94	1.2
95–100	1.2

Figure 5 (100% of the national line): Derivation of estimated poverty likelihoods associated with scores

Score	Households at score and < poverty line		All households at score		Poverty likelihood (%)
0–4	146	÷	146	=	100.0
5–9	229	÷	233	=	98.0
10–14	2,271	÷	2,350	=	96.6
15–19	2,009	÷	2,234	=	89.9
20–24	3,784	÷	4,895	=	77.3
25–29	4,485	÷	6,595	=	68.0
30–34	6,618	÷	14,229	=	46.5
35–39	4,271	÷	10,481	=	40.7
40–44	1,844	÷	11,665	=	15.8
45–49	1,341	÷	11,440	=	11.7
50–54	368	÷	9,061	=	4.1
55–59	298	÷	7,677	=	3.9
60–64	158	÷	5,414	=	2.9
65–69	34	÷	2,689	=	1.3
70–74	9	÷	775	=	1.2
75–79	7	÷	598	=	1.2
80–84	0	÷	19	=	1.2
85–89	0	÷	0	=	1.2
90–94	0	÷	0	=	1.2
95–100	114	÷	9,500	=	1.2

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

Figure 6: Probability that a given household's consumption falls in a range demarcated by two adjacent poverty lines

	Likelihood (%) that daily per-capita consumption is in a range demarcated by adjacent poverty lines								
	<$1.25/day	**≥$1.25/day and <$2.00/day**	**≥$2.00/day and <Median**	**≥Median and <$2.50/day**	**≥$2.50/day and <100% Natl.**	**≥100% Natl. and <150% Natl.**	**≥150% Natl. and <$5.00/day**	**≥$5.00/day and <200% Natl.**	**≥200% Natl.**
Score	**<KGS32.55**	**≥KGS32.55 and <KGS52.08**	**≥KGS52.08 and <KGS60.06**	**≥KGS60.06 and <KGS65.10**	**≥KGS65.10 and <KGS72.40**	**≥KGS72.40 and <KGS108.60**	**≥KGS108.60 and <KGS130.20**	**≥KGS130.20 and <KGS144.80**	**≥KGS144.80**
0–4	5.8	83.1	2.5	8.5	0.0	0.0	0.0	0.0	0.0
5–9	1.6	69.8	17.2	3.4	6.0	2.0	0.0	0.0	0.0
10–14	1.4	49.9	8.6	18.9	17.9	3.4	0.0	0.0	0.0
15–19	1.4	37.3	16.6	19.2	15.4	10.1	0.0	0.0	0.0
20–24	1.4	22.9	15.2	18.8	19.0	22.5	0.2	0.0	0.0
25–29	1.1	14.7	14.6	15.2	22.3	31.0	0.9	0.0	0.0
30–34	0.9	8.4	11.0	6.9	19.4	49.6	3.6	0.3	0.0
35–39	0.3	4.2	8.1	6.6	21.6	48.7	6.6	3.9	0.1
40–44	0.1	3.8	2.3	3.8	5.9	60.8	15.5	3.5	4.4
45–49	0.1	2.4	3.0	1.9	4.4	53.5	20.0	5.3	9.5
50–54	0.0	0.2	1.0	1.9	1.0	39.8	25.6	11.0	19.5
55–59	0.0	0.2	0.5	1.4	1.7	23.4	23.0	9.9	39.8
60–64	0.0	0.2	0.1	1.2	1.4	18.4	18.4	11.2	49.0
65–69	0.0	0.2	0.1	0.5	0.4	5.6	18.1	13.8	61.2
70–74	0.0	0.2	0.1	0.5	0.3	3.3	15.1	12.0	68.4
75–79	0.0	0.2	0.1	0.5	0.3	3.3	7.6	6.8	81.1
80–84	0.0	0.2	0.1	0.5	0.3	3.3	7.6	6.6	81.3
85–89	0.0	0.2	0.1	0.5	0.3	3.3	7.6	6.6	81.3
90–94	0.0	0.2	0.1	0.5	0.3	3.3	7.6	6.6	81.3
95–100	0.0	0.2	0.1	0.5	0.3	3.3	7.6	6.6	81.3

Figure 7 (100% of the national line): Average differences between estimated and true poverty likelihoods for households in 1,000 bootstraps of n = 16,384, with confidence intervals by score range, scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+0.0	0.0	0.0	0.0
5–9	+0.0	0.0	0.0	0.0
10–14	–4.3	2.1	2.1	2.1
15–19	+1.0	2.6	3.0	4.0
20–24	+16.9	3.9	4.7	6.1
25–29	+16.4	3.3	3.9	5.3
30–34	–21.2	11.8	12.0	12.3
35–39	+24.4	1.6	1.9	2.5
40–44	+1.0	1.6	1.9	2.4
45–49	+4.8	1.1	1.3	1.6
50–54	–2.7	2.0	2.1	2.3
55–59	–25.3	14.4	14.9	15.6
60–64	+0.3	0.9	1.1	1.5
65–69	+1.7	0.0	0.0	0.0
70–74	+1.2	0.0	0.0	0.0
75–79	+1.2	0.0	0.0	0.0
80–84	+1.2	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	–3.8	2.6	2.7	3.0

Figure 8 (100% of the national line): Average differences between estimated poverty rates and true values for a group at a point in time in 1,000 bootstraps of various sample sizes, with confidence intervals, scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+1.8	64.3	78.5	87.9
4	+0.9	41.6	48.6	66.9
8	–0.7	31.8	38.1	49.9
16	–1.2	23.6	28.7	35.8
32	–1.4	17.1	20.2	25.4
64	–1.9	12.7	14.9	18.6
128	–2.1	8.7	10.5	14.0
256	–2.0	6.1	7.3	9.7
512	–2.2	4.1	4.9	6.4
1,024	–2.2	3.0	3.4	4.8
2,048	–2.3	2.1	2.5	3.3
4,096	–2.3	1.5	1.8	2.3
8,192	–2.3	1.1	1.2	1.6
16,384	–2.3	0.7	0.9	1.1

Figure 9 (All poverty lines): Average differences between estimates and true values for poverty rates of a group of households at a point in time in 1,000 bootstraps of n = 16,384, precision, and the α factor for precision, scorecard applied to the validation sample

	Poverty line							
	National poverty lines				Intl. 2005 PPP			
	100%	150%	200%	Median	$1.25	$2.00	$2.50	$5.00
Estimate minus true value	–2.3	–1.0	–2.2	–1.2	+0.3	+2.0	–1.4	–2.5
Precision of difference	0.7	0.5	0.4	0.8	0.0	0.4	0.8	0.5
α factor for precision	1.29	0.88	0.88	1.82	0.32	1.14	1.45	0.96

Differences between estimates and true values are displayed in units of percentage points.
Precision is measured as 90-percent confidence intervals in units of ± percentage points.
Differences and precision estimated from 1,000 bootstraps with n = 16,384.
α is estimated from 1,000 bootstrap samples of n = 256, 512, 1,024, 2,048, 4,096, 8,192, and 16,384.

Figure 10 (All poverty lines): Possible targeting outcomes

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

Figure 11 (100% of the national line): Shares of households by cut-off score and targeting classification, along with the hit rate and BPAC, scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.1	28.3	0.0	71.5	71.7	–99.0
≤9	0.4	28.1	0.0	71.5	71.9	–97.3
≤14	2.6	25.9	0.2	71.4	73.9	–81.4
≤19	4.5	24.0	0.5	71.1	75.6	–66.8
≤24	8.0	20.5	1.9	69.7	77.7	–37.3
≤29	12.0	16.5	4.5	67.1	79.1	–0.1
≤34	19.5	9.0	11.2	60.3	79.8	+60.6
≤39	22.4	6.0	18.7	52.8	75.3	+34.2
≤44	25.0	3.4	27.8	43.7	68.8	+2.3
≤49	26.4	2.0	37.8	33.7	60.2	–32.9
≤54	27.1	1.4	46.2	25.3	52.4	–62.5
≤59	28.1	0.4	52.9	18.6	46.7	–85.9
≤64	28.2	0.3	58.3	13.3	41.4	–104.7
≤69	28.2	0.3	61.0	10.6	38.7	–114.1
≤74	28.2	0.3	61.7	9.8	38.0	–116.9
≤79	28.2	0.3	62.3	9.2	37.4	–119.0
≤84	28.2	0.3	62.3	9.2	37.3	–119.0
≤89	28.2	0.3	62.3	9.2	37.3	–119.0
≤94	28.2	0.3	62.3	9.2	37.3	–119.0
≤100	28.5	0.0	71.5	0.0	28.5	–151.3

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

Figure 12 (100% of the national line): By score cut-off, the share of all households who are targeted (that is, score at or below the cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), scorecard applied to the validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	100.0	0.5	Only poor targeted
≤9	0.4	100.0	1.3	Only poor targeted
≤14	2.7	93.8	9.0	15.0:1
≤19	5.0	90.4	15.8	9.5:1
≤24	9.9	81.1	28.1	4.3:1
≤29	16.5	72.9	42.1	2.7:1
≤34	30.7	63.4	68.4	1.7:1
≤39	41.2	54.5	78.9	1.2:1
≤44	52.8	47.4	87.9	0.9:1
≤49	64.3	41.2	92.9	0.7:1
≤54	73.3	36.9	95.1	0.6:1
≤59	81.0	34.7	98.7	0.5:1
≤64	86.4	32.6	98.9	0.5:1
≤69	89.1	31.6	98.9	0.5:1
≤74	89.9	31.3	98.9	0.5:1
≤79	90.5	31.1	98.9	0.5:1
≤84	90.5	31.1	98.9	0.5:1
≤89	90.5	31.1	98.9	0.5:1
≤94	90.5	31.1	98.9	0.5:1
≤100	100.0	28.5	100.0	0.4:1

Tables for
the 150% of the National Poverty Line

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	100.0
15–19	100.0
20–24	99.8
25–29	99.1
30–34	96.1
35–39	89.4
40–44	76.6
45–49	65.2
50–54	43.9
55–59	27.3
60–64	21.4
65–69	6.9
70–74	4.5
75–79	4.5
80–84	4.5
85–89	4.5
90–94	4.5
95–100	4.5

Figure 7 (150% of the national line): Average differences between estimated and true poverty likelihoods for households in 1,000 bootstraps of n = 16,384, with confidence intervals by score range, scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+0.0	0.0	0.0	0.0
5–9	+0.0	0.0	0.0	0.0
10–14	+0.0	0.0	0.0	0.0
15–19	+0.0	0.0	0.0	0.0
20–24	–0.1	0.1	0.1	0.1
25–29	+2.1	1.0	1.1	1.5
30–34	–0.9	0.7	0.8	0.9
35–39	+10.5	1.9	2.3	3.0
40–44	–2.7	2.1	2.3	2.7
45–49	–2.4	2.3	2.9	4.0
50–54	+2.2	2.4	2.9	3.8
55–59	–22.8	12.9	13.3	14.2
60–64	+12.7	1.6	1.9	2.5
65–69	–13.2	8.8	9.2	9.9
70–74	+4.6	0.0	0.0	0.0
75–79	+4.6	0.0	0.0	0.0
80–84	+4.6	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	–3.1	2.3	2.4	2.7

Figure 8 (150% of the national line): Average differences between estimated poverty rates and true values for a group at a point in time in 1,000 bootstraps of various sample sizes, with confidence intervals, scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	–2.6	58.7	75.5	93.6
4	–0.4	36.8	48.2	61.1
8	–0.4	27.3	34.0	46.5
16	–1.2	18.4	22.5	31.3
32	–1.2	12.5	15.8	22.0
64	–1.0	8.6	10.1	13.7
128	–1.2	6.0	7.2	10.1
256	–1.1	4.4	5.1	6.9
512	–1.1	3.1	3.6	5.0
1,024	–1.1	2.1	2.5	3.3
2,048	–1.1	1.5	1.9	2.4
4,096	–1.1	1.1	1.3	1.7
8,192	–1.1	0.8	1.0	1.3
16,384	–1.0	0.5	0.6	0.9

Figure 11 (150% of the national line): Shares of households by cut-off score and targeting classification, along with the hit rate and BPAC, scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.1	64.0	0.0	35.9	36.0	–99.5
≤9	0.4	63.8	0.0	35.9	36.2	–98.8
≤14	2.7	61.4	0.0	35.9	38.6	–91.5
≤19	5.0	59.2	0.0	35.9	40.8	–84.5
≤24	9.8	54.3	0.0	35.9	45.7	–69.3
≤29	16.3	47.8	0.2	35.7	52.0	–48.9
≤34	29.6	34.6	1.1	34.8	64.3	–6.0
≤39	38.3	25.9	2.9	33.0	71.3	+23.9
≤44	47.3	16.8	5.5	30.4	77.7	+56.2
≤49	55.0	9.1	9.3	26.6	81.6	+85.6
≤54	59.1	5.0	14.2	21.7	80.8	+77.9
≤59	61.7	2.4	19.3	16.6	78.4	+70.0
≤64	63.0	1.1	23.4	12.5	75.5	+63.5
≤69	63.2	0.9	25.9	10.0	73.2	+59.7
≤74	63.2	0.9	26.6	9.2	72.5	+58.5
≤79	63.2	0.9	27.2	8.6	71.9	+57.5
≤84	63.2	0.9	27.3	8.6	71.9	+57.5
≤89	63.2	0.9	27.3	8.6	71.9	+57.5
≤94	63.2	0.9	27.3	8.6	71.9	+57.5
≤100	64.1	0.0	35.9	0.0	64.1	+44.1

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

Figure 12 (150% of the national line): By score cut-off, the share of all households who are targeted (that is, score at or below the cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), scorecard applied to the validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	100.0	0.2	Only poor targeted
≤9	0.4	100.0	0.6	Only poor targeted
≤14	2.7	100.0	4.3	Only poor targeted
≤19	5.0	100.0	7.7	Only poor targeted
≤24	9.9	99.9	15.4	1,235.1:1
≤29	16.5	99.0	25.4	101.6:1
≤34	30.7	96.4	46.1	26.7:1
≤39	41.2	93.0	59.7	13.3:1
≤44	52.8	89.6	73.8	8.6:1
≤49	64.3	85.6	85.8	5.9:1
≤54	73.3	80.6	92.2	4.2:1
≤59	81.0	76.2	96.3	3.2:1
≤64	86.4	72.9	98.3	2.7:1
≤69	89.1	71.0	98.6	2.4:1
≤74	89.9	70.4	98.6	2.4:1
≤79	90.5	69.9	98.6	2.3:1
≤84	90.5	69.9	98.6	2.3:1
≤89	90.5	69.9	98.6	2.3:1
≤94	90.5	69.9	98.6	2.3:1
≤100	100.0	64.1	100.0	1.8:1

Tables for
the 200% of the National Poverty Line

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	100.0
15–19	100.0
20–24	100.0
25–29	100.0
30–34	100.0
35–39	99.9
40–44	95.6
45–49	90.5
50–54	80.5
55–59	60.2
60–64	51.0
65–69	38.8
70–74	31.6
75–79	18.9
80–84	18.7
85–89	18.7
90–94	18.7
95–100	18.7

Figure 7 (200% of the national line): Average differences between estimated and true poverty likelihoods for households in 1,000 bootstraps of n = 16,384, with confidence intervals by score range, scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+0.0	0.0	0.0	0.0
5–9	+0.0	0.0	0.0	0.0
10–14	+0.0	0.0	0.0	0.0
15–19	+0.0	0.0	0.0	0.0
20–24	+0.0	0.0	0.0	0.0
25–29	+0.0	0.0	0.0	0.0
30–34	+0.0	0.0	0.0	0.0
35–39	+0.5	0.2	0.3	0.4
40–44	+0.5	0.9	1.1	1.3
45–49	–7.5	4.0	4.0	4.1
50–54	–3.0	2.3	2.5	2.9
55–59	–14.4	8.3	8.5	8.8
60–64	+2.1	3.9	4.8	6.0
65–69	–26.5	15.5	15.8	16.7
70–74	+24.3	3.6	4.5	5.9
75–79	+18.2	0.4	0.5	0.6
80–84	+18.3	0.9	1.1	1.5
85–89	+0.0	0.0	0.0	0.0
90–94	+0.0	0.0	0.0	0.0
95–100	+4.1	1.9	2.4	3.2

Figure 8 (200% of the national line): Average differences between estimated poverty rates and true values for a group at a point in time in 1,000 bootstraps of various sample sizes, with confidence intervals, scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	–1.2	50.0	63.7	88.7
4	–1.3	26.9	35.6	51.4
8	–1.9	19.0	23.6	34.0
16	–2.1	12.7	16.9	23.5
32	–2.3	9.4	11.2	15.4
64	–2.2	6.7	8.0	10.3
128	–2.2	4.6	5.6	7.2
256	–2.3	3.3	4.1	5.1
512	–2.2	2.4	2.8	3.7
1,024	–2.2	1.7	2.1	2.6
2,048	–2.2	1.3	1.4	1.9
4,096	–2.2	0.9	1.0	1.4
8,192	–2.2	0.6	0.8	1.0
16,384	–2.2	0.4	0.5	0.6

Figure 11 (200% of the national line): Shares of households by cut-off score and targeting classification, along with the hit rate and BPAC, scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.1	81.5	0.0	18.3	18.5	–99.6
≤9	0.4	81.3	0.0	18.3	18.7	–99.1
≤14	2.7	78.9	0.0	18.3	21.1	–93.3
≤19	5.0	76.7	0.0	18.3	23.3	–87.8
≤24	9.9	71.8	0.0	18.3	28.2	–75.9
≤29	16.5	65.2	0.0	18.3	34.8	–59.7
≤34	30.7	51.0	0.0	18.3	49.0	–24.9
≤39	40.7	41.0	0.5	17.9	58.6	+0.2
≤44	51.9	29.8	0.9	17.4	69.3	+28.3
≤49	62.9	18.8	1.4	17.0	79.9	+55.7
≤54	70.1	11.6	3.3	15.1	85.1	+75.6
≤59	75.8	5.8	5.2	13.1	89.0	+92.0
≤64	78.7	3.0	7.7	10.6	89.3	+90.5
≤69	79.6	2.1	9.5	8.8	88.5	+88.4
≤74	79.8	1.9	10.1	8.2	88.0	+87.6
≤79	79.8	1.9	10.7	7.6	87.4	+86.9
≤84	79.8	1.9	10.7	7.6	87.4	+86.9
≤89	79.8	1.9	10.7	7.6	87.4	+86.9
≤94	79.8	1.9	10.7	7.6	87.4	+86.9
≤100	81.7	0.0	18.3	0.0	81.7	+77.6

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

Figure 12 (200% of the national line): By score cut-off, the share of all households who are targeted (that is, score at or below the cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), scorecard applied to the validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	100.0	0.2	Only poor targeted
≤9	0.4	100.0	0.5	Only poor targeted
≤14	2.7	100.0	3.3	Only poor targeted
≤19	5.0	100.0	6.1	Only poor targeted
≤24	9.9	100.0	12.1	Only poor targeted
≤29	16.5	100.0	20.1	Only poor targeted
≤34	30.7	100.0	37.6	Only poor targeted
≤39	41.2	98.9	49.8	87.0:1
≤44	52.8	98.3	63.6	57.2:1
≤49	64.3	97.9	77.0	46.0:1
≤54	73.3	95.6	85.8	21.5:1
≤59	81.0	93.6	92.8	14.6:1
≤64	86.4	91.1	96.3	10.2:1
≤69	89.1	89.4	97.5	8.4:1
≤74	89.9	88.8	97.7	7.9:1
≤79	90.5	88.2	97.7	7.5:1
≤84	90.5	88.2	97.7	7.5:1
≤89	90.5	88.2	97.7	7.5:1
≤94	90.5	88.2	97.7	7.5:1
≤100	100.0	81.7	100.0	4.5:1

Tables for the Median Poverty Line

Figure 4 (Median line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	91.5
5–9	88.6
10–14	59.9
15–19	55.3
20–24	39.5
25–29	30.5
30–34	20.2
35–39	12.6
40–44	6.2
45–49	5.4
50–54	1.2
55–59	0.7
60–64	0.4
65–69	0.4
70–74	0.4
75–79	0.4
80–84	0.4
85–89	0.4
90–94	0.4
95–100	0.4

Figure 7 (Median line): Average differences between estimated and true poverty likelihoods for households in 1,000 bootstraps of n = 16,384, with confidence intervals by score range, scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+0.8	0.9	1.1	1.4
5–9	+3.7	8.2	9.5	12.2
10–14	–28.3	15.7	16.0	16.7
15–19	+33.6	3.5	4.2	5.6
20–24	+6.1	3.5	4.2	5.3
25–29	+8.7	2.5	2.9	3.7
30–34	–17.4	10.1	10.4	10.9
35–39	+6.9	0.8	1.0	1.3
40–44	+4.0	0.5	0.6	0.8
45–49	+5.2	0.2	0.2	0.3
50–54	+0.0	0.4	0.5	0.7
55–59	–2.6	1.8	1.8	1.9
60–64	–1.4	1.1	1.2	1.4
65–69	+0.4	0.0	0.0	0.0
70–74	+0.4	0.0	0.0	0.0
75–79	+0.4	0.0	0.0	0.0
80–84	+0.4	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	+0.4	0.0	0.0	0.0

Figure 8 (Median line): Average differences between estimated poverty rates and true values for a group at a point in time in 1,000 bootstraps of various sample sizes, with confidence intervals, scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	–0.1	55.5	63.0	79.2
4	+0.0	32.6	42.6	55.1
8	–0.0	28.2	35.1	45.1
16	+0.2	22.4	26.1	35.4
32	–0.5	15.5	18.8	24.0
64	–0.9	11.5	13.5	17.4
128	–1.1	8.7	10.2	12.8
256	–1.1	6.2	7.5	8.8
512	–1.2	4.4	5.1	6.8
1,024	–1.3	3.0	3.6	4.6
2,048	–1.3	2.2	2.5	3.4
4,096	–1.2	1.6	1.9	2.3
8,192	–1.2	1.1	1.3	1.8
16,384	–1.2	0.8	1.0	1.2

Figure 11 (Median line): Shares of households by cut-off score and targeting classification, along with the hit rate and BPAC, scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.1	12.9	0.0	86.3	86.5	–97.9
≤9	0.3	12.7	0.1	86.3	86.6	–94.8
≤14	2.1	10.9	0.6	85.7	87.8	–63.0
≤19	3.0	10.0	1.9	84.4	87.5	–39.0
≤24	5.1	8.0	4.7	81.6	86.7	+14.3
≤29	6.8	6.3	9.3	77.0	83.8	+28.3
≤34	10.4	2.6	19.7	66.7	77.1	–50.9
≤39	11.5	1.5	29.1	57.3	68.8	–122.9
≤44	12.2	0.8	40.0	46.3	58.5	–207.0
≤49	12.4	0.6	51.3	35.1	47.5	–293.2
≤54	12.6	0.4	60.1	26.3	38.9	–360.9
≤59	13.0	0.0	67.4	19.0	32.0	–416.8
≤64	13.0	0.0	72.8	13.6	26.6	–458.3
≤69	13.0	0.0	75.5	10.9	23.9	–478.9
≤74	13.0	0.0	76.2	10.1	23.2	–484.9
≤79	13.0	0.0	76.8	9.5	22.6	–489.5
≤84	13.0	0.0	76.9	9.5	22.5	–489.6
≤89	13.0	0.0	76.9	9.5	22.5	–489.6
≤94	13.0	0.0	76.9	9.5	22.5	–489.6
≤100	13.0	0.0	86.4	0.0	13.0	–562.5

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

Figure 12 (Median line): By score cut-off, the share of all households who are targeted (that is, score at or below the cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), scorecard applied to the validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	90.9	1.0	9.9:1
≤9	0.4	79.6	2.3	3.9:1
≤14	2.7	76.7	16.1	3.3:1
≤19	5.0	60.7	23.1	1.5:1
≤24	9.9	51.6	39.0	1.1:1
≤29	16.5	41.2	52.0	0.7:1
≤34	30.7	33.9	79.9	0.5:1
≤39	41.2	27.9	88.2	0.4:1
≤44	52.8	23.1	93.6	0.3:1
≤49	64.3	19.3	95.2	0.2:1
≤54	73.3	17.2	97.0	0.2:1
≤59	81.0	16.1	100.0	0.2:1
≤64	86.4	15.1	100.0	0.2:1
≤69	89.1	14.6	100.0	0.2:1
≤74	89.9	14.5	100.0	0.2:1
≤79	90.5	14.4	100.0	0.2:1
≤84	90.5	14.4	100.0	0.2:1
≤89	90.5	14.4	100.0	0.2:1
≤94	90.5	14.4	100.0	0.2:1
≤100	100.0	13.0	100.0	0.1:1

Tables for
the $1.25/day 2005 PPP Poverty Line

Figure 4 ($1.25/day line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	5.8
5–9	1.6
10–14	1.4
15–19	1.4
20–24	1.4
25–29	1.1
30–34	0.9
35–39	0.3
40–44	0.1
45–49	0.1
50–54	0.0
55–59	0.0
60–64	0.0
65–69	0.0
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Figure 7 ($1.25/day line): Average differences between estimated and true poverty likelihoods for households in 1,000 bootstraps of n = 16,384, with confidence intervals by score range, scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+4.2	1.5	2.0	2.6
5–9	+2.0	0.0	0.0	0.0
10–14	+0.6	0.3	0.4	0.5
15–19	+1.2	0.1	0.1	0.2
20–24	+0.9	0.2	0.3	0.3
25–29	+0.9	0.2	0.2	0.3
30–34	+0.7	0.0	0.0	0.1
35–39	+0.3	0.1	0.1	0.2
40–44	–0.2	0.1	0.2	0.2
45–49	+0.1	0.0	0.0	0.0
50–54	+0.0	0.0	0.0	0.0
55–59	+0.0	0.0	0.0	0.0
60–64	+0.0	0.0	0.0	0.0
65–69	+0.0	0.0	0.0	0.0
70–74	+0.0	0.0	0.0	0.0
75–79	+0.0	0.0	0.0	0.0
80–84	+0.0	0.0	0.0	0.0
85–89	+0.0	0.0	0.0	0.0
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Figure 8 ($1.25/day line): Average differences between estimated poverty rates and true values for a group at a point in time in 1,000 bootstraps of various sample sizes, with confidence intervals, scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	+0.2	0.7	0.7	0.7
4	+0.2	0.5	0.6	6.6
8	+0.3	0.4	2.1	3.6
16	+0.3	0.8	1.3	1.9
32	+0.3	0.6	0.8	1.2
64	+0.3	0.4	0.6	0.8
128	+0.3	0.3	0.4	0.5
256	+0.3	0.2	0.3	0.3
512	+0.4	0.1	0.2	0.2
1,024	+0.3	0.1	0.1	0.2
2,048	+0.3	0.1	0.1	0.1
4,096	+0.3	0.1	0.1	0.1
8,192	+0.3	0.0	0.0	0.1
16,384	+0.3	0.0	0.0	0.0

Figure 11 ($1.25/day line): Shares of households by cut-off score and targeting classification, along with the hit rate and BPAC, scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.0	0.4	0.1	99.4	99.5	–61.2
≤9	0.0	0.4	0.3	99.2	99.3	–0.1
≤14	0.1	0.3	2.6	97.0	97.1	–513.3
≤19	0.2	0.3	4.8	94.8	94.9	–1,033.8
≤24	0.2	0.2	9.6	89.9	90.2	–2,170.8
≤29	0.3	0.2	16.2	83.4	83.7	–3,717.2
≤34	0.3	0.1	30.4	69.2	69.6	–7,057.3
≤39	0.4	0.1	40.8	58.8	59.2	–9,518.9
≤44	0.4	0.0	52.4	47.2	47.6	–12,257.5
≤49	0.4	0.0	63.8	35.7	36.2	–14,955.3
≤54	0.4	0.0	72.9	26.7	27.1	–17,092.0
≤59	0.4	0.0	80.6	19.0	19.4	–18,902.3
≤64	0.4	0.0	86.0	13.6	14.0	–20,178.8
≤69	0.4	0.0	88.7	10.9	11.3	–20,812.9
≤74	0.4	0.0	89.5	10.1	10.5	–20,995.7
≤79	0.4	0.0	90.1	9.5	9.9	–21,136.7
≤84	0.4	0.0	90.1	9.5	9.9	–21,141.2
≤89	0.4	0.0	90.1	9.5	9.9	–21,141.2
≤94	0.4	0.0	90.1	9.5	9.9	–21,141.2
≤100	0.4	0.0	99.6	0.0	0.4	–23,381.5

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

Figure 12 ($1.25/day line): By score cut-off, the share of all households who are targeted (that is, score at or below the cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), scorecard applied to the validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	12.9	4.4	0.1:1
≤9	0.4	11.8	10.5	0.1:1
≤14	2.7	4.7	30.2	0.0:1
≤19	5.0	3.1	36.6	0.0:1
≤24	9.9	2.3	53.8	0.0:1
≤29	16.5	1.6	62.6	0.0:1
≤34	30.7	1.1	77.8	0.0:1
≤39	41.2	0.9	87.9	0.0:1
≤44	52.8	0.8	100.0	0.0:1
≤49	64.3	0.7	100.0	0.0:1
≤54	73.3	0.6	100.0	0.0:1
≤59	81.0	0.5	100.0	0.0:1
≤64	86.4	0.5	100.0	0.0:1
≤69	89.1	0.5	100.0	0.0:1
≤74	89.9	0.5	100.0	0.0:1
≤79	90.5	0.5	100.0	0.0:1
≤84	90.5	0.5	100.0	0.0:1
≤89	90.5	0.5	100.0	0.0:1
≤94	90.5	0.5	100.0	0.0:1
≤100	100.0	0.4	100.0	0.0:1

Tables for
the $2.00/day 2005 PPP Poverty Line

Figure 4 ($2.00/day line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	88.9
5–9	71.4
10–14	51.3
15–19	38.7
20–24	24.3
25–29	15.8
30–34	9.3
35–39	4.5
40–44	3.9
45–49	2.5
50–54	0.2
55–59	0.2
60–64	0.2
65–69	0.2
70–74	0.2
75–79	0.2
80–84	0.2
85–89	0.2
90–94	0.2
95–100	0.2

Figure 7 ($2.00/day line): Average differences between estimated and true poverty likelihoods for households in 1,000 bootstraps of n = 16,384, with confidence intervals by score range, scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+4.6	3.2	3.9	5.9
5–9	–4.8	12.8	15.3	20.8
10–14	–27.0	15.5	15.8	16.4
15–19	+19.7	3.5	4.0	5.3
20–24	+4.7	2.8	3.4	4.2
25–29	+5.9	1.8	2.2	2.9
30–34	+4.7	0.7	0.8	1.1
35–39	+0.2	0.9	1.0	1.4
40–44	+1.8	0.4	0.4	0.6
45–49	+2.4	0.1	0.2	0.2
50–54	–0.0	0.2	0.3	0.4
55–59	–0.5	0.4	0.5	0.5
60–64	+0.3	0.0	0.0	0.0
65–69	+0.3	0.0	0.0	0.0
70–74	+0.3	0.0	0.0	0.0
75–79	+0.3	0.0	0.0	0.0
80–84	+0.3	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	+0.3	0.0	0.0	0.0

Figure 8 ($2.00/day line): Average differences between estimated poverty rates and true values for a group at a point in time in 1,000 bootstraps of various sample sizes, with confidence intervals, scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	–0.0	45.6	65.1	74.1
4	+0.8	23.0	31.4	44.9
8	+1.2	17.7	22.1	29.8
16	+1.7	12.8	15.3	18.5
32	+1.8	9.1	10.7	14.2
64	+1.9	6.1	7.4	10.0
128	+1.8	4.3	5.2	6.7
256	+1.8	3.2	3.7	4.8
512	+1.9	2.2	2.7	3.5
1,024	+1.9	1.5	1.8	2.4
2,048	+1.9	1.0	1.3	1.8
4,096	+2.0	0.7	0.9	1.1
8,192	+2.0	0.5	0.6	0.8
16,384	+2.0	0.4	0.4	0.6

Figure 11 ($2.00/day line): Shares of households by cut-off score and targeting classification, along with the hit rate and BPAC, scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.1	7.0	0.0	92.8	92.9	–96.5
≤9	0.3	6.8	0.1	92.8	93.0	–90.8
≤14	1.6	5.5	1.1	91.7	93.3	–39.3
≤19	2.2	4.9	2.7	90.2	92.4	+1.2
≤24	3.2	3.9	6.6	86.3	89.5	+7.1
≤29	4.3	2.8	12.2	80.7	85.0	–71.0
≤34	5.7	1.5	25.0	67.9	73.5	–251.6
≤39	6.4	0.7	34.8	58.1	64.5	–388.5
≤44	6.8	0.3	46.0	46.9	53.7	–546.1
≤49	6.9	0.2	57.4	35.5	42.4	–706.1
≤54	7.0	0.1	66.3	26.5	33.5	–832.0
≤59	7.1	0.0	73.9	19.0	26.1	–938.0
≤64	7.1	0.0	79.3	13.6	20.7	–1,014.0
≤69	7.1	0.0	82.0	10.9	18.0	–1,051.8
≤74	7.1	0.0	82.8	10.1	17.2	–1,062.7
≤79	7.1	0.0	83.4	9.5	16.6	–1,071.1
≤84	7.1	0.0	83.4	9.5	16.6	–1,071.3
≤89	7.1	0.0	83.4	9.5	16.6	–1,071.3
≤94	7.1	0.0	83.4	9.5	16.6	–1,071.3
≤100	7.1	0.0	92.9	0.0	7.1	–1,204.8

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

Figure 12 ($2.00/day line): By score cut-off, the share of all households who are targeted (that is, score at or below the cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), scorecard applied to the validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	71.3	1.5	2.5:1
≤9	0.4	72.1	3.8	2.6:1
≤14	2.7	58.4	22.4	1.4:1
≤19	5.0	45.2	31.5	0.8:1
≤24	9.9	32.9	45.6	0.5:1
≤29	16.5	26.0	60.2	0.4:1
≤34	30.7	18.4	79.4	0.2:1
≤39	41.2	15.5	89.8	0.2:1
≤44	52.8	12.9	96.0	0.1:1
≤49	64.3	10.7	96.8	0.1:1
≤54	73.3	9.5	98.2	0.1:1
≤59	81.0	8.8	100.0	0.1:1
≤64	86.4	8.2	100.0	0.1:1
≤69	89.1	8.0	100.0	0.1:1
≤74	89.9	7.9	100.0	0.1:1
≤79	90.5	7.9	100.0	0.1:1
≤84	90.5	7.9	100.0	0.1:1
≤89	90.5	7.9	100.0	0.1:1
≤94	90.5	7.9	100.0	0.1:1
≤100	100.0	7.1	100.0	0.1:1

Tables for
the $2.50/day 2005 PPP Poverty Line

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	100.0
5–9	92.0
10–14	78.7
15–19	74.5
20–24	58.4
25–29	45.7
30–34	27.1
35–39	19.2
40–44	10.0
45–49	7.3
50–54	3.1
55–59	2.2
60–64	1.6
65–69	0.9
70–74	0.9
75–79	0.9
80–84	0.9
85–89	0.9
90–94	0.9
95–100	0.9

Figure 7 ($2.50/day line): Average differences between estimated and true poverty likelihoods for households in 1,000 bootstraps of n = 16,384, with confidence intervals by score range, scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+0.8	0.9	1.1	1.4
5–9	+0.0	0.0	0.0	0.0
10–14	–23.5	12.1	12.2	12.3
15–19	–5.6	4.3	4.7	5.2
20–24	+14.9	3.6	4.5	5.9
25–29	+14.1	2.8	3.4	4.5
30–34	–14.7	8.7	9.0	9.6
35–39	+9.3	1.2	1.4	1.7
40–44	+1.4	1.2	1.4	1.9
45–49	+5.8	0.5	0.6	0.7
50–54	+1.2	0.5	0.6	0.7
55–59	–1.9	1.4	1.5	1.7
60–64	+0.1	0.8	1.0	1.3
65–69	+1.0	0.0	0.0	0.0
70–74	+1.0	0.0	0.0	0.0
75–79	+1.0	0.0	0.0	0.0
80–84	+1.0	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	+1.0	0.0	0.0	0.0

Figure 8 ($2.50/day line): Average differences between estimated poverty rates and true values for a group at a point in time in 1,000 bootstraps of various sample sizes, with confidence intervals, scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	–0.3	63.2	74.6	86.5
4	+0.5	35.4	43.1	58.3
8	+0.1	27.7	35.4	42.7
16	–0.3	21.3	24.5	32.1
32	–0.7	15.0	17.4	22.8
64	–1.1	11.2	13.3	16.3
128	–1.2	8.0	9.4	12.3
256	–1.2	5.9	6.9	8.5
512	–1.3	4.0	4.8	6.5
1,024	–1.4	2.8	3.3	4.6
2,048	–1.4	2.1	2.4	3.2
4,096	–1.4	1.5	1.7	2.2
8,192	–1.4	1.0	1.2	1.6
16,384	–1.4	0.8	0.9	1.2

Figure 11 ($2.50/day line): Shares of households by cut-off score and targeting classification, along with the hit rate and BPAC, scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.1	18.8	0.0	81.0	81.2	–98.5
≤9	0.4	18.6	0.0	81.0	81.4	–96.1
≤14	2.5	16.5	0.3	80.8	83.2	–72.6
≤19	4.1	14.8	0.8	80.2	84.4	–51.9
≤24	6.8	12.2	3.1	78.0	84.8	–12.2
≤29	9.3	9.7	7.2	73.9	83.1	+35.7
≤34	14.5	4.5	16.2	64.8	79.3	+14.3
≤39	16.2	2.7	24.9	56.1	72.4	–31.6
≤44	17.6	1.4	35.3	45.8	63.4	–86.2
≤49	18.1	0.8	46.2	34.9	53.0	–143.8
≤54	18.5	0.5	54.9	26.2	44.6	–189.7
≤59	18.9	0.0	62.1	19.0	37.9	–227.7
≤64	18.9	0.0	67.5	13.6	32.5	–256.3
≤69	18.9	0.0	70.2	10.9	29.8	–270.5
≤74	18.9	0.0	70.9	10.1	29.1	–274.6
≤79	18.9	0.0	71.5	9.5	28.5	–277.7
≤84	18.9	0.0	71.6	9.5	28.4	–277.8
≤89	18.9	0.0	71.6	9.5	28.4	–277.8
≤94	18.9	0.0	71.6	9.5	28.4	–277.8
≤100	18.9	0.0	81.1	0.0	18.9	–328.0

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

Figure 12 ($2.50/day line): By score cut-off, the share of all households who are targeted (that is, score at or below the cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), scorecard applied to the validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	90.9	0.7	9.9:1
≤9	0.4	92.7	1.9	12.7:1
≤14	2.7	89.9	13.0	8.9:1
≤19	5.0	83.5	21.9	5.1:1
≤24	9.9	68.8	35.8	2.2:1
≤29	16.5	56.2	48.9	1.3:1
≤34	30.7	47.1	76.3	0.9:1
≤39	41.2	39.4	85.7	0.7:1
≤44	52.8	33.2	92.7	0.5:1
≤49	64.3	28.2	95.6	0.4:1
≤54	73.3	25.2	97.4	0.3:1
≤59	81.0	23.4	100.0	0.3:1
≤64	86.4	21.9	100.0	0.3:1
≤69	89.1	21.3	100.0	0.3:1
≤74	89.9	21.1	100.0	0.3:1
≤79	90.5	20.9	100.0	0.3:1
≤84	90.5	20.9	100.0	0.3:1
≤89	90.5	20.9	100.0	0.3:1
≤94	90.5	20.9	100.0	0.3:1
≤100	100.0	18.9	100.0	0.2:1

Tables for
the $5.00/day 2005 PPP Poverty Line

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	100.0
15–19	100.0
20–24	100.0
25–29	100.0
30–34	99.7
35–39	96.0
40–44	92.1
45–49	85.2
50–54	69.5
55–59	50.3
60–64	39.8
65–69	25.0
70–74	19.6
75–79	12.1
80–84	12.1
85–89	12.1
90–94	12.1
95–100	12.1

Figure 7 ($5.00/day line): Average differences between estimated and true poverty likelihoods for households in 1,000 bootstraps of n = 16,384, with confidence intervals by score range, scorecard applied to the validation sample

	Difference between estimate and true value			
		Confidence interval (±percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+0.0	0.0	0.0	0.0
5–9	+0.0	0.0	0.0	0.0
10–14	+0.0	0.0	0.0	0.0
15–19	+0.0	0.0	0.0	0.0
20–24	+0.0	0.0	0.0	0.0
25–29	+0.0	0.0	0.0	0.0
30–34	–0.1	0.1	0.1	0.1
35–39	+1.7	1.0	1.2	1.6
40–44	+1.5	1.2	1.5	2.0
45–49	+1.9	1.9	2.2	2.9
50–54	–6.2	4.1	4.2	4.5
55–59	–19.3	10.9	11.1	11.7
60–64	+1.9	3.9	4.5	6.1
65–69	–36.1	20.3	20.6	21.4
70–74	+22.2	0.3	0.4	0.5
75–79	+12.6	0.4	0.5	0.6
80–84	+11.8	0.9	1.1	1.5
85–89	+0.0	0.0	0.0	0.0
90–94	+0.0	0.0	0.0	0.0
95–100	–0.9	1.9	2.3	3.0

Figure 8 ($5.00/day line): Average differences between estimated poverty rates and true values for a group at a point in time in 1,000 bootstraps of various sample sizes, with confidence intervals, scorecard applied to the validation sample

Sample Size	Difference between estimate and true value			
		Confidence interval (±percentage points)		
n	Diff.	90-percent	95-percent	99-percent
1	–1.3	55.2	71.6	92.4
4	–1.6	32.2	39.4	59.2
8	–2.2	23.0	29.5	44.2
16	–2.4	16.1	19.9	29.4
32	–2.6	11.4	14.0	19.6
64	–2.5	7.9	9.7	13.0
128	–2.6	5.6	6.6	9.4
256	–2.6	4.3	5.1	6.6
512	–2.6	2.9	3.4	4.6
1,024	–2.5	2.1	2.5	3.0
2,048	–2.5	1.5	1.7	2.2
4,096	–2.5	1.0	1.2	1.6
8,192	–2.5	0.7	0.9	1.2
16,384	–2.5	0.5	0.6	0.8

Figure 11 ($5.00/day line): Shares of households by cut-off score and targeting classification, along with the hit rate and BPAC, scorecard applied to the validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Hit rate Inclusion + Exclusion	BPAC See text
≤4	0.1	76.2	0.0	23.7	23.8	–99.6
≤9	0.4	75.9	0.0	23.7	24.1	–99.0
≤14	2.7	73.6	0.0	23.7	26.4	–92.8
≤19	5.0	71.4	0.0	23.7	28.6	–87.0
≤24	9.9	66.5	0.0	23.7	33.5	–74.2
≤29	16.5	59.9	0.0	23.7	40.1	–56.9
≤34	30.5	45.8	0.2	23.5	53.9	–19.9
≤39	40.2	36.1	1.0	22.7	62.9	+6.6
≤44	50.8	25.6	2.1	21.6	72.4	+35.7
≤49	60.7	15.6	3.6	20.1	80.8	+63.7
≤54	67.0	9.3	6.4	17.3	84.3	+83.8
≤59	71.8	4.5	9.2	14.5	86.3	+87.9
≤64	74.0	2.3	12.4	11.3	85.3	+83.7
≤69	74.8	1.6	14.3	9.3	84.1	+81.2
≤74	74.9	1.5	15.0	8.7	83.5	+80.3
≤79	74.9	1.5	15.6	8.1	82.9	+79.5
≤84	74.9	1.5	15.6	8.0	82.9	+79.5
≤89	74.9	1.5	15.6	8.0	82.9	+79.5
≤94	74.9	1.5	15.6	8.0	82.9	+79.5
≤100	76.3	0.0	23.7	0.0	76.3	+69.0

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

Figure 12 ($5.00/day line): By score cut-off, the share of all households who are targeted (that is, score at or below the cut-off), the share of targeted households who are poor (that is, have consumption below the poverty line), the share of poor households who are targeted, and the number of poor households who are successfully targeted (inclusion) per non-poor household mistakenly targeted (leakage), scorecard applied to the validation sample

Targeting cut-off	% all HHs who are targeted	% targeted HHs who are poor	% poor HHs who are targeted	Poor HHs targeted per non-poor HH targeted
≤4	0.1	100.0	0.2	Only poor targeted
≤9	0.4	100.0	0.5	Only poor targeted
≤14	2.7	100.0	3.6	Only poor targeted
≤19	5.0	100.0	6.5	Only poor targeted
≤24	9.9	100.0	12.9	Only poor targeted
≤29	16.5	100.0	21.6	Only poor targeted
≤34	30.7	99.3	39.9	145.3:1
≤39	41.2	97.7	52.7	41.8:1
≤44	52.8	96.1	66.5	24.5:1
≤49	64.3	94.4	79.5	16.9:1
≤54	73.3	91.3	87.7	10.5:1
≤59	81.0	88.6	94.1	7.8:1
≤64	86.4	85.6	97.0	6.0:1
≤69	89.1	83.9	98.0	5.2:1
≤74	89.9	83.3	98.1	5.0:1
≤79	90.5	82.7	98.1	4.8:1
≤84	90.5	82.7	98.1	4.8:1
≤89	90.5	82.7	98.1	4.8:1
≤94	90.5	82.7	98.1	4.8:1
≤100	100.0	76.3	100.0	3.2:1