Simple Poverty Scorecard® Poverty-Assessment Tool Indonesia

Shiyuan Chen and Mark Schreiner

23 April 2009

This document and related tools are at `SimplePovertyScorecard.com`.

Abstract

The Simple Poverty Scorecard®-brand poverty-assessment tool uses ten low-cost indicators from Indonesia's 2007 National Social Economic Survey to estimate the likelihood that a household has expenditure 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 Indonesia to measure poverty rates, to track changes in poverty rates over time, and to segment clients for targeted services.

Acknowledgements

This paper was funded by the Ford Foundation via a grant to the Grameen Foundation. Data come from Indonesia's *Badan Pusat Statistik*. Thanks go to Kathleen Beegle, Nigel Bigger, Erin Connor, Frank DeGiovanni, Rob Driscoll, Mary Jo Kochendorfer, Lina Marliani, Tony Sheldon, Don Sillers, Asep Suryahadi, and Jeff Toohig. This Simple Poverty Scorecard tool was re-branded by Grameen Foundation (GF) as the 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.

Authors

Shiyuan Chen and Mark Schreiner are Senior Analyst and Director with Microfinance Risk Management, L.L.C. Mark Schreiner is also a Senior Scholar at the Center for Social Development at Washington University in Saint Louis.

Simple Poverty Scorecard® Poverty-Assessment Tool

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

Indicator	Value	Points	Score
1. How many members does the household have?	A. Six or more	0	
	B. Five	7	
	C. Four	13	
	D. Three	21	
	E. Two	26	
	F. One	37	
2. How many household members aged 5 to 18 are currently attending school?	A. Not all, or no children aged 5 to 18	0	
	B. All	3	
3. In the past week, how many household members ages 11 or older worked or had a job/work/business?	A. None	0	
	B. One or two	6	
	C. Three	7	
	D. Four or more	10	
4. What is the main source of drinking water of the household?	A. Public utilities retail, safe/unsafe well, safe/unsafe water spring, river, rain water, or other	0	
	B. Public utilities (in pipes), or drilled/pumped well	4	
	C. From manufacturing	9	
5. What type of toilet does the household have?	A. Toilet over water, hole in ground/river, no toilet, or no one uses bathroom facility	0	
	B. Flush/sitting toilet	5	
6. What is the household's main flooring material?	A. Earth/soil	0	
	B. Not earth/soil	6	
7. What is the household's main ceiling material?	A. Bamboo, other, or does not have	0	
	B. Concrete, gypsum, wood, or asbestos	4	
8. Does the household own a refrigerator?	A. No	0	
	B. Yes	12	
9. Does the household own a motorcycle?	A. No	0	
	B. Yes	9	
10. Does the household own a television?	A. No	0	
	B. Yes	5	

Score:

Simple Poverty Scorecard Poverty-Assessment Tool Indonesia

1. Introduction

Pro-poor programs in Indonesia can use the Simple Poverty Scorecard poverty-assessment tool to estimate the likelihood that a household has expenditure below a given poverty line, to measure groups' poverty rates at a point in time, to track changes in groups' poverty rates over time, and to segment clients for targeted services.

The direct approach to poverty measurement via surveys is difficult and costly, asking households about a lengthy list of expenditure items such as "What is the amount of rice expenditure of the household in the last week? What is the amount of root vegetables' expenditure of the household in the last week? . . ."

In contrast, the indirect approach via the scorecard is simple, quick, and inexpensive. It uses 10 verifiable indicators (such as "What is the main source of drinking water of the household?" or "Does the household own a refrigerator?") to get a score that is highly correlated with poverty status as measured by the exhaustive survey.

The scorecard differs from "proxy means tests" (Coady, Grosh, and Hoddinott, 2002) in that it is tailored to the capabilities and purposes not of national governments but rather of local, pro-poor organizations. The feasible poverty-measurement options for these organizations are typically subjective and relative (such as participatory

wealth ranking by skilled field workers) or blunt (such as rules based on land-ownership or housing quality). Results from these approaches are not comparable across organizations nor across countries, they may be costly, and their accuracy is unknown.

If an organization wants to know what share of its participants are below a poverty line (say, USD1.25/day at 2005 purchase-power parity for the Millennium Development Goals, or the poorest half below the national poverty line as required of USAID microenterprise partners), or if it wants to measure movement across a poverty line (for example, to report to the Microcredit Summit Campaign), then it needs an expenditure-based, objective tool with known accuracy. While expenditure surveys are costly even for governments, many small, local organizations can implement an inexpensive poverty-assessment tool that can serve for monitoring, management, and targeting.

The statistical approach here aims to be understood by non-specialists. After all, if managers are to adopt the scorecard on their own and apply it to inform their decisions, they must first trust that it works. Transparency and simplicity build trust. Getting "buy-in" matters; proxy means tests and regressions on the "determinants of poverty" have been around for three decades, but they are rarely used to inform decisions, not because they do not work, but because they are presented (when they are presented at all) as tables of regression coefficients incomprehensible to lay people (with cryptic indicator names such as "`LGHHSZ_2`", negative values, and many decimal places).

Thanks to the predictive-modeling phenomenon known as the "flat max" (discussed later), simple poverty-assessment tools can be about accurate as complex ones.

The technical approach here is also innovative in how it associates scores with poverty likelihoods, in the extent of its accuracy tests, and in how it derives sample-size formulas. Although these techniques are simple and/or standard, they have rarely or never been applied to proxy means tests.

The scorecard is based on the 2007 Indonesia National Social Economic Survey (*Survei Sosial Ekonomi Nasional*, Susenas) conducted by Indonesia's *Badan Pusat Statistik* (BPS). Indicators are selected to be:

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

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

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

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

Third, the scorecard can estimate changes in the poverty rate for a group of households between two points in time. This estimate is defined as the change in the average poverty likelihood of the households in the group over time.

The scorecard can also be used for targeting. To help managers choose a targeting cut-off, this paper reports several measures of targeting accuracy for a range of possible cut-offs.

This paper presents a single scorecard whose indicators and points are derived from the Susenas household expenditure data and Indonesia's national poverty line. Scores from this scorecard are calibrated to poverty likelihoods for six poverty lines.

The scorecard is constructed and calibrated using a sub-sample of data from households that appear in both the 2007 Susenas Core and Housing modules. Its accuracy is validated on a different sub-sample. While all three scoring estimators are unbiased when applied to the population they were derived for (that is, they match the true value on average in repeated samples from the same population from which the scorecard was built), they are—like all predictive models—biased to some extent when applied to a different population.[1]

Thus, while the indirect scoring approach is less costly than the direct survey approach, it is also biased. (The survey approach is unbiased by assumption.) There is bias because scoring must assume that the future relationship between indicators and

[1] For example, a nationally representative sample at a different point in time or a non-representative sub-group (Tarozzi and Deaton, 2007).

poverty will be the same as in the data used to build the scorecard.[2] Of course, this assumption—ubiquitous and inevitable in predictive modeling—holds only partly.

When applied to the validation sample for Indonesia, the absolute difference between scorecard estimates of groups' poverty rates and the true rates is 0.4 percentage points for the national line and 0.5 percentage points on average across all six lines. These differences are due to sampling variation and not bias; the average difference would be zero if the whole 2007 Susenas were to be repeatedly redrawn and divided into sub-samples before repeating the entire process of building and calibrating scoecards.

For sample sizes of n = 16,384, the 90-percent confidence intervals for these estimates are +/–0.6 percentage points or less. For n = 1,024, the 90-percent intervals are +/–2.2 percentage points or less.

Section 2 below describes data and poverty lines. Section 3 places the new scorecard here in the context of existing exercises for Indonesia. Sections 4 and 5 describe scorecard construction and offer practical guidelines for use. Sections 6 and 7 detail the estimation of households' poverty likelihoods and of groups' poverty rates at a point in time. Section 8 discusses estimating changes in poverty rates. Section 9 covers targeting. The final section is a summary.

[2] Differences between estimates and true values may also result from changes in the quality of data collection, from imperfect adjustment of poverty lines across time or geographic regions, or from sampling variation across expenditure surveys.

2. Data and poverty lines

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

2.1 Data

The scorecard is based on data from households that appear in both the 2007 Susenas Core and Housing modules. Households are randomly divided into three sub-samples (Figure 2):

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

2.2 Poverty rates and poverty lines

2.2.1 Rates

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

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

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

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

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

If an organization serves one person per household, however, then the household-level rate is relevant. For example, if a microfinance organization serves only one person in a household, then it might prefer to report household-level poverty rates.

Based on the 2007 Susenas, this paper reports household-level poverty rates and person-level poverty rates for Indonesia by urban/rural in each province (Figure 3). The scorecard here is constructed using household-level rates, scores are calibrated to household-level poverty likelihoods, and accuracy is measured for household-level rates.

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

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

2.2.2 Poverty lines

Indonesia's food poverty line is defined as the expenditure on a 52-item food bundle that provides 2,100 kilocalories per person per day (BPS, 2008). For 2007, the per-person, per-day urban food poverty line is IDR4,348, and the rural line is IDR3,822. Indonesia's national poverty line is defined as the food poverty line plus the "minimum required" expenditure on a 46-item non-food bundle (BPS, 2008). For 2007, the per-person, per-day urban national poverty line is IDR6,179, and the rural line is IDR4,828. BPS (2008) also reports urban/rural poverty lines for each province (Figure 14).

The scorecard here is constructed using the national line. For Indonesia as a whole, the national line implies a household-level poverty rate of 11.6 percent and a person-level poverty rate of 14.5 percent. Figure 3 reports poverty lines and poverty rates at the person- and household-level for urban/rural areas by province.

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

- National
- Food
- USAID "extreme"
- USD1.25/day 2005 PPP
- USD1.75/day 2005 PPP
- USD2.50/day 2005 PPP

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

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

- 2005 PPP exchange rate for "individual consumption expenditure by households":[3] IDR4,192.83 per USD1.00
- Average national Consumer Price Index (CPI) in 2007:[4] 150.55
- Average national CPI in 2005: 125.09

Thus, the USD1.25/day 2005 PPP line for Indonesia on average in 2007 is:[5]

$$(\text{2005 PPP exchange rate}) \cdot \text{USD1.25} \cdot \left(\frac{\text{CPI}_{2007}}{\text{CPI}_{2005}}\right) =$$

$$\left(\frac{\text{IDR4,192.83}}{\text{USD1.00}}\right) \cdot \text{USD1.25} \cdot \left(\frac{150.55}{125.09}\right) = \text{IDR6,307.76.}$$

[3] http://siteresources.worldbank.org/ICPINT/Resources/icp-final-tables.pdf, accessed 1 March 2009.

[4] http://www.bps.go.id/sector/cpi/table3.shtml, accessed 1 March 2009.

[5] Sillers (2006) provides this formula.

The USD1.75/day and USD2.50/day 2005 PPP lines are multiples of the USD1.25/day 2005 PPP line.

The lines just discussed apply to all of Indonesia. The USD 2005 PPP lines are adjusted here for regional differences in cost-of-living as reflected in the national poverty lines by urban/rural in each province (Figure 14). This is done using:

- a, index to areas (u for urban or r for rural)
- i, index to the 33 provinces
- L, a given all-Indonesia poverty line
- p_a, population proportion for urban or rural areas
- π_a, given poverty lines for urban or rural areas
- π_{ai}, given provincial poverty lines for urban and rural areas

The cost-of-living-adjusted poverty line L_{ai} for area a in province i is then:

$$L_{ai} = \frac{L \cdot \pi_a}{p_u \cdot \pi_u + p_r \cdot \pi_r} \cdot \frac{\pi_{ai}}{\pi_a} = \frac{L \cdot \pi_{ai}}{p_u \cdot \pi_u + p_r \cdot \pi_r}.$$

The all-Indonesia line L is the person-weighted average of urban and rural lines π_a. The urban/rural lines for a given province L_{ai} as a proportion of the urban and rural lines L_a is the same as π_{ai} as a proportion of π_a, reflecting the differences in cost-of-living in a given urban/rural area across provinces.

3. Context of poverty-assessment tools for Indonesia

This section reviews five existing poverty-assessment tools for Indonesia. The scorecard adds value because its scores are associated with poverty likelihoods for several absolute poverty lines based on expenditure, because it tests accuracy on data not used in construction, because it uses more recent data, and because it reports accuracy and sample-size formulas for a range of scoring purposes.

3.1 Filmer and Pritchett

Filmer and Pritchett (FP, 2001) use Principal Components Analysis to make an asset index that is assumed to be a proxy for long-term wealth/economic status.[6] Beyond FP, examples of the PCA-index approach are Gwatkin *et al.* (2000, see below), Stifel and Christiaensen (2007), Zeller *et al.* (2006), and Sahn and Stifle (2003 and 2000).

FP's goal is to relate economic status to school enrollment in India. They conclude that, compared with current expenditure, their asset index predicts enrollment better and also measures long-term wealth better.

[6] Because their indicators are so similar, the PCA-based index and the expenditure-based scorecard here probably pick up the same underlying construct (such as "permanent income", see Bollen, Glanville, and Stecklov, 2007) and rank households much the same. Research that tests how well PCA-based indices predict expenditure includes Filmer and Scott (2008), Lindelow (2006), Wagstaff and Watanabe (2003), and Montgomery *et al.* (2000).

FP's India data lacks expenditure, so to check how their asset index relates with expenditure, they build another asset index using the 1994 Indonesia Demographic and Health Survey (DHS). FP check their index against expenditure not because they propose that the index be used as a proxy for expenditure—in fact, they explicitly disavow any such claims—but rather because expenditure is the most common proxy for economic status. FP do not report the indicators in their asset index for Indonesia.

FP rank households in Indonesia's 1994 DHS twice, once based on their index and a second time based on expenditure. For each given measure, they then classify households as bottom 40 percent, middle 40 percent, or top 20 percent. They judge the coherence of the rankings by comparing how households are classified across the three groups by the index and by expenditure.

Besides not proposing their index as a proxy for expenditure, FP differ from this paper in several ways. First, their purpose is not to develop a tool that local pro-poor organizations can use; indeed, they do not report indicators or points for their Indonesia poverty-assessment tool. Rather, they seek a method that researchers can use as a proxy for economic status.

Second, FP's index—unlike poverty scores—is not linked to absolute poverty lines. While this means that their index can be built without expenditure data, it also means that it cannot be used to estimate poverty rates or changes in poverty rates. Also, indices cannot be compared across countries.

Third, while FP check their index against expenditure, they do so using the same data that was used to build the tool. Thus, they overstate targeting accuracy.

Fourth, FP test accuracy by dividing households into three large groups, while this paper reports targeting accuracy (Figures 12 and 13) for a wider range of cut-offs.

Which tool targets better? If households in the 2007 Susenas are grouped as in FP by their expenditure and their scores from the new scorecard here, 28.5 percent are in the bottom group both by expenditure and by scores (compared with 26.1 percent for FP), 22.3 percent are in the middle group by both indicators (21.1 percent for FP), while 12.4 percent are in the top group by both (11.2 percent for FP). Thus, the scorecard has slightly higher targeting accuracy.

3.2 Sumarto, Suryadarma, and Suryahadi

Sumarto, Suryadarma, and Suryahadi ("SSS", 2006) compare three methods for building poverty-assessment tools: regression on poverty status (as in this paper), regression on expenditure, and Principal Components Analysis (as in FP). They aim to see how well inexpensive-to-collect indicators can proxy for expensive-to-collect expenditure in an early-warning system that would alert the government to sudden deterioration in welfare.

Although this purpose would imply that SSS would focus on accuracy in terms of estimating changes in poverty rates, in fact they focus on accuracy in terms of targeting. Their data comes from the 1999 Susenas. For each of the three approaches,

they build urban and rural tools for both the national and food lines. Each tool includes most of the following 48 indicators:

- Ownership of durable assets:
  - Radio
  - Television
  - Jewelry
  - Bicycle or boat
  - Sewing machine
  - Refrigerator
  - Motorcycle
  - Satellite dish
  - Car
  - House
  - Land
- Characteristics of the residence:
  - Type of roof
  - Type of wall
  - Type of floor
  - Presence of electrical connection
  - Type of toilet arrangement
  - Source of drinking water (protected well or pump/other)
- Animal husbandry:
  - Chickens
  - Goats
  - Cows
  - Other animals
- Education:
  - Highest level finished by head
  - Highest level finished by spouse of the head
  - Whether all children ages 6–15 attend school
- Employment:
  - Who works:
    - Head
    - Spouse
    - Any child aged 5–16
  - Whether head works in the formal sector
  - Whether main source of household income is agriculture

- Demographics:
  - Age and age squared of the head
  - Age and age squared of the spouse of the head
  - Household size and household size squared
  - Marital status of head
  - Dependency ratio
- Non-food consumption:
  - Whether each household member has different clothes for different activities
  - Whether modern medicine is used to treat illnesses
- Food consumption:
  - Whether each household member eats at least twice a day
  - Whether in the past week, the household ate:
    - Fresh cassava (*gaplek*)
    - Dried cassava (*tiwul*)
    - Banana
    - Bread
    - Biscuit
    - Egg
    - Milk
    - Beef
- Province

In the regression on poverty status, SSS classify households as poor if their expenditure is below the poverty lines in Pradhan *et al.* (2001). They then use stepwise Probit regression—similar to the Logit here—to select indicators based on statistical significance. Estimates are in terms of poverty likelihoods, and households are targeted if their poverty likelihood exceeds the arbitrary cut-off of 50 percent.

In the regression on expenditure, least-squares is used with stepwise to select statistically significant indicators. Households are targeted if their estimated expenditure from the tool is below a given poverty line.

The PCA of SSS follows FP. Households are targeted if their index is below a cut-off, and the cut-off is set so that the percentage of households who are targeted

matches the actual poverty rate in 1999. Unlike the two regression approaches, PCA produces a relative measure of poverty.

Based on the share of households correctly targeted or correctly not targeted ("Total Accuracy", see Section 9 below), SSS conclude that the regression on expenditure is the most accurate. This may be incorrect, however, as the regression on poverty status would perform better with a cut-off other than 50 percent.

For example, the scorecard here—despite only having 10 indicators and not being segmented by urban/rural—is about as accurate for targeting as the tools in SSS. In particular, when households in the 2007 Susenas are grouped by their expenditure and their scores from the scorecard, 19.4 percent are in the bottom group both by expenditure and by scores (compared with 19.5 percent for regression on expenditure, and 14.7 percent for principal components in SSS, after combining results of their urban and rural tools), 21.9 percent are in the middle group by both indicators (21.9 percent for regression on expenditure, and 18.1 percent for PCA in SSS), while 20.7 percent are in the top group by both (20.0 percent for regression on expenditure, and 15.6 percent for principal components in SSS).

If households in the 2007 Susenas are grouped by poverty status based on the national poverty line and if their poverty status is predicted using Probit (as in SSS) or Logit (as in this paper), then targeting accuracy depends—of course—on the selected cut-off. To compare the scorecard here with the regression on expenditure in SSS, exclusion in urban areas is held fixed at about 92 percent. Then inclusion is 56.8

percent for the scorecard and 49.6 percent for SSS. If exclusion for rural areas is held fixed at about 92 percent, inclusion for the scorecard is 41.1 percent versus 45.7 percent for SSS. Thus, the regression on expenditure in SSS is more accurate than the scorecard here for rural areas but less accurate for urban areas.

To compare with the regression on poverty status (Probit) in SSS, exclusion for urban areas is held fixed at about 97 percent. Then inclusion for the scorecard is 35.2 percent, versus 35.6 percent for SSS. With exclusion for the rural areas held fixed at about 90 percent, inclusion for the scorecard is 46.6 percent, versus 52.7 percent for SSS. In this case, the Probit in SSS is more accurate.

For the final comparison with PCA, exclusion for urban areas is held fixed at about 90 percent. Then inclusion for the scorecard is 62.1 percent versus 35.3 percent for SSS. With exclusion for rural areas held fixed at about 78 percent, inclusion for the scorecard here is 68.6 percent, versus 46.3 percent for SSS. Thus, the scorecard performs better than the PCA in SSS.

Beyond accuracy, the scorecard differs from the tools in SSS in several ways. First, some indicators used by SSS are not verifiable (such as whether the household uses modern medicine when someone is ill, or whether the household ate a certain food in the past week). SSS also uses three to four times as many indicators, and some indicators require computing squares or ratios.

Second, SSS do not discuss the accuracy of estimated poverty rates or changes in poverty rates, nor do they discuss sample-size formula. And they report targeting accuracy for only two cut-offs.

Third, SSS test accuracy with the same data used to build tools. As noted in the discussion of FP, this leads to overstated accuracy.

3.3 IRIS Center

IRIS Center ("IRIS", 2007a) builds a "poverty-assessment tool" (PAT) for Indonesia based on the 2002 Susenas. USAID commissioned the PAT for use by their Indonesian microenterprise partners for reporting on their participants' poverty rates. Thus, IRIS considers only the USAID "extreme" poverty line (IDR3,628 and IDR2,711 per person per day for urban and rural at April 2002 prices).

After comparing several statistical approaches,[7] IRIS settles on quantile regression (Koenker and Hallock, 2001). Their indicators are:[8]

- Household demographics:
  - — Household size
  - — Age of the household head
- Education:
  - — Literacy of the household head
  - — Education of the household head
  - — Highest level of education of members, excluding the household head

[7] All methods have roughly the same accuracy, thanks to the "flat max".
[8] IRIS does not report the actual tool, only the questionnaire used to collect data, so their indicators may differ slightly from those listed here.

- Characteristics of the residence:
  - — Area
  - — Type of floor
  - — Source of drinking water
  - — Type of toilet arrangement
  - — Source of lighting
- Whether a stall/shop is owned or rented outside of the residence
- Whether any food aid was received in the past six months
- Whether any new sets of clothes were purchased in the past year

As in the scorecard here, many of IRIS' indicators are simple to collect and verify. Some IRIS indicators, however, are not verifiable, such as the past receipt of food aid or the past purchase of a new sets of clothes. Furthermore, IRIS does not report the tool's indicators or points.

IRIS' accuracy tests focus on the difference between the estimated poverty rate and its true value. IRIS also discusses targeting accuracy in terms of successful "hits" (*coverage* when a household truly below a poverty line is predicted to have per capita expenditure below the line, or *exclusion* when a household truly above a line is predicted to be above) versus unsuccessful "misses" (*undercoverage* when a household truly below a line is predicted to be above, or *leakage* when a household truly above a line is predicted to be below).

IRIS' preferred measure of accuracy is the Balanced Poverty Accuracy Criterion (BPAC), the criterion USAID adopted for certifying poverty-assessment tools for use by its microenterprise partners (IRIS Center, 2005). The BPAC formula is:

(Inclusion – |Undercoverage – Leakage|) x [100 ÷ (Inclusion + Undercoverage)].

A higher BPAC means more accuracy; BPAC for IRIS for the USAID "extreme" line (with a poverty rate of 7.7 percent) with the 2002 SUSENAS is 33.9. For the scorecard with the 2007 Susenas, BPAC for the USAID "extreme" line (with a poverty rate of 5.6 percent) is 9.5, while BPAC for the national line (with a poverty rate of 11.6 percent) is 17.5 (Figure 12).[9]

3.4 Gwatkin *et al.*

Like FP, Gwatkin *et al.* (2000) apply PCA to make an asset index from simple, low-cost indicators, this time using Indonesia's 1997 DHS. USAID uses this same approach in 56 countries that have a DHS (Rutstein and Johnson, 2004).

A strength of PCA-based indices is that, because they do not require expenditure data, they can be applied to a wide array of "light" surveys such as censuses, DHS, Welfare Monitoring Surveys, and Core Welfare Indicator Questionnaires. The flip side is that, without expenditure data, they can only rank households and thus provide only relative—not absolute—measures of poverty. Thus, while PCA-based indices can be used for targeting, they cannot estimate households' poverty likelihoods or groups' poverty rates.

[9] If scores are not grouped in ranges, BPAC for the national line is 37.8, while BPAC for the USAID "extreme" line is 25.7. In any case, these comparisons are imperfect because BPAC depends on the overall poverty rate and is more sensitive to small differences in accuracy as the overall poverty rate is lower.

The 15 indicators in Gwatkin *et al.* are similar to those here:

- Characteristics of the residence:
  - Type of floor
  - Type of roof
  - Type of wall
  - Presence of electricity
  - Source of drinking water
  - Type of toilet facility
- Ownership of consumer durables:
  - Radio or tape recorder
  - Television
  - Refrigerator
  - Bicycle
  - Motorcycle or motorboat
  - Car
  - Gas stove
  - Kerosene stove
  - Electric stove
- Whether household members work their own or their family's agricultural land

Gwatkin *et al.* have three basic goals for their asset index:

- Segment people by quintiles in order to see how health measures vary with socio-economic status
- Monitor (via exit surveys) how well health-service points reach the poor
- Measure coverage of services via small-scale local surveys

Of course, these last two goals are the same as the monitoring and targeting goals of this paper, and the first goal of ranking household be quintiles is akin to targeting. As here, Gwatkin *et al.* present the index in a format that could be photocopied and taken to the field, although their index cannot be computed by hand in the field because the points have four decimal places and are sometimes negative.

The central contrast between the PCA-based index and the scorecard here is that because the scorecard is linked to an absolute line, it not only can rank households

but also link them to quantitative levels of expenditure. Without being based on data that includes expenditure, the PCA index cannot do this and so cannot estimate of poverty rates. Furthermore, relative accuracy (that is, targeting accuracy) is tested more completely here than in Gwatkin *et al.*; generally, discussion of the accuracy of PCA-based indices rests on how well they produce segments that are correlated with health or education.

3.5 Suryahadi *et al.*

Suryahadi *et al.* (2005) use "poverty mapping" (Elbers, Lanjouw, and Lanjouw, 2003) to estimate poverty rates at the level of Indonesia's villages. They first construct 59 expenditure-based poverty-assessment tools (one per urban/rural by province) using only indicators found both in the 1999 Susenas and in the 2000 Population Census. The tools are then applied to the census data to estimate poverty rates for smaller areas than would be possible with only the 1999 Susenas. Finally, Suryahadi *et al.* make "poverty maps" that quickly show how estimated poverty rates vary across areas in a way that makes sense to lay people.

Poverty mapping in Suryahadi *et al.* has much in common with the scorecard here in that they both:

- Build tools with nationally representative survey data and then apply them to other data on groups that may not be nationally representative
- Use simple, verifiable indicators that are quick and inexpensive to collect
- Provide unbiased estimates
- Report standard errors for their estimates (or, equivalently, confidence intervals)
- Estimate poverty likelihoods for individual households or persons
- Estimate poverty rates for groups as averages of individual poverty likelihoods
- Seek to be useful in practice and so aim to be understood by non-specialists

The strengths of poverty mapping include that it:

- Has formally established theoretical properties
- Can be applied straightforwardly to measures of well-being beyond poverty rates
- Requires less data to construct and calibrate a tool
- Uses only indicators that appear in a census

The strengths of the scorecard include that it:

- Is simpler in terms of both construction and application
- Tests accuracy and precision empirically
- Associates poverty likelihoods with scores non-parametrically
- Reports sample-size formulas (or equivalently, standard-error formulas)

The basic difference between the two approaches is that poverty mapping seeks to help governments design pro-poor policies, while the scorecard seeks to help small, local pro-poor organizations to manage their outreach when implementing policies.[10]

[10] Poverty mapping also appears to differ in that its developers say that it is inappropriate for targeting individual households. In contrast, the scorecard supports such targeting as a legitimate, potentially useful application (Schreiner, 2008a).

For Indonesia, Suryahadi *et al.* report their indicators, but that specific volume of the report is not available on the internet. While they do report standard errors for estimated poverty rates, they do not relate those to sample sizes, so the precision of their poverty mapping cannot be compared to that of the scorecard here.

4. Scorecard construction

About 100 potential indicators are initially prepared in the areas of:

- Family composition (such as household size)
- Education (such as school attendance of children)
- Employment (such as number of household members who had a job/work/business in the past week)
- Housing (such as main flooring material)
- Ownership of durable goods (such as refrigerators and televisions)

Each indicator is first screened with the entropy-based "uncertainty coefficient" (Goodman and Kruskal, 1979) that measures how well it predicts poverty on its own. Figure 4 lists the best indicators, ranked by uncertainty coefficient. Responses for each indicator are ordered starting with those most strongly associated with poverty.

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

The scorecard itself is built using the national poverty line and Logit regression on the construction sub-sample (Figure 2). Indicator selection uses both judgment and statistics (forward stepwise, based on "c"). The first step is to use Logit to build one scorecard for each candidate indicator. Each scorecard's accuracy is taken as "c", a measure of ability to rank by poverty status (SAS Institute Inc., 2004).

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

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

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

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

The single scorecard here applies to all of Indonesia. Evidence from India and Mexico (Schreiner, 2006a and 2005a), Sri Lanka (Narayan and Yoshida, 2005), and Jamaica (Grosh and Baker, 1995) suggests that segmenting poverty-assessment tools by urban/rural does not improve accuracy much.

5. Practical guidelines for scorecard use

The main challenge of scorecard design is not to maximize accuracy but rather to improve the chances that scoring is actually used (Schreiner, 2005b). When scoring projects fail, the reason is not usually technical inaccuracy but rather the failure of an organization to decide to do what is needed to integrate scoring in its processes and to learn to use it properly (Schreiner, 2002). After all, most reasonable scorecards predict tolerably well, thanks to the empirical phenomenon known as the "flat max" (Hand, 2006; Baesens *et al.*, 2003; Lovie and Lovie, 1986; Kolesar and Showers, 1985; Stillwell, Hutton, 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 adopt it and use it properly. Of course, accuracy matters, but it is balanced against simplicity, ease-of-use, and "face validity". Programs are more likely to collect data, compute scores, and pay attention to the results if, in their view, scoring does not make a lot of "extra" work and if the whole process generally seems to make sense.

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

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

A field worker using the paper scorecard would:

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

Of course, field workers must be trained. Quality outputs depend on quality inputs. If organizations or field workers gather their own data and have an incentive to exaggerate poverty rates (for example, if funders reward them for higher poverty rates), then it is wise to do on-going quality control via data review and random audits (Matul and Kline, 2003).[11] IRIS Center (2007b) 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 easy, it is still absolutely difficult. Training and explicit definitions of terms and concepts in the scorecard is essential. For the example of Nigeria, Onwujekwe, Hanson, and Fox-Rushby (2006) find distressingly low inter-rater and test-retest correlations for indicators as seemingly simple and obvious as whether the household owns an automobile. In contrast for Mexico, Martinelli and Parker (2007) find that errors by

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

interviewers and lies by respondents have negligible effects on targeting accuracy. For now, it is unknown whether these results are universal or country-specific.

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

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

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

- Employees of the organization
- Third-party contractors

Responses, scores, and poverty likelihoods can be recorded:

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

The subjects to be scored can be:

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

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

Frequency of application can be:

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

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

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

An example set of choices are illustrated by BRAC and ASA, two microlenders in Bangladesh (each with 7 million participants) who are applying the Simple Poverty Scorecard tool for Bangladesh (Schreiner, 2006b). Their design is that loan officers in a random sample of branches will score all participants each time they visit a homestead (about once a year) as part of their standard due diligence prior to loan disbursement. Responses are recorded on paper in the field before being sent to a central office to be entered into a database. ASA's and BRAC's sampling plans cover 50,000–100,000 participants each.

6. Estimates of household poverty likelihoods

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

To get absolute units, scores must be converted to *poverty likelihoods*, that is, probabilities of being below a poverty line. This is done via simple look-up tables. For the example of the national line, scores of 10–14 have a poverty likelihood of 56.9 percent, and scores of 40–44 have a poverty likelihood of 7.1 percent (Figure 5).

The poverty likelihood associated with a score varies by poverty line. For example, scores of 40–44 are associated with a poverty likelihood of 7.1 percent for the national line but 1.1 percent for the food line.[12]

6.1 Calibrating scores with poverty likelihoods

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

[12] Starting with Figure 5, most figures have six versions, one for each poverty line. To keep them straight, they are grouped by poverty line. Single tables that pertain to all poverty lines are placed with the tables for the national line.

For the example of the national line (Figure 6), there are 4,575 (normalized) households in the calibration sub-sample with a score of 20–24, of whom 1,638 (normalized) are below the poverty line. The estimated poverty likelihood associated with a score of 20–24 is then 35.8 percent, because 1,638 ÷ 4,575 = 35.8 percent.

To illustrate for a score of 40–44, there are 11,580 (normalized) households in the calibration sample, of whom 826 (normalized) are below the national line (Figure 6). Thus, the poverty likelihood for this score is 826 ÷ 11,580 = 7.1 percent.

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

Figure 7 shows, for all scores, the likelihood that expenditure falls in a range demarcated by two adjacent poverty lines. For example, the daily expenditure of someone with a score of 35–39 falls in the following ranges with probability:

- 2.4 percent below the food line
- 2.8 percent between the food and USAID "extreme" lines
- 8.0 percent between the USAID "extreme" and national lines
- 13.4 percent between the national and USD1.25/day 2005 PPP lines
- 39.3 percent between the USD1.25/day and USD 1.75/day 2005 PPP lines
- 27.0 percent between the USD1.75/day and USD 2.50/day 2005 PPP lines
- 7.1 percent above the USD2.50/day 2005 PPP line

Even though the scorecard is constructed partly based on judgment, the calibration process produces poverty likelihoods that are objective, that is, derived from survey data on expenditure and quantitative poverty lines. The poverty likelihoods would be objective even if indicators and/or points were selected without any data at all. In fact, objective scorecards of proven accuracy are often based only on judgment

(Fuller, 2006; Caire, 2004; Schreiner *et al.*, 2004). Of course, the scorecard here is constructed with both data and judgment. The fact that this paper acknowledges that some choices in scorecard construction—as in any statistical analysis—are informed by judgment in no way impugns the objectivity of the poverty likelihoods, as this depends on using data in score calibration, not on using data (and nothing else) in scorecard construction.

Although the points in Indonesia's scorecard are transformed coefficients from a Logit regression, scores are not converted to poverty likelihoods via the Logit formula of 2.718281828^{score} x $(1+ 2.718281828^{score})^{-1}$. This is because the Logit formula is esoteric and difficult to compute by hand. Non-specialists find it more intuitive to define the poverty likelihood as the share of households with a given score in the calibration sample who are below a poverty line. In the field, converting scores to poverty likelihoods requires no arithmetic at all, just a look-up table. This non-parametric calibration can also improve accuracy, especially with large calibration samples.

6.2 Accuracy of estimates of poverty likelihoods

As long as the relationship between indicators and poverty does not change and the scorecard is applied to households from the same population from which it was constructed, this calibration process produces unbiased estimates of poverty likelihoods. *Unbiased* means that in repeated samples from the same population, the average estimate matches the true poverty likelihood. The scorecard also produces unbiased

estimates of poverty rates at a point in time and of changes in poverty rates between two points in time.[13]

Of course, the relationship between indicators and poverty changes with time and across sub-groups within Indonesia's population, so the scorecard will generally be biased when applied after July 2007 (the end date of fieldwork for the 2007 Susenas) and/or to non-nationally representative groups.

How accurate are estimates of poverty likelihoods? To measure, the scorecard is applied to 1,000 bootstrap samples of size $n = 16{,}384$ from the validation sub-sample (Figure 2). Bootstrapping entails (Efron and Tibshirani, 1993):

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

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

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

For the national line, the average poverty likelihood across bootstrap samples for scores of 20–24 in the validation sample is too low by 1.3 percentage points (Figure 8). For scores of 10–14, the estimate is too high by 2.6 percentage points.[14]

For the validation sample, the 90-percent confidence interval for the differences for scores of 20–24 is +/–3.0 percentage points (Figure 8). This means that in 900 of 1,000 bootstraps, the difference between the estimate and the true value is between –4.3 and +1.7 percentage points (because –1.3 – 3.0 = –4.3, and –1.3 + 3.0 = +1.7). In 950 of 1,000 bootstraps (95 percent), the difference is –1.3 +/–3.6 percentage points, and in 990 of 1,000 bootstraps (99 percent), the difference is –1.3 +/–4.5 percentage points.

For most score ranges, Figure 8 shows some small differences between estimated poverty likelihoods and true values. These differences exist because the validation sub-sample is a single sample that—thanks to sampling variation—differs in distribution from the construction/calibration sub-samples and from Indonesia'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 9 below looks at targeting accuracy in detail.

[14] There are differences, in spite of the estimator's unbiasedness, because the scorecard comes from a single sample. The average difference by score would be zero if Susenas samples were repeatedly drawn from the population and split into sub-samples before repeating the entire process of scorecard building and calibration.

Of course, if estimates of groups' poverty rates are to be usefully accurate, then errors for individual households must largely cancel out. As discussed later, this is generally the case.

By construction, the scorecard here is unbiased. It may still, however, be *overfit* when applied after the July 2007 end date of Susenas fieldwork. That is, it may fit the 2007 Susenas data so closely that it captures not only some timeless patterns but also some random patterns that, due to sampling variation, show up only in the 2007 Susenas. Or the scorecard may be overfit in the sense that it becomes biased as the relationships between indicators and poverty change or when it is applied to non-nationally representative samples.

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

Most errors in individual households' likelihoods, however, cancel out in the estimates of groups' poverty rates (see later sections). Furthermore, much of the differences may come from non-scorecard sources such as changes in the relationship between indicators and poverty, sampling variation, changes in poverty lines, inconsistencies in data quality across time, and inconsistencies/imperfections in cost-of-living adjustments. 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).

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

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

To illustrate, suppose a program samples three households on Jan. 1, 2009 and that they have scores of 20, 30, and 40, corresponding to poverty likelihoods of 35.8, 18.3, and 7.1 percent (national line, Figure 5). The group's estimated poverty rate is the households' average poverty likelihood of $(35.8 + 18.3 + 7.1) \div 3 = 20.4$ percent.[15]

7.1 Accuracy of estimated poverty rates at a point in time

How accurate is this estimate? For a range of sample sizes, Figure 10 reports average differences between estimated and true poverty rates as well as precision (confidence intervals for the differences) for the scorecard applied to 1,000 bootstrap samples from the validation sample. For the national line, the scorecard's estimate is generally too low by about 0.4 percentage points; it estimates a poverty rate of 11.1 percent for the validation sample, but the true value is 11.5 percent (Figure 2). For all poverty lines, absolute differences for the validation sample are 1.1 percentage points or less, with an average of about 0.5 percentage points (Figure 9, summarizing Figure 10 across poverty lines).

[15] The group's poverty rate is *not* the poverty likelihood associated with the average score. Here, the average score is $(20 + 30 + 40) \div 3 = 30$, and the poverty likelihood associated with the average score is 18.3 percent. This is not the 20.4 percent found as the average of the three poverty likelihoods associated with each of the three scores.

As before, these differences are due to sampling variation in the validation sample and in the random division of the 2007 Susenas into three sub-samples.

In terms of precision, the 90-percent confidence interval for a group's estimated poverty rate at a point in time and $n = 16{,}384$ is 0.6 percentage points or less (Figure 9). This means that in 900 of 1,000 bootstraps of this size, the difference between the estimate and the true value is within 0.6 percentage points of the average difference. In the specific case of the national line and the validation sample, 90 percent of all samples of $n = 16{,}384$ produce estimates that differ from the true value in the range of –0.4 – 0.4 = –0.8 to –0.4 + 0.4 = 0.0 percentage points. (–0.4 is the average difference, and +/–0.4 is its 90-percent confidence interval.)

7.2 Sample-size formula for estimates of poverty rates at a point in time

How many households should an organization sample if it wants to estimate their poverty rate at a point in time for a desired confidence interval and confidence level? This practical question was first addressed in Schreiner (2008b).[16]

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

With direct measurement, the poverty rate can be estimated as the number of households observed to be below the poverty line, divided by the number of all observed households. The formula for sample size n in this textbook case is (Cochran, 1977):

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

where

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

c is the confidence interval as a proportion (for example, 0.02 for an interval of +/–2 percentage points), and

$\hat{p}$ is the expected (before measurement) proportion of households below the poverty line.

The scorecard, however, does not measure poverty directly, so this formula is not applicable. To derive a similar sample-size formula for the Indonesia scorecard, consider the scorecard applied to the validation sample. Figure 2 shows that the expected (before measurement) poverty rate $\hat{p}$ for the national line is 11.6 percent (that is, the average poverty rate in the construction and calibration sub-samples). In turn, a sample size n of 16,384 and a 90-percent confidence level correspond to a confidence interval of +/– 0.385 percentage points (Figure 10).[17] Plugging these into the direct-measurement sample-size formula (1) above gives not $n = 16{,}384$ but rather

[17] Due to rounding, Figure 10 displays 0.4, not 0.385.

$n = \left(\frac{1.64}{0.00385}\right)^2 \cdot 0.116 \cdot (1 - 0.116) = 18{,}608$. The ratio of the sample size for scoring (derived empirically) to the sample size for direct measurement (derived from theory) is 16,384 ÷ 18,608 = 0.88.

Applying the same method to $n = 8{,}192$ (confidence interval of +/–0.55 percentage points) gives $n = \left(\frac{1.64}{0.0055}\right)^2 \cdot 0.116 \cdot (1 - 0.116) = 9{,}118$. This time, the ratio of the sample size using scoring to the sample size using direct measurement is 8,192 ÷ 9,118 = 0.90. This ratio for $n = 8{,}192$ is close to that for $n = 16{,}384$. Indeed, applying this same procedure for all $n \geq 256$ in Figure 10 gives ratios that average to 0.90. This can be used to define a sample-size formula for the scorecard applied to the population in the validation sample:

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

where $\alpha = 0.90$ and z, c, and $\hat{p}$ are defined as in (1) above. It is this α that appears in Figure 9 as "α for sample size".

To illustrate the use of (2), suppose $c = 0.03095$ (confidence interval of +/–3.095 percentage points) and $z = 1.64$ (90-percent confidence). Then (2) gives $n = 0.90 \cdot \left(\frac{1.64}{0.03095}\right)^2 \cdot 0.116 \cdot (1 - 0.116) = 260$, which is close to the sample size of 256 for these parameters in Figure 10.

When the sample-size factor α is less than 1.0, it means that scoring is more precise than direct measurement. This occurs for five of six poverty lines in Figure 9.

Of course, the sample-size formulas here are specific to Indonesia, its poverty lines, its poverty rates, and this scorecard. The derivation method, however, is valid for any poverty-assessment tool following the approach in this paper.

In practice after July 2007 (the end date of fieldwork for the 2007 Susenas), an organization would select a poverty line (say, the national line), select a desired confidence level (say, 90 percent, or $z = 1.64$), select a desired confidence interval (say, +/–2.0 percentage points, or $c = 0.02$), make an assumption about $\hat{p}$ (perhaps based on a previous measurement such as the 11.6 percent national average for the 2007 Susenas in Figure 2), look up α (here, 0.90 for the national line), assume that the scorecard will still work in the future and/or for non-nationally representative subgroups,[18] and then compute the required sample size. In this illustration,

$$n = 0.90 \cdot \left(\frac{1.64}{0.02} \right)^2 \cdot 0.116 \cdot (1 - 0.116) = 621.$$

If the scorecard has already been applied to a sample n, then $\hat{p}$ is the scorecard's estimated poverty rate, and the confidence interval c is

$$+/- z \cdot \sqrt{\frac{\alpha \cdot \hat{p} \cdot (1 - \hat{p})}{n}}.$$

[18] This paper reports accuracy for the scorecard applied to the validation sample, but it cannot test accuracy for later years or for other groups. Still, performance after the 2007 Susenas will probably resemble that in the 2007 Susenas, with some deterioration as time passes.

8. Estimates of changes in group poverty rates over time

The change in a group's poverty rate between two points in time is estimated as the change in the average poverty likelihood of the households in the group. With data only from the 2007 Susenas, this paper cannot estimate changes over time, nor can it derive sample-size formula. Nevertheless, the relevant concepts are presented here because, in practice, pro-poor organizations can generate their own data and measure change through time.

8.1 Warning: Change is not impact

Scoring can estimate change. Of course, poverty could get better or worse, and scoring does not indicate what caused change. This point is often forgotten or confused, so it bears repeating: the scorecard simply estimates change, and it does not, in and of itself, indicate the reason for the change. In particular, estimating the impact of program participation requires knowing what would have happened to participants if they had not been participants (Moffitt, 1991). Knowing this requires either strong assumptions or a control group that resembles participants in all ways except participation. To belabor the point, the scorecard can help estimate program impact only if there is some way to know what would have happened in the absence of the program. And that information must come from somewhere beyond the scorecard. Even measuring simple change usually requires assuming that the population is constant over time and that program drop-outs do not differ from non-drop-outs.

8.2 Calculating estimated changes in poverty rates over time

Consider the illustration begun in the previous section. On Jan. 1, 2009, a program samples three households who score 20, 30, and 40 and so have poverty likelihoods of 35.8, 18.3, and 7.1 percent (national line, Figure 5). The group's baseline estimated poverty rate is the households' average poverty likelihood of (35.8 + 18.3 + 7.1) ÷ 3 = 20.4 percent.

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

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

By way of illustration, suppose that a year later on Jan. 1, 2010, the program samples three additional households who are in the same cohort as the three households originally sampled (or suppose that the program scores the same three original households a second time) and finds that their scores are 25, 35, and 45 (poverty likelihoods of 27.1, 13.1, and 4.3 percent, national line, Figure 5). Their average poverty likelihood at follow-up is now (27.1 + 13.1 + 4.3) ÷ 3 = 14.8 percent, an improvement of 20.4 – 14.8 = 5.6 percentage points.

This suggests that about one of eighteen participants crossed the poverty line in 2008. (This is a net figure; some people start above the line and end below it, and vice versa.) Among those who started below the line, about one in four (5.6 ÷ 20.4 = 27.5 percent) ended up above the line. Of course, the scorecard does not reveal the reasons for this change.

8.3 Accuracy for estimated change in two independent samples

With data only for 2007, it is not possible to measure the accuracy of scorecard estimates of changes in groups' poverty rates over time. In practice, of course, Indonesia's scorecard can still be applied to estimate change. The following sub-sections suggest approximate sample-size formula that may be used until there is additional data.

Under direct measurement, the textbook sample-size formula for direct estimates of changes in poverty rates in two equal-sized independent samples is:

$$n = 2 \cdot \left(\frac{z}{c}\right)^2 \cdot \hat{p} \cdot (1 - \hat{p}), \tag{3}$$

where z, c, and $\hat{p}$ are defined as in (1). Before measurement, $\hat{p}$ is assumed equal at both baseline and follow-up. n is the sample size at both baseline and follow-up.[19]

The method developed in the previous section can be used again to derive a sample-size formula for indirect measurement via the scorecard:

$$n = \alpha \cdot 2 \cdot \left(\frac{z}{c}\right)^2 \cdot \hat{p} \cdot (1 - \hat{p}). \tag{4}$$

As before, α is the average across sample sizes ≥ 256 of the ratio between the empirical sample size required by scoring for a given precision and the theoretical sample size required under direct measurement.

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

For the only countries for which this α has been measured (Peru, the Philippines, and India, see Schreiner, 2009a, 2009b, and 2008c), the average α across poverty lines is 0.77, 0.77, and 1.40, so 1.00 may be a reasonably conservative figure for Indonesia.

To illustrate the use of (4) to determine sample size for estimating changes in poverty rates across two independent samples, suppose the desired confidence level is 90 percent ($z = 1.64$), the desired confidence interval is 2 percentage points ($c = 0.02$), the poverty line is the national line, $\alpha = 1.00$, and $\hat{p} = 0.116$ (from Figure 2). Then the baseline sample size is $n = 1.00 \cdot 2 \cdot \left(\frac{1.64}{0.02}\right)^2 \cdot 0.116 \cdot (1 - 0.116) = 1{,}380$, and the follow-up sample size is also 1,380.

8.4 Accuracy for estimated change for one sample, scored twice

In general, the direct-measurement sample-size formula for this case is:[20]

$$n = \left(\frac{z}{c}\right)^2 \cdot \left[\hat{p}_{12} \cdot (1 - \hat{p}_{12}) + \hat{p}_{21} \cdot (1 - \hat{p}_{21}) + 2 \cdot \hat{p}_{12} \cdot \hat{p}_{21}\right], \qquad (5)$$

where z and c are defined as in (1), $\hat{p}_{12}$ is the expected (before measurement) share of all sampled cases that move from below the poverty line to above it, and $\hat{p}_{21}$ is the expected share of all sampled cases that move from above the line to below it.

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

How can a user set $\hat{p}_{12}$ and $\hat{p}_{21}$? Before measurement, a reasonable assumption is that the change in the poverty rate is zero. Then $\hat{p}_{12} = \hat{p}_{21} = \hat{p}_*$ and (5) becomes:

$$n = 2 \cdot \left(\frac{z}{c}\right)^2 \hat{p}_* . \tag{6}$$

Still, $\hat{p}_*$ could be anything between 0–1, so (6) is not enough to compute sample size. The estimate of $\hat{p}_*$ must be based on data available before baseline measurement.

Suppose that as in Peru (Schreiner, 2009a), the observed relationship between $\hat{p}_*$, the number of years between measurements y, and $p_{baseline} \cdot (1 - p_{baseline})$ is:

$$\hat{p}_* = -0.02 + 0.016 \cdot y + 0.47 \cdot [p_{baseline} \cdot (1 - p_{baseline})] \tag{7}$$

Of course, $p_{baseline}$ is not known before baseline measurement, but it is reasonable to use as its expected value a previously observed poverty rate. Given this, a poverty line, and the α indirect-scoring adjustment factor, then the scorecard's sample-size formula for a single sample directly measured twice for Indonesia (once in 2007 and then again y years later) is:

$$n = \alpha \cdot 2 \cdot \left(\frac{z}{c}\right)^2 \cdot \{-0.02 + 0.016 \cdot y + 0.47 \cdot [p_{2007} \cdot (1 - p_{2007})]\} . \tag{8}$$

In Peru (the only other country for which there is an estimate, Schreiner 2009a), the average α across years and poverty lines is about 0.90.

To illustrate the use of (8), suppose the desired confidence level is 90 percent (z = 1.64), the desired confidence interval is 2.0 percentage points (c = 0.02), the poverty line is the national line, and the sample will first be scored at baseline in 2008 and then

again at follow-up in 2007 ($y = 1$). The before-baseline poverty rate is 11.6 percent ($p_{2007} = 0.116$, Figure 2), and suppose $\alpha = 0.9$. Then the baseline sample size is $n = 0.90 \cdot 2 \cdot \left(\frac{1.64}{0.02}\right)^2 \cdot \left\{-0.02 + 0.016 \cdot 1 + 0.47 \cdot \left[0.116 \cdot (1 - 0.116)\right]\right\} = 535$. Of course, the same group of 535 households is scored at follow-up as well.

9. Targeting

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

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

Targeting is successful when households truly below a poverty line are targeted (*inclusion*) and when households truly above a poverty line are not targeted (*exclusion*). Of course, no scorecard is perfect, and targeting is unsuccessful when households truly below a poverty line are not targeted (*undercoverage*) or when households truly above a poverty line are targeted (*leakage*). Figure 11 depicts these four possible targeting outcomes. Targeting accuracy varies by cut-off; a higher cut-off has better inclusion (but greater leakage), while a lower cut-off has better exclusion (but higher undercoverage).

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

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

Figure 12 shows the distribution of households by targeting outcome for the scorecard applied to the validation sample. For an example cut-off of 15–19, outcomes for the national line applied to the validation sample are:

- Inclusion: 2.5 percent are below the line and correctly targeted
- Undercoverage: 9.0 percent are below the line and mistakenly not targeted
- Leakage: 2.3 percent are above the line and mistakenly targeted
- Exclusion: 86.2 percent are above the line and correctly not targeted

Increasing the cut-off to 20–24 improves inclusion and undercoverage but worsens leakage and exclusion:

- Inclusion: 4.1 percent are below the line and correctly targeted
- Undercoverage: 7.4 percent are below the line and mistakenly not targeted
- Leakage: 5.2 percent are above the line and mistakenly targeted
- Exclusion: 83.3 percent are above the line and correctly not targeted

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

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

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

- Assign benefits and costs to possible outcomes, based on its values and mission
- Tally total net benefits for each cut-off using Figure 12 for a given poverty line
- Select the cut-off with the highest total net benefit

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

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

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

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

Figure 12 shows "Total Accuracy" for all cut-offs for the Indonesia scorecard. For the national line in the validation sample, total net benefit is greatest (88.7) for a cut-off of 10–14, with about nine in ten Indonesian households correctly classified.

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

Figure 12 also reports BPAC, the criterion adopted by USAID for certifying poverty-assessment tools. The formula is:

BPAC = (Inclusion – |Undercoverage – Leakage|) x [100 ÷ (Inclusion+Undercoverage)].

As an alternative to assigning benefits and costs to targeting outcomes and then choosing a cut-off to maximize total net benefit, a program could set a cut-off to

achieve a desired poverty rate among targeted households. The third column of Figure 13 ("% targeted who are poor") shows, for the Indonesia scorecard applied to the validation sample, the expected poverty rate among households who score at or below a given cut-off. For the example of the national line, targeting households who score 39 or less would target 40.0 percent of all Indonesian households and produce a poverty rate among those targeted of 24.3 percent.

Figure 13 also reports two other measures of targeting accuracy. The first is a version of coverage ("% of poor who are targeted"). For the example of the national line and a cut-off of 39 or less, 84.6 percent of all poor households are covered.

The final targeting measure in Figure 13 is the number of successfully targeted poor households for each non-poor household mistakenly targeted (right-most column). For the national line and a cut-off of 39 or less, covering 3 poor households means leaking to 10 non-poor household (a ratio of 0.3:1).

10. Conclusion

Pro-poor programs in Indonesia can use the scorecard to segment clients for targeted services as well as to estimate:

- The likelihood that a household has expenditure 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 between two points in time

The scorecard is inexpensive to use and can be understood by non-specialists. It is designed to be practical for local pro-poor organizations who want to improve how they monitor and manage their social performance in order to speed up their participants' progress out of poverty.

The scorecard is built with a sub-sample of data from the 2007 Susenas, tested with a different sub-sample, and calibrated to six poverty lines (national, food, USAID "extreme", USD1.25/day 2005 PPP, USD1.75/day 2005 PPP, and USD2.50/day 2005 PPP).

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

When the scorecard is applied to the validation sample, the absolute difference between estimates versus true poverty rates for groups of households at a point in time is always less than 1.1 percentage points and averages—across the six poverty lines—about 0.5 percentage points. For $n = 16{,}384$ and 90-percent confidence, the precision of

these differences is +/–0.6 percentage points or less, and for $n = 1{,}024$, precision is +/–2.2 percentage points or less.

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

Although the statistical technique is innovative, and although technical accuracy is important, the design of the scorecard here focuses on transparency and ease-of-use. After all, a perfectly accurate scorecard is worthless if programs feel so daunted by its complexity or its cost that they do not even try to use it. For this reason, the scorecard is kept simple, using 10 indicators that are inexpensive to collect and that are straightforward to verify. Points are all zeros or positive integers, and scores range from 0 (most likely below a poverty line) to 100 (least likely below a poverty line). Scores are related to poverty likelihoods via simple look-up tables, and targeting cut-offs are likewise simple to apply. The design attempts to facilitate adoption by helping managers understand and trust scoring and by allowing non-specialists to generate scores quickly in the field.

In sum, the Simple Poverty Scorecard tool is a practical, objective way for pro-poor programs in Indonesia to monitor poverty rates, track changes in poverty rates over time, and target services. The same approach can be applied to any country with similar data from a national expenditure survey.

References

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

Badan Pusat Statistik. (2008) *Statistical Yearbook of Indonesia 2008*.

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`, accessed 2 March 2009.

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.

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

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

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

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.

Filmer, Deon; and Kinnon Scott. (2008) "Assessing Asset Indices", World Bank Policy Research Working Paper No. 4605, papers.ssrn.com/sol3/papers.cfm?abstract_id=1149108, accessed 1 March 2009.

Fuller, Rob. (2006) "Measuring the Poverty of Microfinance Clients in Haiti", microfinance.com/English/Papers/Scoring_Poverty_Haiti_Fuller.pdf, accessed 2 March 2009.

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

Grootaert, Christiaan; and Jeanine Braithwaite. (1998) "Poverty Correlates and Indicator-Based Targeting in Eastern Europe and the Former Soviet Union", World Bank Policy Research Working Paper No. 1942, dx.doi.org/10.1596/1813-9450-1942, retrieved 15 May 2016.

Grosh, Margaret; and Judy L. Baker. (1995) "Proxy Means Tests for Targeting Social Programs: Simulations and Speculation", World Bank LSMS Working Paper No. 118, poverty2.forumone.com/library/view/5496/, accessed 2 March 2009.

Gwatkin, Davidson R.; Rutstein, Shea; Johnson, Kiersten; Pande, Rohini; and Adam Wagstaff. (2000) "Socio-Economic Differences in Health, Nutrition, and Population in Indonesia", siteresources.worldbank.org/INTINDONESIA/Resources/Human/socio-economic.pdf, accessed 1 March 2009.

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

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

IRIS Center. (2007a) "Client Assessment Survey—Indonesia", povertytools.org/USAID_documents/Tools/Current_Tools/USAID_PAT_Indo_7-2007.xls, accessed 2 March 2009.

_____. (2007b) "Manual for the Implementation of USAID Poverty Assessment Tools", povertytools.org/training_documents/Manuals/USAID_PAT_Manual_Eng.pdf, accessed 2 March 2009.

_____. (2007c) "Introduction to Sampling for the Implementation of PATs", povertytools.org/training_documents/Sampling/Introduction_Sampling.ppt, accessed 2 March 2009.

_____. (2005) "Notes on Assessment and Improvement of Tool Accuracy", povertytools.org/other_documents/AssessingImproving_Accuracy.pdf, accessed 2 March 2009.

Johnson, Glenn. (2007) "Lesson 3: Two-Way Tables—Dependent Samples", www.stat.psu.edu/online/development/stat504/03_2way/53_2way_compare.htm, accessed 2 March 2009.

Koenker, Roger; and Kevin F. Hallock. (2001) "Quantile Regression", *Journal of Economic Perspectives*, Vol. 15, No. 4, pp. 143–156.

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

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

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

Martinelli, César; and Susan W. Parker. (2007) "Deception and Misreporting in a Social Program", ciep.itam.mx/~martinel/lies4.pdf, accessed 2 March 2009.

Matul, Michal; and Sean Kline. (2003) "Scoring Change: Prizma's Approach to Assessing Poverty", Microfinance Centre for Central and Eastern Europe and the New Independent States Spotlight Note No. 4, www.mfc.org.pl/doc/Research/ImpAct/SN/MFC_SN04_eng.pdf, accessed 2 March 2009.

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

Moffitt, Robert. (1991) "Program Evaluation with Non-experimental Data", *Evaluation Review*, Vol. 15, No. 3, pp. 291–314.

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

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

Narayan, Ambar; and Nobuo Yoshida. (2005) "Proxy Means Tests for Targeting Welfare Benefits in Sri Lanka", World Bank Report No. SASPR–7, `documents.worldbank.org/curated/en/2005/07/6209268/proxy-means-test-targeting-welfare-benefits-sri-lanka`, retrieved 5 May 2016.

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

Pradhan, Menno; Suryahadi, Asep; Sumarto, Sudarno; and Lant Pritchett. (2001) "Eating like Which 'Joneses'? An Iterative Solution to the Choice of Poverty Line Reference Group", *Review of Income and Wealth*, Vol. 47, No. 4, pp. 473–487.

Rutstein, Shea Oscar; and Kiersten Johnson. (2004) "The DHS Wealth Index", DHS Comparative Reports No. 6, `measuredhs.com/pubs/pdf/CR6/CR6.pdf`, accessed 2 March 2009.

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*.

Schreiner, Mark. (2009a) "Simple Poverty Scorecard Poverty-Assessent Tool: Peru", `SimplePovertyScorecard.com/PER_2007_ENG.pdf`, accessed 14 February 2016.

_____. (2009b) "Simple Poverty Scorecard Poverty-Assessent Tool: Philippines", `SimplePovertyScorecard.com/PHL_2002_ENG.pdf`, accessed 14 February 2016.

_____. (2008a) "Simple Poverty Scorecard Poverty-Assessent Tool: Ecuador", `SimplePovertyScorecard.com/ECU_2005_ENG.pdf`, accessed 14 February 2016.

_____. (2008b) "Simple Poverty Scorecard Poverty-Assessent Tool: Peru", `SimplePovertyScorecard.com/PER_2003_ENG.pdf`, accessed 14 February 2016.

_____. (2008c) "Simple Poverty Scorecard Poverty-Assessent Tool: India", SimplePovertyScorecard.com/IND_2005_ENG.pdf, accessed 14 February 2016.

_____. (2006a) "Is One Simple Poverty Scorecard Poverty-Assessent Tool Enough for India?", microfinance.com/English/Papers/Scoring_Poverty_India_Segments.pdf, accessed 2 March 2009.

_____. (2006b) "Simple Poverty Scorecard Poverty-Assessent Tool: Bangladesh", SimplePovertyScorecard.com/BGD_2000_ENG.pdf, accessed 14 February 2016.

_____. (2005a) "Herramienta Índice de Calificación de la Pobreza™: México", SimplePovertyScorecard.com/MEX_2002_SPA.pdf, accessed 14 February 2016.

_____. (2005b) "IRIS Questions on the Simple Poverty Scorecard Poverty-Assessent Tool", microfinance.com/English/Papers/Scoring_Poverty_Response_to_IRIS.pdf, accessed 2 March 2009.

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

_____; Matul, Michal; Pawlak, Ewa; and Sean Kline. (2004) "Poverty Scoring: Lessons from a Microlender in Bosnia-Herzegovina", microfinance.com/English/Papers/Scoring_Poverty_in_BiH_Short.pdf, accessed 2 March 2009.

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

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.

Sumarto, Sudarno; Suryadarma, Daniel; and Asep Suryahadi. (2006) "Predicting Consumption Poverty Using Non-Consumption Indicators: Experiments Using Indonesian Data", *Social Indicators Research*, Vol. 81, pp. 543–578.

Suryahadi, Asep; Widyanti, Wenefrida; Artha, Rima Prama; Perwira, Daniel; and Sudarno Sumarto. (2005) "Developing a Poverty Map for Indonesia: A Tool for Better Targeting in Poverty Reduction and Social Protection Programs, Book I: Technical Report", www.smeru.or.id/report/research/povertymapping1/buku%20satu%20poverty%20mapping.pdf, accessed 1 March 2009.

Tarozzi, Alessandro; and Angus Deaton. (2007) "Using Census and Survey Data to Estimate Poverty and Inequality for Small Areas", princeton.edu/~deaton/downloads/20080301SmallAreas_FINAL.pdf, accessed 2 March 2009.

Toohig, Jeff. (2008) "PPI Pilot Training Guide", progressoutofpoverty.org/toolkit, accessed 2 March 2009.

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

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

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

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

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

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

		% with expenditure below a poverty line					
			National	USAID	International (2005 PPP)		
Sub-sample	**Households**	**National**	**Food**	**'Extreme'**	**$1.25/day**	**$1.75/day**	**$2.50/day**
All Indonesia	68,634	11.6	3.0	5.6	20.7	48.5	75.2
Construction							
Selecting indicators and weights	22,922	11.6	3.0	5.5	20.8	48.5	75.3
Calibration							
Associating scores with likelihoods	22,970	11.6	3.0	5.5	20.9	48.6	75.3
Validation							
Measuring accuracy	22,742	11.5	3.1	5.7	20.4	48.4	74.9
Change between construction and calibration to validation (percentage points)							
		0.1	-0.1	-0.2	0.4	0.2	0.4

Source: 2007 Susenas Core and Housing modules.

Figure 3: Average poverty lines and poverty rates by region (household level)

		Poverty line (IDR/person/day) and poverty rate (%)											
				National		USAID		International (2005 PPP)					
	Line or	National		Food		'Extreme'		$1.25/day		$1.75/day		$2.50/day	
Region	rate	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural
Nanggroe Aceh Darussalam	Line	8,100	6,767	5,700	5,381	6,234	5,509	9,426	7,875	13,196	11,025	18,852	15,750
	Rate	22.3	25.2	8.0	11.4	10.8	12.5	33.1	38.8	56.9	70.1	78.9	91.0
Sumatera Utara	Line	6,752	5,092	4,752	4,030	5,694	4,462	7,858	5,926	11,001	8,296	15,716	11,851
	Rate	10.7	11.6	1.6	3.2	5.1	5.7	19.6	21.3	49.4	51.1	76.5	82.2
Sumatera Barat	Line	7,034	5,369	4,950	4,251	6,197	4,665	8,185	6,248	11,459	8,747	16,370	12,496
	Rate	6.8	11.5	1.1	3.5	3.4	5.8	15.7	19.9	39.4	50.8	65.9	78.9
Riau	Line	7,684	6,379	5,408	5,051	6,950	5,415	8,943	7,423	12,520	10,392	17,886	14,846
	Rate	7.6	9.6	1.0	3.2	3.5	4.5	13.4	17.4	38.4	46.4	66.0	81.0
Jambi	Line	7,061	4,998	4,969	3,957	6,041	4,288	8,217	5,816	11,504	8,142	16,434	11,632
	Rate	14.5	8.7	3.0	2.4	6.9	3.9	24.0	13.7	52.6	36.4	79.3	73.2
Sumatera Selatan	Line	6,744	5,300	4,746	4,196	5,569	4,867	7,849	6,168	10,989	8,635	15,698	12,336
	Rate	8.8	10.0	1.5	1.8	3.9	5.0	17.4	20.6	44.1	55.7	74.0	86.5
Bengkulu	Line	6,907	4,914	4,860	3,891	5,859	4,393	8,038	5,719	11,253	8,007	16,076	11,438
	Rate	13.2	15.4	3.1	3.5	6.3	7.2	24.3	27.7	49.7	64.3	77.1	89.3
Lampung	Line	6,178	4,788	4,348	3,791	5,223	4,099	7,190	5,572	10,066	7,801	14,380	11,144
	Rate	10.9	15.3	1.9	4.6	4.9	7.6	21.8	26.5	50.5	58.4	76.4	86.9
Kepulauan Bangka Belitung	Line	7,787	7,694	5,480	6,092	6,773	6,820	9,062	8,954	12,687	12,536	18,124	17,908
	Rate	2.2	10.6	0.3	2.1	0.8	5.3	6.8	18.5	26.6	58.1	65.2	88.9
Kepulauan Riau	Line	9,164	7,035	6,449	5,570	7,481	5,401	10,665	8,187	14,931	11,462	21,330	16,374
	Rate	9.1	23.3	1.9	13.1	4.2	10.8	13.9	31.3	30.9	59.1	56.1	92.0
DKI Jakarta	Line	8,774	5,945	6,174	4,183	7,436	4,846	10,211	6,918	14,295	9,685	20,422	13,836
	Rate	3.5	8.4	0.6	2.3	1.7	3.9	6.6	16.2	23.8	38.1	53.9	62.6
Jawa Barat	Line	5,945	4,741	4,183	3,754	4,846	4,086	6,918	5,517	9,685	7,724	13,836	11,034
	Rate	8.4	11.2	2.3	3.2	3.9	5.4	16.2	20.3	38.1	49.1	62.6	79.1

Figure 3 (cont.): Average poverty lines and poverty rates by region (household level)

		Poverty line (IDR/person/day) and poverty rate (%)											
				National		USAID		International (2005 PPP)					
	Line or	National		Food		'Extreme'		$1.25/day		$1.75/day		$2.50/day	
Region	rate	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural
Jawa Tengah	Line	5,529	4,629	3,891	3,665	4,843	4,072	6,435	5,387	9,009	7,542	12,870	10,774
	Rate	13.0	14.2	1.8	3.7	6.2	6.9	22.7	26.1	52.5	62.7	77.2	87.8
DI Yogyakarta	Line	6,603	5,140	4,647	4,070	5,672	4,496	7,685	5,982	10,759	8,375	15,370	11,964
	Rate	11.7	15.4	2.5	4.4	5.8	7.2	18.0	28.7	36.6	64.1	55.5	85.3
Jawa Timur	Line	5,475	4,613	3,853	3,653	4,728	4,009	6,372	5,369	8,921	7,517	12,744	10,738
	Rate	11.1	15.9	1.2	4.1	5.3	7.7	18.7	29.2	45.2	63.8	70.9	87.4
Banten	Line	6,194	4,632	4,359	3,667	5,089	4,108	7,208	5,390	10,091	7,546	14,416	10,780
	Rate	5.1	11.3	1.2	3.1	2.6	5.4	8.9	21.0	22.7	49.0	47.5	77.3
Bali	Line	5,890	4,865	4,145	3,852	5,104	4,404	6,854	5,661	9,596	7,925	13,708	11,322
	Rate	2.1	1.9	0.6	0.2	1.1	1.0	5.0	6.8	20.3	25.9	47.2	60.0
Nusa Tenggara Barat	Line	5,806	4,302	4,086	3,407	4,901	3,784	6,756	5,007	9,458	7,010	13,512	10,014
	Rate	21.2	11.3	3.6	2.8	10.6	5.4	31.4	22.1	56.5	55.9	77.4	82.2
Nusa Tenggara Timur	Line	6,114	3,725	4,303	2,950	4,854	3,135	7,115	4,335	9,961	6,069	14,230	8,670
	Rate	27.2	22.4	9.6	8.3	13.6	10.8	37.1	37.8	54.4	68.8	72.8	88.6
Kalimantan Barat	Line	5,465	4,386	3,846	3,473	4,987	4,015	6,360	5,104	8,904	7,146	12,720	10,208
	Rate	7.1	5.0	0.6	1.0	3.7	2.7	12.9	10.9	30.4	38.3	58.0	74.0
Kalimantan Tengah	Line	5,899	5,044	4,151	3,994	5,394	4,439	6,865	5,870	9,611	8,218	13,730	11,740
	Rate	6.0	5.1	0.3	0.9	2.8	2.4	9.7	13.3	27.8	41.3	61.4	76.3
Kalimantan Selatan	Line	6,092	4,756	4,287	3,765	4,989	4,063	7,089	5,534	9,925	7,748	14,178	11,068
	Rate	3.5	3.8	0.3	1.4	1.3	1.7	8.6	9.6	28.2	36.7	56.1	73.5
Kalimantan Timur	Line	7,876	6,207	5,542	4,914	7,066	4,997	9,166	7,223	12,832	10,112	18,332	14,446
	Rate	6.3	10.5	1.2	4.0	3.2	4.4	11.0	21.2	31.9	40.5	62.0	63.3
Sulawesi Utara	Line	5,452	4,913	3,837	3,890	4,291	4,114	6,344	5,718	8,882	8,005	12,688	11,436
	Rate	3.4	7.5	0.2	2.3	1.4	3.7	8.2	15.8	24.0	42.9	53.8	74.9

Figure 3 (cont.): Average poverty lines and poverty rates by region (household level)

		Poverty line (IDR/person/day) and poverty rate (%)											
	Line or	National		National Food		USAID 'Extreme'		International (2005 PPP)					
								$1.25/day		$1.75/day		$2.50/day	
Region	rate	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural
Sulawesi Tengah	Line	5,969	4,822	4,200	3,818	5,051	4,191	6,946	5,612	9,724	7,857	13,892	11,224
	Rate	7.6	17.9	2.2	5.4	4.0	8.8	14.7	30.9	37.1	63.1	56.7	85.1
Sulawesi Selatan	Line	4,913	3,807	3,457	3,014	4,358	3,340	5,718	4,430	8,005	6,202	11,436	8,860
	Rate	7.2	5.3	0.8	1.2	3.7	2.7	11.6	12.1	27.7	37.0	52.1	70.5
Sulawesi Tenggara	Line	4,672	4,182	3,288	3,311	3,460	3,552	5,437	4,867	7,612	6,814	10,874	9,734
	Rate	0.5	11.2	0.0	4.0	0.0	5.8	2.9	19.0	10.6	53.8	35.1	84.4
Gorontalo	Line	4,815	4,419	3,388	3,499	4,305	3,609	5,603	5,143	7,844	7,200	11,206	10,286
	Rate	8.9	20.9	1.0	8.1	4.7	9.9	13.0	35.6	33.9	63.4	56.8	85.7
Sulawesi Barat	Line	4,762	4,288	3,351	3,395	4,300	3,608	5,542	4,990	7,759	6,986	11,084	9,980
	Rate	11.5	13.0	0.0	4.4	5.2	6.3	15.6	23.4	39.6	58.4	64.6	78.9
Maluku	Line	6,741	5,607	4,744	4,440	6,207	4,430	7,845	6,525	10,983	9,135	15,690	13,050
	Rate	6.7	31.1	0.0	14.0	3.8	13.5	14.9	46.1	39.9	76.2	67.3	92.7
Maluku Utara	Line	6,322	5,047	4,449	3,997	5,479	4,197	7,357	5,874	10,300	8,224	14,714	11,748
	Rate	0.9	15.5	0.0	6.0	0.0	7.0	1.8	24.8	25.9	51.0	44.6	74.5
Papua Barat	Line	6,888	6,738	4,847	5,335	6,379	5,056	8,016	7,842	11,222	10,979	16,032	15,684
	Rate	4.0	37.8	0.0	21.7	2.8	17.8	10.2	50.0	26.7	74.0	60.2	88.8
Papua	Line	7,974	6,263	5,612	4,959	6,589	4,673	9,280	7,289	12,992	10,205	18,560	14,578
	Rate	10.0	51.2	2.0	30.0	4.8	24.6	16.8	59.5	36.0	76.3	58.4	89.3
All Indonesia	Line	6,287	4,832	4,424	3,826	5,328	4,183	7,317	5,623	10,243	7,872	14,633	11,245
	Rate	9.2	13.4	1.6	4.1	4.4	6.5	16.3	24.0	39.4	55.4	65.2	82.8

Source: 2007 Susenas Core and Housing modules.

Note: See text for details of deriving urban and rural regional poverty lines.

Figure 3: Average poverty lines and poverty rates by region (person level)

		Poverty line (IDR/person/day) and poverty rate (%)											
				National		USAID		International (2005 PPP)					
	Line or	National		Food		'Extreme'		$1.25/day		$1.75/day		$2.50/day	
Region	rate	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural
Nanggroe Aceh Darussalam	Line	8,100	6,760	5,700	5,381	6,234	5,509	9,426	7,866	13,196	11,013	18,852	15,732
	Rate	25.7	30.9	9.5	14.8	12.8	15.9	36.8	46.3	60.5	76.4	81.4	94.1
Sumatera Utara	Line	6,752	5,092	4,752	4,030	5,694	4,462	7,858	5,926	11,001	8,296	15,716	11,851
	Rate	14.0	15.3	2.3	4.5	6.9	7.8	24.1	27.6	57.0	59.0	82.0	86.8
Sumatera Barat	Line	7,034	5,369	4,950	4,251	6,197	4,665	8,185	6,248	11,459	8,747	16,370	12,496
	Rate	9.5	14.9	1.4	4.6	4.7	7.4	20.3	25.7	47.6	58.2	74.3	83.4
Riau	Line	7,684	6,379	5,408	5,051	6,950	5,415	8,943	7,423	12,520	10,392	17,886	14,846
	Rate	10.2	13.1	1.6	4.7	5.0	6.6	16.6	21.6	44.6	53.3	72.2	85.7
Jambi	Line	7,061	4,998	4,969	3,957	6,041	4,288	8,217	5,816	11,504	8,142	16,434	11,632
	Rate	17.3	11.0	3.7	3.1	8.5	5.4	28.0	16.4	57.8	40.8	83.4	77.5
Sumatera Selatan	Line	6,744	5,300	4,746	4,196	5,569	4,867	7,849	6,168	10,989	8,635	15,698	12,336
	Rate	10.1	13.3	2.0	2.3	5.0	6.7	19.9	25.8	50.0	62.2	79.1	90.7
Bengkulu	Line	6,907	4,914	4,860	3,891	5,859	4,393	8,038	5,719	11,253	8,007	16,076	11,438
	Rate	15.8	19.3	4.2	4.8	7.9	9.5	29.3	32.8	55.6	71.5	82.7	92.6
Lampung	Line	6,178	4,788	4,348	3,791	5,223	4,099	7,190	5,572	10,066	7,801	14,380	11,144
	Rate	12.6	17.9	2.0	5.5	6.2	8.9	25.0	31.0	55.2	63.4	79.7	89.6
Kepulauan Bangka Belitung	Line	7,787	7,694	5,480	6,092	6,773	6,820	9,062	8,954	12,687	12,536	18,124	17,908
	Rate	3.4	14.2	0.5	2.7	1.5	7.2	9.8	22.9	33.7	66.0	72.8	91.8
Kepulauan Riau	Line	9,164	7,035	6,449	5,570	7,481	5,401	10,665	8,187	14,931	11,462	21,330	16,374
	Rate	11.4	29.6	2.9	18.3	5.5	15.1	17.1	38.2	37.3	67.7	64.7	95.2
DKI Jakarta	Line	8,774	5,945	6,174	4,183	7,436	4,846	10,211	6,918	14,295	9,685	20,422	13,836
	Rate	4.8	10.7	0.8	3.0	2.4	5.3	9.0	19.7	28.2	42.9	58.7	67.0
Jawa Barat	Line	5,945	4,741	4,183	3,754	4,846	4,086	6,918	5,517	9,685	7,724	13,836	11,034
	Rate	10.7	14.4	3.0	4.5	5.3	7.2	19.7	24.7	42.9	55.3	67.0	83.2

Figure 3 (cont.): Average poverty lines and poverty rates by region (person level)

		Poverty line (IDR/person/day) and poverty rate (%)											
				National		USAID		International (2005 PPP)					
	Line or	National		Food		'Extreme'		$1.25/day		$1.75/day		$2.50/day	
Region	rate	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural
Jawa Tengah	Line	5,529	4,629	3,891	3,665	4,843	4,072	6,435	5,387	9,009	7,542	12,870	10,774
	Rate	16.1	17.7	2.3	5.1	8.1	8.8	27.5	31.1	58.4	68.0	81.7	90.2
DI Yogyakarta	Line	6,603	5,140	4,647	4,070	5,672	4,496	7,685	5,982	10,759	8,375	15,370	11,964
	Rate	15.7	19.7	3.7	6.1	8.0	9.9	23.6	34.5	44.7	67.7	65.6	86.3
Jawa Timur	Line	5,475	4,613	3,853	3,653	4,728	4,009	6,372	5,369	8,921	7,517	12,744	10,738
	Rate	13.2	19.3	1.5	5.3	6.6	9.6	21.8	33.8	50.1	68.3	74.8	89.6
Banten	Line	6,194	4,632	4,359	3,667	5,089	4,108	7,208	5,390	10,091	7,546	14,416	10,780
	Rate	6.9	14.7	1.6	4.0	3.4	7.2	11.8	26.6	27.1	55.8	52.8	81.0
Bali	Line	5,890	4,865	4,145	3,852	5,104	4,404	6,854	5,661	9,596	7,925	13,708	11,322
	Rate	2.5	2.3	0.7	0.2	1.2	1.1	6.0	8.2	24.1	30.6	53.4	65.6
Nusa Tenggara Barat	Line	5,806	4,302	4,086	3,407	4,901	3,784	6,756	5,007	9,458	7,010	13,512	10,014
	Rate	25.8	14.5	5.2	3.7	13.0	7.2	36.3	27.2	59.3	62.0	80.7	85.4
Nusa Tenggara Timur	Line	6,114	3,725	4,303	2,950	4,854	3,135	7,115	4,335	9,961	6,069	14,230	8,670
	Rate	31.3	27.2	11.5	10.3	15.9	13.6	43.0	43.9	62.3	75.4	80.8	92.3
Kalimantan Barat	Line	5,465	4,386	3,846	3,473	4,987	4,015	6,360	5,104	8,904	7,146	12,720	10,208
	Rate	10.2	6.7	0.9	1.3	5.0	3.5	18.4	13.7	37.2	44.2	64.9	78.9
Kalimantan Tengah	Line	5,899	5,044	4,151	3,994	5,394	4,439	6,865	5,870	9,611	8,218	13,730	11,740
	Rate	8.4	7.1	0.3	1.5	4.2	3.6	13.9	17.5	34.6	47.8	70.0	81.5
Kalimantan Selatan	Line	6,092	4,756	4,287	3,765	4,989	4,063	7,089	5,534	9,925	7,748	14,178	11,068
	Rate	5.2	5.6	0.5	2.2	2.4	2.8	11.4	12.7	34.8	43.9	62.9	78.6
Kalimantan Timur	Line	7,876	6,207	5,542	4,914	7,066	4,997	9,166	7,223	12,832	10,112	18,332	14,446
	Rate	8.7	14.3	1.7	6.3	4.3	7.0	13.9	27.0	38.0	50.1	68.6	71.6
Sulawesi Utara	Line	5,452	4,913	3,837	3,890	4,291	4,114	6,344	5,718	8,882	8,005	12,688	11,436
	Rate	5.4	10.0	0.6	3.1	2.4	5.1	11.5	20.4	30.2	49.7	60.5	79.6

Figure 3 (cont.): Average poverty lines and poverty rates by region (person level)

		Poverty line (IDR/person/day) and poverty rate (%)											
	Line or	National		National Food		USAID 'Extreme'		International (2005 PPP)					
								$1.25/day		$1.75/day		$2.50/day	
Region	rate	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural
Sulawesi Tengah	Line	5,969	4,822	4,200	3,818	5,051	4,191	6,946	5,612	9,724	7,857	13,892	11,224
	Rate	9.0	23.7	2.4	7.0	4.4	12.0	19.0	38.1	43.5	70.1	64.0	88.5
Sulawesi Selatan	Line	4,913	3,807	3,457	3,014	4,358	3,340	5,718	4,430	8,005	6,202	11,436	8,860
	Rate	9.5	6.4	1.1	1.4	4.8	3.2	15.3	14.9	34.4	42.8	60.7	76.0
Sulawesi Tenggara	Line	4,672	4,182	3,288	3,311	3,460	3,552	5,437	4,867	7,612	6,814	10,874	9,734
	Rate	0.6	14.3	0.0	4.9	0.0	7.3	3.6	23.3	13.1	60.0	39.7	88.1
Gorontalo	Line	4,815	4,419	3,388	3,499	4,305	3,609	5,603	5,143	7,844	7,200	11,206	10,286
	Rate	10.7	26.8	1.2	11.2	5.4	13.4	15.3	42.9	37.5	70.6	62.5	89.7
Sulawesi Barat	Line	4,762	4,288	3,351	3,395	4,300	3,608	5,542	4,990	7,759	6,986	11,084	9,980
	Rate	15.7	18.2	0.0	6.2	7.5	8.9	21.3	29.9	50.4	65.5	75.0	83.6
Maluku	Line	6,741	5,607	4,744	4,440	6,207	4,430	7,845	6,525	10,983	9,135	15,690	13,050
	Rate	9.7	40.2	0.0	20.6	5.3	20.0	19.7	56.5	48.1	84.0	73.1	95.7
Maluku Utara	Line	6,322	5,047	4,449	3,997	5,479	4,197	7,357	5,874	10,300	8,224	14,714	11,748
	Rate	1.0	19.1	0.0	8.0	0.0	9.3	2.7	29.5	30.8	59.6	53.6	83.0
Papua Barat	Line	6,888	6,738	4,847	5,335	6,379	5,056	8,016	7,842	11,222	10,979	16,032	15,684
	Rate	5.1	47.4	0.0	27.7	3.2	23.4	11.5	60.1	31.7	81.0	70.1	92.6
Papua	Line	7,974	6,263	5,612	4,959	6,589	4,673	9,280	7,289	12,992	10,205	18,560	14,578
	Rate	12.2	57.2	2.8	34.9	5.9	28.4	20.0	66.0	43.7	81.5	65.0	92.0
All Indonesia	Line	6,304	4,840	4,436	3,833	5,340	4,187	7,336	5,632	10,270	7,885	14,672	11,265
	Rate	11.5	16.8	2.2	5.4	5.7	8.4	19.9	28.9	44.8	61.2	70.2	86.3

Source: 2007 Susenas Core and Housing modules.

Note: See text for details of deriving urban and rural regional poverty lines.

Figure 4: Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly indicative of poverty)
771	Does any household member have a cell phone? (No; Yes)
717	What is the main fuel/energy source for cooking? (Charcoal or wood; Electricity, kerosene, or other; Gas/LPG)
643	What type of toilet does the household have? (Toilet over water, hole in ground/river, no toilet, or no one uses bathroom facility; Flush/sitting toilet)
641	How many members does the household have? (Six or more; Five; Four; Three; Two; One)
595	Does the household own a refrigerator? (No; Yes)
560	Does the household own a motorcycle? (No; Yes)
527	Where does the household dispose of waste (sewage)? (Beach/field/farmland, or other; River/lake/sea, or hole in ground; Man-made pond/rice paddy; Tanks)
513	What is the highest education level attained by any household member? (Junior high school, *Madrasah* junior high school. or lower; High school or *Madrasah*/vocational high school; Diploma from one-, two- or three-year college; Diploma from university or higher)
470	What is the household's main ceiling material? (Bamboo, other, or does not have; Concrete, gypsum, wood, or asbestos)
460	Does the household own a television? (No; Yes)
419	Does the household own a video? (No; Yes)
411	How many household members aged 5 to 18 are currently attending school? (Not all, or no children aged 5 to 18; All)
409	What is the main source of drinking water of the household? (Public utilities retail, safe/unsafe well, safe/unsafe water spring, river, rain water, or other; Public utilities (in pipes), or drilled/pumped well; From manufacturing)
359	Does the household own a stove? (No; Yes)
355	What is the household's main wall material? (Bamboo or other; Wood; Cement)
336	Does the household own a telephone? (No; Yes)

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

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly associated with poverty)
266	What is the household's main flooring material? (Earth/soil; Not earth/soil)
261	Is the household's garbage picked up by garbage men? (No; Yes)
222	What type of lighting does the household have? (Kerosene lighting, torch, or other; PLN or non-PLN (electricity company))
197	Does the household own a computer? (No; Yes)
163	What is the floor area in the house in square meters? (55 or less; 56 to 85; 86 or more)
156	Does the household own a radio/tape? (No; Yes)
149	Does the household own a car? (No; Yes)
83	What is the housing ownership/rental status (residential status) of the household? (Owned, free rent, or owned by parents/other relatives; Contract, rented, government/company housing, or other)
78	Does the household own a satellite dish? (No; Yes)
78	In the past week, how many household members ages 11 or older worked or had a job/work/business? (None; One or two; Three; Four or more)
76	Does the household throw away garbage in the river? (Yes; No)
58	Where is the household's wastewater reservoir from bathing/kitchen/washing? (Open reservoir in the yard; Out-yard reservoir, or without reservoir/directly to the drain/river; Closed reservoir in the yard)
34	Does the household turn the garbage into compost? (Yes; No)
30	Does the household pile up the garbage? (Yes; No)
25	What is the household's main roofing material? (Ceramic, thatch, or other; Concrete, shingles, zinc (corrugated iron), or asbestos)
24	What is the highest class ever attained by any household member? (Class 7 or lower; Class 8)

Source: 2007 Susenas Core and Housing modules, national poverty line.

National Poverty Line Tables

(and tables pertaining to all six poverty lines)

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

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	60.1
10–14	56.9
15–19	45.5
20–24	35.8
25–29	27.1
30–34	18.3
35–39	13.1
40–44	7.1
45–49	4.3
50–54	2.2
55–59	1.0
60–64	0.6
65–69	0.4
70–74	0.1
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Surveyed cases weighted to represent Indonesia's households.

Based on the Core and Housing modules of Susenas 2007.

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

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	1	÷	1	=	100.0
5–9	217	÷	361	=	60.1
10–14	748	÷	1,316	=	56.9
15–19	1,415	÷	3,108	=	45.5
20–24	1,638	÷	4,575	=	35.8
25–29	2,033	÷	7,491	=	27.1
30–34	1,994	÷	10,909	=	18.3
35–39	1,601	÷	12,196	=	13.1
40–44	826	÷	11,580	=	7.1
45–49	499	÷	11,566	=	4.3
50–54	236	÷	10,794	=	2.2
55–59	87	÷	8,610	=	1.0
60–64	44	÷	7,016	=	0.6
65–69	20	÷	4,847	=	0.4
70–74	3	÷	3,306	=	0.1
75–79	0	÷	1,689	=	0.0
80–84	0	÷	504	=	0.0
85–89	0	÷	79	=	0.0
90–94	0	÷	53	=	0.0
95–100	0	÷	0	=	0.0

Surveyed cases weighted to represent Indonesia's households.
Based on the Core and Housing modules of Susenas 2007.

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

	Likelihood of having expenditure in range demarcated by poverty lines per day per capita						
Score	<Food <IDR4,097	≥Food and <USAID ≥IDR4,097 and <IDR4,692	≥USAID and <National ≥IDR4,692 and <IDR5,481	≥National and <$1.25/day ≥IDR5,481 and <IDR6,378	≥$1.25/day and <$1.75/day ≥IDR6,378 and <IDR8,930	≥$1.75/day and <$2.50/day ≥IDR8,930 and <IDR12,757	≥$2.50/day ≥IDR12,757
0–4	100.0	0.0	0.0	0.0	0.0	0.0	0.0
5–9	22.0	11.8	26.3	18.7	19.3	1.9	0.0
10–14	28.5	8.8	19.6	15.9	21.3	5.9	0.0
15–19	17.6	9.3	18.6	19.1	26.6	8.4	0.4
20–24	11.2	7.8	16.7	18.1	34.0	11.0	1.1
25–29	6.7	6.1	14.3	19.2	36.4	15.0	2.2
30–34	3.4	4.1	10.8	16.2	39.9	20.6	5.1
35–39	2.4	2.8	8.0	13.4	39.3	27.0	7.1
40–44	1.1	1.6	4.4	9.3	37.0	33.2	13.4
45–49	0.5	1.2	2.5	7.1	30.5	36.6	21.5
50–54	0.4	0.4	1.4	3.9	24.5	38.1	31.4
55–59	0.1	0.3	0.6	1.8	15.0	37.7	44.5
60–64	0.1	0.2	0.4	1.1	11.7	29.1	57.6
65–69	0.0	0.0	0.4	0.6	6.1	24.0	68.9
70–74	0.0	0.0	0.1	0.2	2.9	16.0	80.8
75–79	0.0	0.0	0.0	0.0	1.3	8.4	90.3
80–84	0.0	0.0	0.0	0.0	0.5	5.2	94.3
85–89	0.0	0.0	0.0	0.0	0.0	0.0	100.0
90–94	0.0	0.0	0.0	0.0	0.0	0.0	100.0
95–100	0.0	0.0	0.0	0.0	0.0	0.0	100.0

Note: All poverty likelihoods in percentage units.
All the US dollar lines are in 2005 PPP.

Figure 8 (National poverty line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) from the validation sample, with confidence intervals

	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	+4.3	11.6	14.5	19.0
10–14	+2.6	6.0	7.1	9.9
15–19	–6.2	5.1	5.4	5.9
20–24	–1.3	3.0	3.6	4.5
25–29	–1.3	2.3	2.8	3.5
30–34	–2.4	2.1	2.2	2.8
35–39	+1.5	1.3	1.5	1.9
40–44	–1.0	1.1	1.3	1.7
45–49	+0.0	0.8	0.9	1.2
50–54	+0.3	0.6	0.7	0.9
55–59	–0.1	0.5	0.5	0.7
60–64	+0.2	0.3	0.4	0.5
65–69	+0.4	0.1	0.1	0.2
70–74	+0.1	0.0	0.0	0.0
75–79	+0.0	0.0	0.0	0.0
80–84	+0.0	0.0	0.0	0.0
85–89	+0.0	0.0	0.0	0.0
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Based on scorecard applied to the validation sample.

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

	Poverty line					
		National	USAID	International (2005 PPP)		
	National	Food	'Extreme'	$1.25/day	$1.75/day	$2.50/day
Estimate minus true value	-0.4	-0.2	-0.6	-0.5	-1.1	-0.4
Precision of difference	0.4	0.2	0.3	0.5	0.6	0.5
α for sample size	0.90	0.95	1.03	0.84	0.76	0.77

Precision is measured as 90-percent confidence intervals in units of +/− percentage points.
Differences and precision estimated from 1,000 bootstraps of size n = 16,384.
α is estimated from 1,000 bootstrap samples of n = 256, 512, 1,024, 2,048, 4,096, 8,192, and 16,384.

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

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Sample size (n)	Diff.	90-percent	95-percent	99-percent
2	–1.3	36.8	44.4	59.1
4	–0.9	24.6	30.0	39.4
8	–0.7	17.6	21.3	27.0
16	–0.5	12.2	15.1	21.0
32	–0.6	8.6	10.1	13.8
64	–0.3	6.1	7.4	9.4
128	–0.5	4.5	5.4	6.6
256	–0.5	3.1	3.7	5.2
512	–0.4	2.1	2.5	3.4
1,024	–0.4	1.6	1.9	2.4
2,048	–0.4	1.1	1.4	1.7
4,096	–0.4	0.8	1.0	1.2
8,192	–0.4	0.6	0.7	0.9
16,384	–0.4	0.4	0.5	0.6

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

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

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

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.0	11.5	0.0	88.5	88.5	–100.0
5–9	0.2	11.3	0.1	88.4	88.6	–95.0
10–14	0.9	10.6	0.8	87.8	88.7	–77.3
15–19	2.5	9.0	2.3	86.2	88.6	–37.0
20–24	4.1	7.4	5.2	83.3	87.4	+17.5
25–29	6.2	5.3	10.6	77.9	84.1	+7.3
30–34	8.3	3.1	19.4	69.1	77.4	–69.1
35–39	9.7	1.8	30.2	58.3	68.0	–163.3
40–44	10.6	0.8	40.9	47.6	58.3	–256.1
45–49	11.1	0.3	52.0	36.6	47.7	–352.4
50–54	11.3	0.1	62.5	26.0	37.3	–444.6
55–59	11.4	0.0	71.1	17.5	28.9	–518.8
60–64	11.5	0.0	78.0	10.5	22.0	–579.6
65–69	11.5	0.0	82.9	5.6	17.1	–621.7
70–74	11.5	0.0	86.2	2.3	13.8	–650.5
75–79	11.5	0.0	87.9	0.6	12.1	–665.2
80–84	11.5	0.0	88.4	0.1	11.6	–669.6
85–89	11.5	0.0	88.5	0.1	11.5	–670.3
90–94	11.5	0.0	88.5	0.0	11.5	–670.8
95–100	11.5	0.0	88.5	0.0	11.5	–670.8

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

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

Targeting cut-off	% all households who are targeted	% targeted who are poor	% of poor who are targeted	Poor households targeted per non-poor household targeted
0–4	0.0	100.0	0.0	Only poor targeted
5–9	0.4	58.9	1.9	1.4:1
10–14	1.7	55.1	8.0	1.2:1
15–19	4.8	51.3	21.4	1.1:1
20–24	9.4	44.1	36.0	0.8:1
25–29	16.9	36.9	54.1	0.6:1
30–34	27.8	30.1	72.6	0.4:1
35–39	40.0	24.3	84.6	0.3:1
40–44	51.5	20.6	92.6	0.3:1
45–49	63.1	17.7	97.0	0.2:1
50–54	73.9	15.4	98.8	0.2:1
55–59	82.5	13.9	99.7	0.2:1
60–64	89.5	12.8	100.0	0.1:1
65–69	94.4	12.2	100.0	0.1:1
70–74	97.7	11.8	100.0	0.1:1
75–79	99.4	11.6	100.0	0.1:1
80–84	99.9	11.5	100.0	0.1:1
85–89	99.9	11.5	100.0	0.1:1
90–94	100.0	11.5	100.0	0.1:1
95–100	100.0	11.5	100.0	0.1:1

Figure 14: Indonesia national poverty line for urban/rural areas in each province in 2007

	National poverty line (IDR/person/day)	
Province	Urban	Rural
Nanggroe Aceh Darussalam	8,100	6,796
Sumatera Utara	6,752	5,090
Sumatera Barat	7,034	5,369
Riau	7,684	6,379
Jambi	7,061	4,998
Sumatera Selatan	6,744	5,300
Bengkulu	6,907	4,914
Lampung	6,178	4,788
Kepulauan Bangka Belitung	7,787	7,694
Kepulauan Riau	9,164	7,035
DKI Jakarta	8,774	—
Jawa Barat	5,945	4,741
Jawa Tengah	5,529	4,629
DI Yogyakarta	6,603	5,140
Jawa Timur	5,475	4,613
Banten	6,194	4,632
Bali	5,890	4,865
Nusa Tenggara Barat	5,806	4,302
Nusa Tenggara Timur	6,114	3,725
Kalimantan Barat	5,465	4,386
Kalimantan Tengah	5,899	5,044
Kalimantan Selatan	6,092	4,756
Kalimantan Timur	7,876	6,207
Sulawesi Utara	5,452	4,913
Sulawesi Tengah	5,969	4,822
Sulawesi Selatan	4,913	3,807
Sulawesi Tenggara	4,672	4,182
Gorontalo	4,815	4,419
Sulawesi Barat	4,762	4,288
Maluku	6,741	5,607
Maluku Utara	6,322	5,047
Papua Barat	6,888	6,738
Papua	7,974	6,263
Indonesia	6,179	4,828

Source: Badan Pusat Statistik (2008)

Food Poverty Line Tables

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

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	22.0
10–14	28.5
15–19	17.6
20–24	11.2
25–29	6.7
30–34	3.4
35–39	2.4
40–44	1.1
45–49	0.5
50–54	0.4
55–59	0.1
60–64	0.1
65–69	0.0
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Surveyed cases weighted to represent Indonesia's households.

Based on the Core and Housing modules of Susenas 2007.

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

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	1	÷	1	=	100.0
5–9	79	÷	361	=	22.0
10–14	375	÷	1,316	=	28.5
15–19	548	÷	3,108	=	17.6
20–24	513	÷	4,575	=	11.2
25–29	500	÷	7,491	=	6.7
30–34	365	÷	10,909	=	3.4
35–39	291	÷	12,196	=	2.4
40–44	131	÷	11,580	=	1.1
45–49	62	÷	11,566	=	0.5
50–54	39	÷	10,794	=	0.4
55–59	9	÷	8,610	=	0.1
60–64	8	÷	7,016	=	0.1
65–69	0	÷	4,847	=	0.0
70–74	0	÷	3,306	=	0.0
75–79	0	÷	1,689	=	0.0
80–84	0	÷	504	=	0.0
85–89	0	÷	79	=	0.0
90–94	0	÷	53	=	0.0
95–100	0	÷	0	=	0.0

Surveyed cases weighted to represent Indonesia's households.
Based on the Core and Housing modules of Susenas 2007.

Figure 8 (Food line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) from the validation sample, with confidence intervals

	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	–1.1	9.4	11.1	15.0
10–14	+6.6	5.2	6.0	7.8
15–19	–1.4	3.1	3.6	4.9
20–24	–2.1	2.2	2.6	3.6
25–29	–1.8	1.6	1.7	2.0
30–34	–1.1	1.0	1.0	1.2
35–39	+0.4	0.6	0.7	0.9
40–44	+0.6	0.3	0.4	0.5
45–49	+0.1	0.3	0.3	0.4
50–54	+0.2	0.2	0.2	0.3
55–59	+0.0	0.1	0.2	0.2
60–64	+0.1	0.0	0.0	0.0
65–69	+0.0	0.0	0.0	0.0
70–74	+0.0	0.0	0.0	0.0
75–79	+0.0	0.0	0.0	0.0
80–84	+0.0	0.0	0.0	0.0
85–89	+0.0	0.0	0.0	0.0
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Based on scorecard applied to the validation sample.

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

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Sample size (n)	Diff.	90-percent	95-percent	99-percent
2	–0.0	17.8	29.9	41.4
4	–0.0	14.2	17.4	24.4
8	–0.1	9.2	11.1	17.4
16	–0.2	5.9	7.3	10.5
32	–0.1	4.4	5.7	7.3
64	–0.1	3.3	3.9	5.1
128	–0.1	2.5	2.9	3.9
256	–0.2	1.7	2.0	2.7
512	–0.2	1.2	1.5	1.8
1,024	–0.2	0.8	1.0	1.3
2,048	–0.2	0.6	0.7	1.0
4,096	–0.2	0.4	0.5	0.7
8,192	–0.2	0.3	0.4	0.5
16,384	–0.2	0.2	0.3	0.3

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

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.0	3.1	0.0	96.9	96.9	–99.9
5–9	0.1	3.0	0.3	96.7	96.8	–85.1
10–14	0.4	2.7	1.3	95.7	96.0	–32.3
15–19	1.0	2.1	3.8	93.1	94.1	–25.3
20–24	1.6	1.5	7.8	89.1	90.7	–155.2
25–29	2.2	0.9	14.7	82.3	84.5	–379.7
30–34	2.7	0.4	25.1	71.9	74.5	–721.1
35–39	2.9	0.2	37.1	59.9	62.8	–1,112.5
40–44	3.0	0.1	48.6	48.4	51.4	–1,489.2
45–49	3.0	0.0	60.1	36.9	39.9	–1,865.9
50–54	3.0	0.0	70.8	26.1	29.1	–2,218.4
55–59	3.1	0.0	79.5	17.5	20.5	–2,499.9
60–64	3.1	0.0	86.5	10.5	13.5	–2,729.4
65–69	3.1	0.0	91.3	5.6	8.7	–2,888.0
70–74	3.1	0.0	94.6	2.3	5.4	–2,996.2
75–79	3.1	0.0	96.3	0.6	3.7	–3,051.5
80–84	3.1	0.0	96.8	0.1	3.2	–3,068.0
85–89	3.1	0.0	96.9	0.1	3.1	–3,070.6
90–94	3.1	0.0	96.9	0.0	3.1	–3,072.3
95–100	3.1	0.0	96.9	0.0	3.1	–3,072.3

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

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

Targeting cut-off	% all households who are targeted	% targeted who are poor	% of poor who are targeted	Poor households targeted per non-poor household targeted
0–4	0.0	100.0	0.0	Only poor targeted
5–9	0.4	26.1	3.1	0.4:1
10–14	1.7	23.3	12.8	0.3:1
15–19	4.8	20.0	31.3	0.2:1
20–24	9.4	16.7	51.1	0.2:1
25–29	16.9	13.0	71.7	0.1:1
30–34	27.8	9.6	87.3	0.1:1
35–39	40.0	7.3	95.0	0.1:1
40–44	51.5	5.8	97.2	0.1:1
45–49	63.1	4.8	99.0	0.1:1
50–54	73.9	4.1	99.7	0.0:1
55–59	82.5	3.7	100.0	0.0:1
60–64	89.5	3.4	100.0	0.0:1
65–69	94.4	3.2	100.0	0.0:1
70–74	97.7	3.1	100.0	0.0:1
75–79	99.4	3.1	100.0	0.0:1
80–84	99.9	3.1	100.0	0.0:1
85–89	99.9	3.1	100.0	0.0:1
90–94	100.0	3.1	100.0	0.0:1
95–100	100.0	3.1	100.0	0.0:1

USAID "Extreme" Poverty Line Tables

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

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	33.8
10–14	37.3
15–19	27.0
20–24	19.1
25–29	12.8
30–34	7.5
35–39	5.2
40–44	2.7
45–49	1.8
50–54	0.8
55–59	0.4
60–64	0.3
65–69	0.0
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Surveyed cases weighted to represent Indonesia's households.

Based on the Core and Housing modules of Susenas 2007.

Figure 6 (USAID "extreme" line): Derivation of estimated poverty likelihoods associated with scores

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	1	÷	1	=	100.0
5–9	122	÷	361	=	33.8
10–14	491	÷	1,316	=	37.3
15–19	838	÷	3,108	=	27.0
20–24	872	÷	4,575	=	19.1
25–29	960	÷	7,491	=	12.8
30–34	813	÷	10,909	=	7.5
35–39	631	÷	12,196	=	5.2
40–44	314	÷	11,580	=	2.7
45–49	206	÷	11,566	=	1.8
50–54	86	÷	10,794	=	0.8
55–59	35	÷	8,610	=	0.4
60–64	20	÷	7,016	=	0.3
65–69	0	÷	4,847	=	0.0
70–74	0	÷	3,306	=	0.0
75–79	0	÷	1,689	=	0.0
80–84	0	÷	504	=	0.0
85–89	0	÷	79	=	0.0
90–94	0	÷	53	=	0.0
95–100	0	÷	0	=	0.0

Surveyed cases weighted to represent Indonesia's households.
Based on the Core and Housing modules of Susenas 2007.

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

	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	–5.5	11.5	14.1	17.6
10–14	+4.7	5.5	6.7	9.2
15–19	–4.4	4.1	4.5	5.4
20–24	–1.3	2.5	3.0	3.8
25–29	–3.6	2.7	2.9	3.4
30–34	–1.7	1.5	1.6	1.8
35–39	+0.1	0.9	1.1	1.4
40–44	–0.2	0.7	0.8	1.0
45–49	+0.2	0.5	0.6	0.8
50–54	+0.0	0.4	0.4	0.6
55–59	+0.0	0.3	0.3	0.4
60–64	+0.3	0.0	0.0	0.0
65–69	+0.0	0.0	0.0	0.0
70–74	+0.0	0.0	0.0	0.0
75–79	+0.0	0.0	0.0	0.0
80–84	+0.0	0.0	0.0	0.0
85–89	+0.0	0.0	0.0	0.0
90–94	+0.0	0.0	0.0	0.0
95–100	+0.0	0.0	0.0	0.0

Based on scorecard applied to the validation sample.

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

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Sample size (n)	Diff.	90-percent	95-percent	99-percent
2	–1.0	30.3	37.3	55.9
4	–0.6	18.4	22.2	32.7
8	–0.5	11.9	15.5	21.0
16	–0.5	8.6	10.3	15.0
32	–0.5	6.3	7.4	10.5
64	–0.4	4.5	5.5	7.1
128	–0.6	3.4	3.9	5.0
256	–0.6	2.4	3.0	3.9
512	–0.6	1.7	2.0	2.7
1,024	–0.6	1.2	1.4	1.8
2,048	–0.6	0.9	1.0	1.3
4,096	–0.6	0.6	0.7	0.9
8,192	–0.6	0.4	0.5	0.7
16,384	–0.6	0.3	0.3	0.4

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

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.0	5.7	0.0	94.3	94.3	–100.0
5–9	0.1	5.6	0.2	94.1	94.2	–91.2
10–14	0.6	5.1	1.1	93.2	93.7	–60.8
15–19	1.5	4.2	3.3	91.0	92.4	+9.5
20–24	2.4	3.3	7.0	87.3	89.7	–22.6
25–29	3.5	2.2	13.3	81.0	84.5	–133.4
30–34	4.5	1.2	23.3	71.0	75.5	–307.9
35–39	5.1	0.6	34.9	59.4	64.5	–511.6
40–44	5.4	0.3	46.1	48.2	53.6	–708.9
45–49	5.6	0.1	57.5	36.8	42.4	–908.5
50–54	5.7	0.0	68.2	26.1	31.7	–1,096.4
55–59	5.7	0.0	76.8	17.5	23.2	–1,246.7
60–64	5.7	0.0	83.8	10.5	16.2	–1,369.7
65–69	5.7	0.0	88.7	5.6	11.3	–1,454.7
70–74	5.7	0.0	92.0	2.3	8.0	–1,512.7
75–79	5.7	0.0	93.7	0.6	6.3	–1,542.3
80–84	5.7	0.0	94.2	0.1	5.8	–1,551.1
85–89	5.7	0.0	94.2	0.1	5.8	–1,552.5
90–94	5.7	0.0	94.3	0.0	5.7	–1,553.5
95–100	5.7	0.0	94.3	0.0	5.7	–1,553.5

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

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

Targeting cut-off	% all households who are targeted	% targeted who are poor	% of poor who are targeted	Poor households targeted per non-poor household targeted
0–4	0.0	100.0	0.0	Only poor targeted
5–9	0.4	39.4	2.5	0.7:1
10–14	1.7	33.2	9.8	0.5:1
15–19	4.8	30.6	25.6	0.4:1
20–24	9.4	25.3	41.6	0.3:1
25–29	16.9	21.0	62.1	0.3:1
30–34	27.8	16.2	78.8	0.2:1
35–39	40.0	12.7	89.0	0.1:1
40–44	51.5	10.5	94.7	0.1:1
45–49	63.1	8.9	98.0	0.1:1
50–54	73.9	7.7	99.4	0.1:1
55–59	82.5	6.9	100.0	0.1:1
60–64	89.5	6.4	100.0	0.1:1
65–69	94.4	6.0	100.0	0.1:1
70–74	97.7	5.8	100.0	0.1:1
75–79	99.4	5.7	100.0	0.1:1
80–84	99.9	5.7	100.0	0.1:1
85–89	99.9	5.7	100.0	0.1:1
90–94	100.0	5.7	100.0	0.1:1
95–100	100.0	5.7	100.0	0.1:1

$1.25/Day 2005 PPP Poverty Line Tables

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

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	78.8
10–14	72.8
15–19	64.6
20–24	53.9
25–29	46.4
30–34	34.4
35–39	26.6
40–44	16.4
45–49	11.4
50–54	6.1
55–59	2.8
60–64	1.7
65–69	1.1
70–74	0.3
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Surveyed cases weighted to represent Indonesia's households.

Based on the Core and Housing modules of Susenas 2007.

Figure 6 ($1.25/day 2005 PPP line): Derivation of estimated poverty likelihoods associated with scores

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	1	÷	1	=	100.0
5–9	284	÷	361	=	78.8
10–14	958	÷	1,316	=	72.8
15–19	2,008	÷	3,108	=	64.6
20–24	2,467	÷	4,575	=	53.9
25–29	3,473	÷	7,491	=	46.4
30–34	3,757	÷	10,909	=	34.4
35–39	3,239	÷	12,196	=	26.6
40–44	1,898	÷	11,580	=	16.4
45–49	1,317	÷	11,566	=	11.4
50–54	657	÷	10,794	=	6.1
55–59	244	÷	8,610	=	2.8
60–64	119	÷	7,016	=	1.7
65–69	51	÷	4,847	=	1.1
70–74	11	÷	3,306	=	0.3
75–79	0	÷	1,689	=	0.0
80–84	0	÷	504	=	0.0
85–89	0	÷	79	=	0.0
90–94	0	÷	53	=	0.0
95–100	0	÷	0	=	0.0

Surveyed cases weighted to represent Indonesia's households.
Based on the Core and Housing modules of Susenas 2007.

Figure 8 ($1.25/day 2005 PPP line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) from the validation sample, with confidence intervals

	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	–0.5	9.6	11.6	15.6
10–14	–3.6	5.3	6.0	8.1
15–19	–4.0	3.8	4.4	5.3
20–24	–2.6	3.2	3.7	5.0
25–29	–0.3	2.4	3.0	3.7
30–34	–3.0	2.5	2.7	3.4
35–39	+1.7	1.7	2.0	2.6
40–44	–1.8	1.7	1.9	2.7
45–49	+1.0	1.2	1.4	1.8
50–54	+0.5	0.9	1.1	1.4
55–59	–0.6	0.8	1.0	1.3
60–64	–0.0	0.7	0.8	1.0
65–69	+0.6	0.4	0.5	0.6
70–74	+0.1	0.3	0.3	0.4
75–79	–0.2	0.3	0.4	0.4
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

Based on scorecard applied to the validation sample.

Figure 10 ($1.25/day 2005 PPP line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Sample size (n)	Diff.	90-percent	95-percent	99-percent
2	–0.2	43.5	52.7	64.8
4	–0.4	31.3	36.9	48.1
8	–0.9	21.8	25.4	34.5
16	–0.9	15.5	18.2	23.3
32	–0.6	10.6	12.5	16.3
64	–0.5	7.3	9.1	11.7
128	–0.6	5.3	6.4	8.6
256	–0.6	3.8	4.6	6.4
512	–0.5	2.6	3.2	4.1
1,024	–0.5	1.9	2.3	2.9
2,048	–0.5	1.4	1.6	2.1
4,096	–0.4	1.0	1.1	1.5
8,192	–0.5	0.7	0.8	1.0
16,384	–0.5	0.5	0.6	0.7

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

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.0	20.4	0.0	79.6	79.6	–100.0
5–9	0.3	20.1	0.1	79.5	79.8	–96.8
10–14	1.3	19.1	0.4	79.2	80.5	–85.5
15–19	3.4	17.0	1.4	78.2	81.5	–60.1
20–24	5.9	14.5	3.5	76.1	82.0	–25.2
25–29	9.3	11.1	7.6	72.0	81.3	+28.1
30–34	13.2	7.2	14.6	65.0	78.2	+28.4
35–39	16.1	4.3	23.9	55.7	71.8	–17.0
40–44	18.1	2.3	33.4	46.2	64.3	–63.8
45–49	19.3	1.1	43.8	35.8	55.2	–114.6
50–54	19.9	0.5	54.0	25.6	45.6	–164.5
55–59	20.2	0.2	62.3	17.3	37.6	–205.3
60–64	20.4	0.0	69.2	10.4	30.8	–239.0
65–69	20.4	0.0	74.0	5.6	26.0	–262.6
70–74	20.4	0.0	77.3	2.3	22.7	–278.8
75–79	20.4	0.0	79.0	0.6	21.0	–287.1
80–84	20.4	0.0	79.5	0.1	20.5	–289.5
85–89	20.4	0.0	79.5	0.1	20.5	–289.9
90–94	20.4	0.0	79.6	0.0	20.4	–290.2
95–100	20.4	0.0	79.6	0.0	20.4	–290.2

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

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

Targeting cut-off	% all households who are targeted	% targeted who are poor	% of poor who are targeted	Poor households targeted per non-poor household targeted
0–4	0.0	100.0	0.0	Only poor targeted
5–9	0.4	79.1	1.4	3.8:1
10–14	1.7	76.8	6.3	3.3:1
15–19	4.8	70.2	16.5	2.4:1
20–24	9.4	63.0	28.9	1.7:1
25–29	16.9	55.1	45.5	1.2:1
30–34	27.8	47.4	64.5	0.9:1
35–39	40.0	40.3	78.9	0.7:1
40–44	51.5	35.1	88.8	0.5:1
45–49	63.1	30.6	94.8	0.4:1
50–54	73.9	27.0	97.7	0.4:1
55–59	82.5	24.5	99.2	0.3:1
60–64	89.5	22.7	99.8	0.3:1
65–69	94.4	21.6	100.0	0.3:1
70–74	97.7	20.9	100.0	0.3:1
75–79	99.4	20.5	100.0	0.3:1
80–84	99.9	20.4	100.0	0.3:1
85–89	99.9	20.4	100.0	0.3:1
90–94	100.0	20.4	100.0	0.3:1
95–100	100.0	20.4	100.0	0.3:1

$1.75/Day 2005 PPP Poverty Line Tables

Figure 5 ($1.75/day 2005 PPP line): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	98.1
10–14	94.1
15–19	91.2
20–24	87.9
25–29	82.8
30–34	74.3
35–39	65.9
40–44	53.4
45–49	41.8
50–54	30.5
55–59	17.8
60–64	13.4
65–69	7.1
70–74	3.2
75–79	1.3
80–84	0.5
85–89	0.0
90–94	0.0
95–100	0.0

Surveyed cases weighted to represent Indonesia's households.

Based on the Core and Housing modules of Susenas 2007.

Figure 6 ($1.75/day 2005 PPP line): Derivation of estimated poverty likelihoods associated with scores

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	1	÷	1	=	100.0
5–9	354	÷	361	=	98.1
10–14	1,238	÷	1,316	=	94.1
15–19	2,833	÷	3,108	=	91.2
20–24	4,022	÷	4,575	=	87.9
25–29	6,201	÷	7,491	=	82.8
30–34	8,107	÷	10,909	=	74.3
35–39	8,035	÷	12,196	=	65.9
40–44	6,184	÷	11,580	=	53.4
45–49	4,839	÷	11,566	=	41.8
50–54	3,296	÷	10,794	=	30.5
55–59	1,534	÷	8,610	=	17.8
60–64	937	÷	7,016	=	13.4
65–69	344	÷	4,847	=	7.1
70–74	107	÷	3,306	=	3.2
75–79	21	÷	1,689	=	1.3
80–84	2	÷	504	=	0.5
85–89	0	÷	79	=	0.0
90–94	0	÷	53	=	0.0
95–100	0	÷	0	=	0.0

Surveyed cases weighted to represent Indonesia's households.
Based on the Core and Housing modules of Susenas 2007.

Figure 8 ($1.75/day 2005 PPP line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) from the validation sample, with confidence intervals

	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	–0.8	1.5	1.7	2.6
10–14	–2.1	2.0	2.4	3.2
15–19	–2.8	2.3	2.4	2.8
20–24	+1.9	2.3	2.7	3.4
25–29	–1.9	1.8	2.1	2.7
30–34	–4.4	3.0	3.1	3.3
35–39	–0.3	1.8	2.2	2.7
40–44	–3.2	2.6	2.8	3.2
45–49	–1.4	1.9	2.3	3.0
50–54	+0.7	1.9	2.3	2.8
55–59	–0.7	1.7	2.0	2.9
60–64	+1.4	1.7	2.0	2.4
65–69	–0.7	1.6	2.0	2.6
70–74	+0.0	1.3	1.6	2.1
75–79	+0.2	1.1	1.2	1.6
80–84	+0.5	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

Based on scorecard applied to the validation sample.

Figure 10 ($1.75/day 2005 PPP line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Sample size (n)	Diff.	90-percent	95-percent	99-percent
2	+0.8	50.1	58.4	71.8
4	+0.2	35.4	43.0	51.8
8	–0.6	25.4	29.9	40.6
16	–0.8	17.5	20.7	26.5
32	–0.8	12.4	15.1	18.9
64	–0.8	8.4	10.1	13.0
128	–1.0	6.2	7.5	10.5
256	–1.0	4.3	5.2	7.3
512	–1.1	3.2	3.7	5.0
1,024	–1.1	2.2	2.7	3.6
2,048	–1.1	1.7	2.0	2.7
4,096	–1.1	1.1	1.3	1.9
8,192	–1.1	0.8	0.9	1.2
16,384	–1.1	0.6	0.7	0.9

Figure 12 ($1.75/day 2005 PPP line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to validation sample

Figure 13 ($1.75/day 2005 PPP line): Households below the poverty line and all households at a given score or at or below a given score cut-off, scorecard applied to validation sample

Targeting cut-off	% all households who are targeted	% targeted who are poor	% of poor who are targeted	Poor households targeted per non-poor household targeted
0–4	0.0	100.0	0.0	Only poor targeted
5–9	0.4	98.2	0.7	53.6:1
10–14	1.7	95.7	3.3	22.4:1
15–19	4.8	93.8	9.3	15.2:1
20–24	9.4	89.9	17.4	8.9:1
25–29	16.9	87.3	30.4	6.9:1
30–34	27.8	83.2	47.7	4.9:1
35–39	40.0	77.6	64.1	3.5:1
40–44	51.5	72.5	77.2	2.6:1
45–49	63.1	67.0	87.4	2.0:1
50–54	73.9	61.5	94.0	1.6:1
55–59	82.5	57.0	97.3	1.3:1
60–64	89.5	53.5	99.0	1.1:1
65–69	94.4	51.1	99.7	1.0:1
70–74	97.7	49.5	100.0	1.0:1
75–79	99.4	48.7	100.0	0.9:1
80–84	99.9	48.4	100.0	0.9:1
85–89	99.9	48.4	100.0	0.9:1
90–94	100.0	48.4	100.0	0.9:1
95–100	100.0	48.4	100.0	0.9:1

$2.50/Day 2005 PPP Poverty Line Tables

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

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	100.0
15–19	99.6
20–24	98.9
25–29	97.8
30–34	94.9
35–39	92.9
40–44	86.6
45–49	78.5
50–54	68.6
55–59	55.5
60–64	42.4
65–69	31.1
70–74	19.2
75–79	9.7
80–84	5.7
85–89	0.0
90–94	0.0
95–100	0.0

Surveyed cases weighted to represent Indonesia's households.

Based on the Core and Housing modules of Susenas 2007.

Figure 6 ($2.50/day 2005 PPP line): Derivation of estimated poverty likelihoods associated with scores

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	1	÷	1	=	100.0
5–9	361	÷	361	=	100.0
10–14	1,316	÷	1,316	=	100.0
15–19	3,095	÷	3,108	=	99.6
20–24	4,525	÷	4,575	=	98.9
25–29	7,325	÷	7,491	=	97.8
30–34	10,356	÷	10,909	=	94.9
35–39	11,327	÷	12,196	=	92.9
40–44	10,024	÷	11,580	=	86.6
45–49	9,075	÷	11,566	=	78.5
50–54	7,404	÷	10,794	=	68.6
55–59	4,780	÷	8,610	=	55.5
60–64	2,975	÷	7,016	=	42.4
65–69	1,507	÷	4,847	=	31.1
70–74	635	÷	3,306	=	19.2
75–79	164	÷	1,689	=	9.7
80–84	29	÷	504	=	5.7
85–89	0	÷	79	=	0.0
90–94	0	÷	53	=	0.0
95–100	0	÷	0	=	0.0

Surveyed cases weighted to represent Indonesia's households.
Based on the Core and Housing modules of Susenas 2007.

Figure 8 ($2.50/day 2005 PPP line): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) from the validation sample, with confidence intervals

	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	+0.5	1.1	1.3	1.7
10–14	+0.9	1.1	1.2	1.6
15–19	+0.1	0.4	0.6	0.7
20–24	+0.9	0.9	1.1	1.5
25–29	+0.3	0.8	0.9	1.1
30–34	–0.6	0.8	1.0	1.3
35–39	–0.3	0.9	1.2	1.5
40–44	–1.4	1.4	1.5	1.9
45–49	–2.4	1.9	2.0	2.4
50–54	–0.8	1.9	2.3	2.8
55–59	–0.6	2.2	2.6	3.5
60–64	+2.1	2.5	3.0	3.7
65–69	+1.3	2.8	3.4	4.5
70–74	–0.1	3.0	3.5	4.7
75–79	–1.9	3.3	3.9	5.0
80–84	+3.8	2.3	2.7	3.5
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

Based on scorecard applied to the validation sample.

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

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Sample size (n)	Diff.	90-percent	95-percent	99-percent
2	+0.2	44.4	53.0	67.9
4	–0.1	31.1	39.7	50.8
8	–0.5	22.2	25.5	35.5
16	–0.9	15.4	18.8	24.9
32	–0.5	10.8	13.1	17.5
64	–0.4	7.8	9.5	12.7
128	–0.6	5.7	6.9	8.9
256	–0.5	4.0	4.8	6.1
512	–0.4	2.7	3.2	4.2
1,024	–0.5	1.9	2.4	3.0
2,048	–0.4	1.4	1.6	2.1
4,096	–0.4	0.9	1.1	1.4
8,192	–0.4	0.7	0.8	1.1
16,384	–0.4	0.5	0.5	0.7

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

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: ≥ poverty line mistakenly targeted	Exclusion: ≥ poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.0	74.9	0.0	25.1	25.1	–100.0
5–9	0.4	74.6	0.0	25.1	25.4	–99.0
10–14	1.7	73.3	0.0	25.1	26.7	–95.5
15–19	4.7	70.2	0.0	25.0	29.8	–87.3
20–24	9.2	65.7	0.1	24.9	34.2	–75.2
25–29	16.5	58.4	0.3	24.7	41.3	–55.5
30–34	26.9	48.0	0.8	24.2	51.2	–27.0
35–39	38.2	36.7	1.8	23.3	61.5	+4.3
40–44	48.3	26.6	3.2	21.8	70.1	+33.2
45–49	57.5	17.4	5.6	19.5	77.0	+61.0
50–54	65.0	10.0	8.9	16.1	81.1	+85.3
55–59	69.8	5.2	12.8	12.3	82.1	+83.0
60–64	72.6	2.3	16.9	8.2	80.8	+77.4
65–69	74.1	0.9	20.3	4.8	78.8	+72.9
70–74	74.7	0.2	23.0	2.1	76.8	+69.4
75–79	74.9	0.0	24.4	0.6	75.5	+67.4
80–84	74.9	0.0	24.9	0.1	75.1	+66.7
85–89	74.9	0.0	25.0	0.1	75.0	+66.6
90–94	74.9	0.0	25.1	0.0	74.9	+66.5
95–100	74.9	0.0	25.1	0.0	74.9	+66.5

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

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

Targeting cut-off	% all households who are targeted	% targeted who are poor	% of poor who are targeted	Poor households targeted per non-poor household targeted
0–4	0.0	100.0	0.0	Only poor targeted
5–9	0.4	99.2	0.5	128.6:1
10–14	1.7	99.2	2.2	122.9:1
15–19	4.8	99.3	6.3	134.6:1
20–24	9.4	98.7	12.3	73.7:1
25–29	16.9	98.0	22.0	49.7:1
30–34	27.8	97.0	35.9	32.0:1
35–39	40.0	95.6	51.0	21.7:1
40–44	51.5	93.7	64.5	14.9:1
45–49	63.1	91.1	76.8	10.3:1
50–54	73.9	87.9	86.7	7.3:1
55–59	82.5	84.5	93.1	5.5:1
60–64	89.5	81.1	96.9	4.3:1
65–69	94.4	78.5	98.9	3.6:1
70–74	97.7	76.5	99.7	3.3:1
75–79	99.4	75.4	100.0	3.1:1
80–84	99.9	75.0	100.0	3.0:1
85–89	99.9	75.0	100.0	3.0:1
90–94	100.0	74.9	100.0	3.0:1
95–100	100.0	74.9	100.0	3.0:1