Simple Poverty Scorecard® Poverty-Assessment Tool Vietnam

Shiyuan Chen, Mark Schreiner, and Gary Woller

August 20, 2008

This document is at `SimplePovertyScorecard.com`.

Abstract

The Simple Poverty Scorecard-brand poverty-assessment tool uses 10 low-cost indicators from Vietnam's 2004 Household Living Standards Survey to estimate the likelihood that a household has consumption below a given poverty line. Field workers can collect responses in about ten minutes. The scorecard's accuracy is reported for a range of poverty lines. The scorecard is a practical way for pro-poor programs in Vietnam to measure poverty rates, to track changes in poverty rates over time, and to segment clients for differentiated treatment.

Acknowledgements

This paper was funded by the Microcredit Summit Campaign. Data are from Vietnam's General Statistical Office. Thanks to Malika Anand, Nhan Phan Cu, Frank DeGiovanni, Robert Driscoll, Paul Glewwe, Phan Ha, Phi Ha Le, Tony Sheldon, and Don Sillers.

Simple Poverty Scorecard® Poverty-Assessment Tool

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

Indicator	Value	Points	Score
1. How many household members are 14 years old or younger?	A. Three or more	0	
	B. Two	7	
	C. One	14	
	D. None	21	
2. How many household members participate in the household production or services in planting, breeding, forestry, or aquaculture?	A. Four or more	0	
	B. Three	1	
	C. Two	4	
	D. One	6	
	E. None	9	
3. What type of residence does the household have?	A. No data; temporary; other	0	
	B. Semi-permanent house	4	
	C. House with a shared kitchen or shared bathroom/toilet	7	
	D. Villa; house with a private kitchen and private bathroom/toilet	13	
4. What type of toilet does the household have?	A. None; other	0	
	B. Double-vault compost latrine; toilet directly over water	4	
	C. *Suilabh*; flush toilet with septic tank or sewage pipes	13	
5. What is the main source of cooking/drinking water for the household?	A. Filtered spring water; water from hand-dug, reinforced wells; water from hand-dug, non-reinforced, covered wells; water from hand-dug, non-reinforced, uncovered wells; water from river, lake, or pond; other	0	
	B. Rain water; water pumped from deep drilled wells	3	
	C. Private or public tap water; or bought water (in tank or bottle)	10	
6. What kind of cooker does the household have?	A. None	0	
	B. Electric cooker; rice cooker; pressure cooker (no gas cooker)	5	
	C. Gas cooker	12	
7. Does the household have a motorcycle?	A. No	0	
	B. Yes	7	
8. Does the household have a video player?	A. No	0	
	B. Yes	7	
9. Does the household have a wardrobe of any kind?	A. No	0	
	B. Yes	4	
10. Does the household have a pump?	A. None or unknown	0	
	B. 1 or more	4	
SimplePovertyScorecard.com		**Score:**	

Simple Poverty Scorecard® Poverty-Assessment Tool Vietnam

1. Introduction

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

The direct approach to poverty measurement via consumption surveys is difficult, lengthy, and costly, asking about a long list of consumption items ("Which of the following things has your household consumed during holidays in the last 12 months? Fragrant, specialty rice? Bought or bartered? What is the quantity and value? . . ."). In contrast, the indirect approach via the scorecard is simple, quick, and inexpensive. It uses 10 verifiable indicators (such as "Where do most household members go to the bathroom?" or "What is the main fuel your household uses for cooking?") to get a score that is highly correlated with poverty status as measured by the exhaustive consumption 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 land-ownership cut-offs or housing indices). 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, $1/day for the Millennium Development Goals, or the poorest half below the national poverty line as required of USAID microenterprise grantees, see U.S. Congress, 2004), or if it wants to measure movement across a poverty line (for example, movement across $1/day to report to the Microcredit Summit Campaign), then it needs an consumption-based, objective tool with known accuracy. While most organizations lack the resources to field consumption surveys—and even governments cannot survey large shares of all households—many 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 "`HHSIZE_2`", negative values, decimal places, and

standard errors). Thanks to the predictive-modeling phenomenon known as the "flat max", simple poverty-assessment tools are almost as 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 formula. 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 2004 Vietnam Living Standard Survey (VLSS). 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-specialist field workers can collect data and tally scores on paper in about 5 minutes.

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

Second, the scorecard can estimate a group's poverty rate at a point in time. (The "poverty rate" is also known as the "poverty prevalence", "head-count index", or "share below the poverty line".) This is simply the average poverty likelihood among individuals in the group.

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

The scorecard can also be used for targeting. To help managers choose a targeting cut-off, this paper reports the share of Vietnam's population who are at or below a given score cut-off and who are also below a given poverty line.

This paper presents a single scorecard whose indicators and points were derived from Vietnamese household consumption data and the international $3/person/day poverty line. Scores from this scorecard are calibrated to poverty likelihoods for eight poverty lines.

The scorecard is constructed using a sub-sample of data from the 2004 VLSS. Its accuracy is validated on a different sub-sample from the 2004 VLSS. While all three scoring estimators are unbiased when applied to the validation sample (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.

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

poverty will be the same as in the data used to build the scorecard.[1] Of course, this assumption—ubiquitous and inevitable in predictive modeling—holds only partly.

When applied to the validation sample, the absolute difference between scorecard estimates of groups' poverty rates and the true rates is 1.1 percentage points for the national line, 0.7 percentage points for $3/day, and 0.6 percentage points on average across all eight lines (Figure 11). This difference is due to sampling variation and not bias because its average would be zero if the VLSS data were to be repeatedly redrawn and divided into sub-samples before repeating the entire scorecard-building process.

For sample sizes of $n = 16{,}384$, the 90-percent confidence intervals for these estimates are ± 0.5 or less (Figure 11). For $n = 1{,}024$, the 90-percent intervals are ± 2.0 percentage points or less (Figure 12).

Section 2 below describes data and poverty lines. Section 3 places the scorecard in the context of existing tools for Vietnam. Sections 4 and 5 describe scorecard construction and offer practical guidelines for use. Sections 6 and 7 detail the estimation of individuals' 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.

[1] Bias 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 consumption surveys.

2. Data and poverty lines

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

2.1 Data

The scorecard is based on data from the 2004 VLSS. The data is 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 lines

The Vietnam General Statistical Office (GSO) reports a national poverty line and a food poverty line. The GSO food line is based on the cost of a food basket that provides 2,100 calories/person/day, and the GSO national poverty line is equal to the GSO food line plus the non-food consumption observed for people whose consumption of calories is close to the food line. This study calls the GSO lines the "national line" and the "food line". Glewwe (2005) also calculates a national poverty line and a food poverty line for Vietnam, which this study calls the "Glewwe national line" and the "Glewwe food line".

The scorecard here is constructed using the international $3/day poverty line, adjusted for urban/rural cost-of-living by region using price deflators provided by

Glewwe (personal communication). The $3/day line implies a poverty rate of 36.5 percent (Figure 3).

Because local pro-poor organizations may want to use different or various poverty lines, this paper calibrates scores from its single scorecard to poverty likelihoods for eight lines (figures in parentheses are per-capita daily poverty lines in units of 1000 dong, with poverty rates from Figure 3):

- National line (5.65, 12.0 percent)
- USAID "extreme" line (4.69, 6.0 percent)
- Food line (5.21, 9.3 percent)
- $1/day (2.86, 0.6 percent)
- $2/day (5.71, 12.4 percent)
- $3/day (8.57, 36.5 percent)
- Glewwe national line (6.52, 18.8 percent)
- Glewwe food line (3.63, 2.1 percent)

The USAID "extreme" line (U.S. Congress, 2002) is the median consumption of individuals below the national line, by urban/rural within provinces.

The $1/day line is derived from this data:

- 2000 purchase-power parity exchange rate: 2,320 dong per $1
- 2000 CPI: 100
- 2004 CPI: 114.9

The $1/day line for 2000 then (Sillers, 2006) is:

$$2{,}320 \times (114.9 \div 100) \times 1.08 = \text{Dong } 2{,}879.$$

The $1/day line is adjusted for differences in cost-of-living by rural/urban area within provinces using:

- L, the all-Vietnam $1/day line in dong
- p_{iu}, the urban population for region i
- p_{ir}, the rural population for region i
- π_i, Glewwe (2005) regional price deflator in region i
- π_u, Glewwe price deflator for urban areas
- π_r, Glewwe price deflator for rural areas

L_{ir} is the $1/day line adjusted for cost-of-living for the rural areas in region i:

$$L_{ir} = L \cdot \pi_i \cdot \frac{\pi_r}{100 \cdot (p_{iu}\pi_u + p_{ir}\pi_r)/(p_{iu} + p_{ir})}.$$

The lines for $2/day, $2.50/day, $3/day, and $4/day are multiples of the $1/day lines L_{ir}. Switching r and u in the formula gives the urban lines.

3. Existing poverty-assessment tools for Vietnam

This section reviews four existing poverty-assesssment tools for Vietnam. The main aspects of interest are the purpose of the study, methods, relative/absolute poverty estimation, poverty lines, indicators, accuracy/precision, and sample-size formula.

3.1 Minot

Minot (2000) seeks to build disaggregated poverty maps for targeting in rural Vietnam. Using Probit (akin to Logit here) and setting the poverty line at the 30th percentile of consumption, Minot first creates a poverty-assessment tool based on the 1992–3 Vietnam Living Standards Survey (VLSS). All of the tool's indicators appear in both the VLSS and the 1994 Agricultural Census. Each rural district's poverty rate is then estimated by applying the district means of the indicators from the Agricultural Census to the tool. The indicators are:

- Household size
- Share of household members of working age
- Ethnic group
- Household headship
- Main occupation
- Per-capita ownership of farmland
- Farming of perennial crops
- Per-capita annual production of paddy in kilograms
- Ownership of livestock (cattle, chickens, pigs)
- Area of residence
- Type of residence (permanent or semi-permanent)
- Source of drinking water
- Ownership of durables (television, radio, motorcycle)
- Region

Like the scorecard here, Minot adjusts for regional differences in cost-of-living and uses simple, inexpensive, verifiable indicators (except for food production in kilograms and area of residence). The two poverty-assessment tools differ (beyond the focus on estimating poverty rates for districts rather than individuals) in that, due to lack of household-level data from the Agricultural Census, Minot does not report accuracy, precision, or sample-size formulas.

3.2 Baulch

Baulch (2002) builds a household-level poverty-assessment tool for monitoring and targeting. Baulch is similar to the scorecard here in its focus and in that it:

- Uses a few simple, inexpensive-to-collect, and verifiable indicators
- Uses only categorical indicators
- Uses "c" to measure of how well the tool ranks households[2]
- Chooses indicators based on both statistics and judgment
- Uses a Probit regression on poverty status (like the Logit here), producing estimates of poverty likelihoods
- Uses the VLSS (although 1997/8 rather than 2004)
- Discusses the selection of a targeting cut-off based on the benefits and costs of successfully inclusion/exclusion versus mistaken undercoverage/leakage

[2] Baulch calls this the "area under the Receiver Operator Characteristic" curve.

Based a national poverty line of 4,904 dong/person/day, Baulch builds an urban poverty-assessment tool (six indicators) and a rural tool (nine indicators):

- Number of children (urban and rural)
- Number of women (urban and rural)
- Ethnicity (rural)
- Floor of residence (rural)
- Cooking fuel (urban and rural)
- Ownership of consumer durables:
  — Color television (urban and rural)
  — Black-and-white television (urban and rural)
  — Radio (rural)
  — Car or motorcycle (urban and rural)

Overall, the spirit and analysis of Baulch closely resemble that of the paper here. Still, Baulch does not test accuracy on data different from that used in tool construction, and he does not discuss estimates of poverty rates (despite having a goal that includes monitoring).

3.3 Sahn and Stifel

Like this paper, Sahn and Stifel (2003) seek a low-cost, practical way to measure poverty. They use factor analysis and the 1993 and 1998 VLSS to construct an "asset index" that "(a) is consistent with the financial means and technical capabilities of government statistical offices, and (b) provides sufficient information to identify and profile the poor [and] target transfers" (p. 465).

As here, Sahn and Stifle's indicators are simple, inexpensive, and verifiable:

- Ownership of consumer durables:
  - Radio
  - Stereo
  - Television
  - Sewing machine
  - Stove
  - Refrigerator
  - Bicycle
  - Motorcycles or cars
- Household quality:
  - Source of drinking water (piped/surface or well)
  - Toilet facilities (flush, pit/latrine, or none)
  - Cooking fuel (gas or electricity)
  - Quality of construction material of floor (low or high)
- Human capital (education of the household head)

To check coherency between the asset index and reported consumption[3] (or child nutrition), Sahn and Stifel rank Vietnamese households once based on the index and a second time based on consumption (or height-for-age). For each pair of proxies, they judge the coherence of the rankings by the distance between a given household's decile ranks. They conclude that the asset index predicts long-term nutritional status no worse than does current consumption. They also report that the asset index predicts expenditure worse than does predicted expenditure from a least-squares regression on household demographics, education, housing quality, and access to public services.

[3] They check the index against consumption because it is a common proxy for living standards, not because they believe consumption should be the benchmark.

3.4 IRIS Center

USAID commissioned IRIS Center ("IRIS", 2007a) to build a poverty-assessment tool for use by their Vietnamese microenterprise partners for reporting on their participants' poverty rates. Thus, IRIS considers only the USAID "extreme" poverty line (3,818 dong/person/day at January 1999 prices).

After comparing several statistical approaches,[4] IRIS settles on quantile regression (Koenker and Hallock, 2001). Their indicators[5] are:

- Household size
- Age of household head
- Number of household members with no education
- Number of rooms occupied
- Type of toilet facility
- Main source of lighting
- Main cooking fuel
- Type of roofing material
- Ownership of:
  — Refrigerator or freezer
  — Motorcycle
  — Radio, radio receiver, phonograph, or cassette player
  — Gas stove, electric stove, rice cooker, or pressure cooker
  — Television
- Number of chickens owned
- Whether any household member managed agricultural or forestry land or participated in agricultural or forestry cultivation, or raised livestock, or seafood on land managed or used by the household during the past 12 months
- Whether any household member has worked on any annual crop land belonging to the household
- Total land area of all of the plots owned by the household

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

With the possible exception of total land area owned, these indicators are simple, inexpensive, and verifiable.

IRIS' accuracy tests focus on the difference between the estimated poverty rate and its true value. IRIS also reports targeting accuracy (for a single cut-off) in terms of successful "hits" (*inclusion* when a household truly below a poverty line is predicted to have per capita consumption 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 adopted by USAID for certifying poverty-assessment tools (IRIS Center, 2005a). BPAC depends on accuracy in terms of the poverty rate at a point in time.[6] A higher BPAC means more accuracy; BPAC for IRIS for the USAID "extreme" line is 61.7 (IRIS Center, 2008). For the scorecard, BPAC for the USAID "extreme" line with the 2004 VLSS data is 19.7 (Figure 14). IRIS does not report,

[6] BPAC does not indicate precision, nor does IRIS report it. IRIS Center (2005a) says that BPAC is designed to consider accuracy both in terms of the estimated poverty rate and in terms of inclusion (successful classification of households below the poverty line). In fact, for a given poverty line and a single-step poverty-assessment tool, BPAC is maximized when the difference between the estimated poverty rate and its true value is minimized, regardless of inclusion. Thus, selecting a poverty-assessment tool on the basis of BPAC is equivalent to selecting on the basis of the difference between the estimated poverty rate and its true value. Thus, it would be clearer to drop the BPAC nomenclature and simply discuss directly the accuracy and precision of the estimated poverty rate.

however, what data they use to build and test their tool nor whether they weight by individuals or households.

3.5 The new scorecard for Vietnam

This study uses the 2004 VLSS to build a scorecard for Vietnam and to test its accuracy. Compared to the tools above, it has seven strengths.

First, it measures accuracy using different data than that used to construct the tool. This mimics how the scorecard is actually used in practice.

Second, this study reports scorecard indicators and points. This means that local pro-poor organizations in Vietnam can pick up the scorecard and use it.

Third, the scorecard here is designed to be practical for local pro-poor organizations. It has 10 indicators, all of them categorical and selected not only to be highly predictive of poverty but also verifiable, quick to answer, and liable to change over time. This facilitates data collection and improves data quality, which in turn improves accuracy. Baulch, Sahn and Stifel, and IRIS also use simple and inexpensive indicators; Minot includes some indicators that are difficult to verify. The scorecard here also has the most straightforward derivation and the simplest point scheme.

Fourth, the scorecard here is based on an absolute poverty line. While Minot, Baulch, and IRIS also do this, Sahn and Stifle do not. While this means that their index can be built without consumption 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 (unless built with pooled data as in Sahn and Stifle, 2000).

Fifth, this study adjusts poverty lines for differences in cost-of-living across urban/rural and provinces. Also, this study considers eight poverty lines, providing users with the flexibility to use the line most relevant for their purposes.

Sixth, this study reports sample-size formulas, and seventh, it uses the most recent data.

4. Scorecard construction

About 200 potential indicators are prepared in the areas of:

- Family composition (such as household headship and number of children)
- Education (such as the highest grade passed by a household member)
- Employment (such as male head/spouse working for a state-owned enterprise)
- Housing (such as residence type and toilet type)
- Asset ownership (such as video player and motorcycle)

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 video player or ownership of a pump is probably more likely to change in response to changes in poverty status than is the highest grade passed by any family member.

The scorecard itself is built using 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 build one scorecard for each candidate indicator, using Logit to derive points. 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 Vietnam. Evidence from India and Mexico (Schreiner, 2006 and 2005a), Sri Lanka (Narayan and Yoshida, 2005), and Jamaica (Grosh and Baker, 1995) suggests that segmenting poverty-assessment tools by urban/rural does not improve accuracy much.

5. Practical guidelines for scorecard use

The main challenge of scorecard design is not to squeeze out the last drops of accuracy but rather to improve the chances that scoring is actually used (Schreiner, 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 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 by hand in the field because the scorecard has:

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

The scorecard is ready to be photocopied. 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 to the far-right
- 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 results depend on quality inputs. If organizations or field workers gather their own data and have an incentive to exaggerate poverty rates (for example, if they are rewarded for reaching poorer participants), then it is wise to implement on-going quality control via data review and random audits (Matul and Kline, 2003).[7] IRIS Center (2007a) and Toohig (2007) are useful nuts-and-bolts guides for budgeting, training field workers and supervisors, logistics, sampling, interviewing, piloting, recording data, and quality control.

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

[7] If an organization does not want field workers to know the points associated with indicators, then it is a simple matter to remove the points from the paper scorecard and apply them later in a spreadsheet or database at the central office.

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

- Employees of the organization
- Third-party contractors

Scores 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
- In 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 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) using a desired level of confidence and a desired confidence interval.

The scorecard's frequency of application can be:

- At in-take 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)
- Each time a field worker visits a participant at home

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

- With a different, representative sets of participants
- With the same set of participants

An example set of choices were made by BRAC and ASA, two microlenders in Bangladesh (each with 7 million participants) who are applying the Simple Poverty Scorecard tool for Bangladesh (Schreiner, 2013). Their design is that loan officers in a random sample of branches score all participants each time they visit a homestead as part of their standard due diligence prior to loan disbursement (about once a year). Scores 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 individual poverty likelihoods

The sum of scorecard points for a household is called the *score*. For Vietnam, 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 national line, scores of 0–4 have a poverty likelihood of 90.9 percent (Figure 5), and scores of 45–49 have a poverty likelihood of 1.5 percent.

The poverty likelihood associated with a score varies by poverty line. For example, scores of 45–49 are associated with a poverty likelihood of 1.5 percent for the national line but 0.0 percent for the food line (Figure 5).[8]

[8] Starting with Figure 5, most figures have eight 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.

6.1 Calibrating scores with poverty likelihoods

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

Figure 6 illustrates this. For the example for the national line, there are 6,381 people with a score of 0–4, of whom 5,799 are below the poverty line. The estimated poverty likelihood associated with a score of 0–4 is then 90.9 percent, because 5,799 ÷ 6,381 = 90.9 percent.

As another illustration, with the national line and a score of 45–49, there are 5,294 people in the calibration sample, of whom 80 are below the line (Figure 6). Thus, the estimated poverty likelihood for a score of 45–49 is 80 ÷ 5,294 = 1.5 percent.

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

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

- 1.5 percent below the USAID "extreme" line
- 2.6 percent between the USAID "extreme" line and the food line
- 1.4 percent between the food line and national line
- 7.1 percent between the national line and the Glewwe national line
- 27.8 percent between the Glewwe national line and $3/day
- 59.7 percent above $3/day

The calibration process produces poverty likelihoods that are objective (that is, derived from data and consumption-based poverty lines) even though the scorecard is constructed partly based on judgment. 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 was 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 Vietnam'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 people 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. 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, 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.[9]

Of course, the relationship between indicators and poverty changes over time, so any scorecard applied out-of-sample (as all are in practice) will generally be biased. Still, estimators that are unbiased in-sample should have less bias out-of-sample.

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:[10]

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

[9] This follows because these estimates of groups' poverty rates are linear functions of the unbiased estimates of individuals' poverty likelihoods.
[10] Efron and Tibshirani, 1993.

For each of the 20 score ranges, Figure 8 shows the average difference between estimated and true poverty likelihoods as well as average confidence intervals around the differences.

For the national line, the average poverty likelihood across bootstrap samples for scores of 0–4 in the validation sample is too high by 2.9 percentage points (Figure 8). For scores of 5–9, the estimate is too low by 7.9 percentage points (Figure 8).[11]

For the validation sample, the 90-percent confidence interval for the differences for scores of 0–4 is ±3.8 percentage points (Figure 8).[12] This means that in 900 of 1,000 bootstraps, the difference between the estimate and the true value is between –0.9 and 6.7 percentage points (because 2.9 – 3.8 = –0.9, and 2.9 + 3.8 = 6.7). In 950 of 1,000 bootstraps (95 percent), the difference is 2.9 ±4.7 percentage points, and in 990 of 1,000 bootstraps (99 percent), the difference is 2.9 ±6.0 percentage points.

For almost all score ranges, Figure 8 shows differences—sometimes large ones—between estimated poverty likelihoods and true values. This is because the validation sub-sample is a single sample that—thanks to sampling variation—differs in distribution from the construction/calibration sub-samples and from Vietnam's population. For targeting, however, what matters is less the bias in all score ranges and more the bias in score ranges just above and below the targeting cut-off. This mitigates

[11] There are differences, in spite of the estimator's unbiasedness, because the estimates come from a single sample. Their average difference would be zero if samples were repeatedly drawn from Vietnam's population and split into sub-samples before repeating the entire scorecard-building process.

[12] Confidence intervals are a standard, widely understood measure of precision.

the effects of bias and sampling variation on targeting (Friedman, 1997). Section 9 below looks at targeting accuracy in detail.

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

Figure 9 (summarizing Figure 10 across poverty lines) shows that differences, when averaged across score ranges for a given poverty line, are typically 1.3 percentage points or less for the validation sample. The differences are due to sampling variation.

By construction, the scorecard here is unbiased when applied to the validation data. However, it may still be *overfit* when applied after 2004. That is, it may fit the 2004 VLSS 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 2004 VLSS. Or the scorecard may be overfit in that becomes biased as the relationship between indicators and poverty changes over time.

Overfitting can be mitigated by simplifying the scorecard and by relying not only on data but also on 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, but that would increase complexity too much in this context.

Most errors in individual households' likelihoods, however, cancel out in the estimates of groups' poverty rates (see later sections). Furthermore, much of the difference 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 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 group poverty rates 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 individuals in the group.

To illustrate, suppose a program samples three participants on Jan. 1, 2007 and that they have scores of 20, 30, and 40, corresponding to poverty likelihoods of 23.7, 9.7, and 1.4 percent (national line, Figure 5). The group's estimated poverty rate is the participants' average poverty likelihood of $(23.7 + 9.7 + 1.4) \div 3 = 11.6$ percent.[13]

7.1 Accuracy of estimated poverty rates at a point in time

How accurate is this estimate? For a range of sample sizes, Figure 12 reports average differences between estimated and true poverty rates as well as average confidence intervals for the differences for the scorecard applied to 1,000 bootstraps from the validation sample. For the national poverty line and a sample size of $n =$ 16,384, the scorecard estimate is too high by about 1.1 percentage points; it estimates a poverty rate of 12.0 percent, but the true value for the validation sample is 10.9 percent (Figure 2). For all poverty lines, differences for the validation sample are 1.1 percentage points or less, with an average of about 0.6 percentage points (Figure 11).[14]

[13] The group's poverty rate is not the poverty likelihood associated with the average score. Here, the average score is $(20 + 30 + 40) \div 3 = 30$, so the poverty likelihood associated with the average score is 9.7 percent. This is not the 11.6 percent found as the average of the three poverty likelihoods associated with each of the three scores.
[14] Figure 11 summarizes Figure 12.

As before, these differences are due to sampling variation in the validation sample and in the random division of the 2004 VLSS 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 0.3 percentage points or less (Figure 11). This means that in 900 of 1,000 bootstraps of this size, the difference between the estimate and the true value is within 0.3 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 1.1 – 0.3 = 0.8 to 1.1 + 0.3 = 1.4 percentage points. (1.1 percentage points is the average difference, and ±0.3 is its 90-percent confidence interval.)

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

How many participants should an organization sample if it wants to estimate their poverty rate at a point in time for a desired confidence interval and desired confidence level? The first paper in the poverty-assessment literature to address practical question is Schreiner (2008a).[15]

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

With direct measurement, the poverty rate can be estimated as the number of people observed to be below the poverty line, divided by the number of all observed people. The formula for sample size n is then (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 people 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 Vietnam scorecard, consider the national poverty line and the scorecard applied to the validation sample. Figure 2 shows that the expected (before measurement) poverty rate $\hat{p}$ in the construction sample is 0.125, the weighted average of 0.124 and 0.126. In turn, a sample size n of 16,384 and a 90-percent confidence level correspond to a confidence interval of ±0.35 percentage points (Figure 12).[16] Plugging these into the direct-measurement sample-size formula (1) above gives not n = 16,384 but rather $n = \left(\frac{1.64}{0.0035}\right)^2 \cdot 0.125 \cdot (1 - 0.125) =$

[16] Due to rounding, Figure 12 displays 0.3, not 0.35.

24,015. The ratio of the sample size for scoring (derived empirically) to the sample size for direct measurement (derived from theory) is 16,384 ÷ 24,015 = 0.68.

Applying the same method to n = 8,192 gives $n = \left(\frac{1.64}{0.0050}\right)^2 \cdot 0.125 \cdot (1 - 0.125) =$ 11,767. This time, the ratio of the sample size using scoring to the sample size using direct measurement is 8,192 ÷ 11,767 = 0.70. This ratio of 0.70 for n = 8,192 is close to the ratio of 0.68 for n = 16,384. Indeed, applying this same procedure for all $n \geq 256$ in Figure 11 gives ratios that average to 0.70. This can be used to define a sample-size formula for the Vietnam scorecard applied to the validation sample:

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

where $\alpha = 0.70$ and z, c, and $\hat{p}$ are defined as in (1) above.

To illustrate, if c = 0.029 (confidence interval of ±2.9 percentage points) and z = 1.64 (90-percent confidence), then (2) gives $n = 0.70 \cdot \left(\frac{1.64}{0.029}\right)^2 \cdot 0.125 \cdot (1 - 0.125) = 245$ which is close to the sample size of 256 for these parameters in Figure 12.

If the sample-size factor α is less than 1.0, it means that the scorecard is more precise than direct measurement. For Vietnam, α ranges from 0.65 to 0.88 for the estimates of groups' poverty rates at a point in time (Figure 11). Thus, the scorecard is more precise than direct measurement.

Of course, the sample-size formula here is specific to Vietnam, 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, an organization would select a poverty line (say, Vietnam's national poverty line), select a desired confidence level (say, 90 percent, or $z = 1.64$), select a desired confidence interval (say, $\pm$ 2 percentage points, or $c = 0.02$), make an assumption about $\hat{p}$ (perhaps based on a previous measurement or national figures), assume that the scorecard works out-of-sample,[17] and compute the required sample size. In this illustration, $n = 0.70 \cdot \left(\frac{1.64}{0.02} \right)^2 0.125 \cdot (1 - 0.125) = 515$.

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

[17] This paper reports accuracy for the scorecard applied to the 2004 validation sample, but it cannot test accuracy for later years. Still, performance after 2004 will most likely resemble performance in 2004, 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 individuals in the group. With data for 2004, this paper cannot estimate changes over time nor present sample-size formula specific to Vietnam. Nevertheless, the 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 the strong assumptions that the population is constant over time and that program drop-outs do not differ from others.

8.2 Calculating estimated changes in poverty rates over time

Consider the illustration begun in the previous section. On Jan. 1, 2007, a program samples three participants who score 20, 30, and 40 and so have poverty likelihoods of 23.7, 9.7, and 1.4 percent (national line, Figure 5). The group's baseline estimated poverty rate is the participants' average poverty likelihood of (23.7 + 9.7 + 1.4) ÷ 3 = 11.6 percent).

After baseline, two sampling approaches are possible:

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

By way of illustration, suppose that on Jan. 1, 2008, the program samples three additional people who are in the same cohort as the three people originally sampled (or scores the same three original people) and finds that their scores are 26, 35, and 45 and so have poverty likelihoods of 15.7, 5.4, and 1.5 percent (national line, Figure 5). Their average poverty likelihood at follow-up is now (15.7 + 5.4 + 1.5) ÷ 3 = 7.5 percent, an improvement of 11.6 – 7.5 = 4.1 percentage points.

This suggests that about 41 of 1000 participants crossed the poverty line in 2007.[18] Among those who started below the line, about one in three (4.1 ÷ 11.6 = 35.3 percent) ended up above the line.[19]

[18] This is a net figure; some people started above the poverty line and ended below it, and vice versa.

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

8.3 Accuracy for estimated change for two independent samples

With data only for 2004, so it is not possible to measure the accuracy of scorecard estimates of changes in groups' poverty rates over time. In practice, of course, Vietnam'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 sample-size formula for 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.[20]

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

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

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.

[20] This means that, for a given precision, estimating the change in a poverty rate between two points in time requires 4 times as many measurements (not twice as many) as does estimating a poverty rate at a point in time.

For Peru and India (Schreiner, 2008a and 2008b), the average α across poverty lines is 1.6 and 1.2, so 1.5 may be a reasonably conservative figure for Vietnam.

To illustrate how to use (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 \$2/day, $\alpha = 1.5$, and $\hat{p} = 0.125$ (from Figure 2). Then baseline sample size is $n = 1.5 \cdot 2 \cdot \left(\frac{1.64}{0.02}\right)^2 \cdot 0.125 \cdot (1 - 0.125) = 2{,}207$, and follow-up sample size is also 2,207.

8.4 Accuracy for estimated change for one sample, scored twice

In general, the direct-measurement sample-size formula for this case is:[21]

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

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.

[21] 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 poverty rate is zero. Then $\hat{p}_{12} = \hat{p}_{21}$ and (5) becomes:

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

where $\hat{p}_* = \hat{p}_{12} = \hat{p}_{21}$.

Still, $\hat{p}_*$ could take any value between 0 and 1, so (6) cannot determine sample size. The estimate of $\hat{p}_*$ must be based on data available before baseline measurement.

Suppose that the observed relationship between $\hat{p}_*$ and the variance of the baseline poverty rate $p_{baseline} \cdot (1 - p_{baseline})$ is—as in Peru, see Schreiner, 2008a—close to $\hat{p}_* = 0.0085 + 0.206 \cdot [p_{baseline} \cdot (1 - p_{baseline})]$. Of course, $p_{baseline}$ is not known before the measurement, but it is reasonable to use as its expected value the observed poverty rate from the previous year. Given this and a poverty line, a sample-size formula for a single sample directly measured twice for Vietnam after 2004 is:

$$n = 2 \cdot \left(\frac{z}{c}\right)^2 \cdot \{0.0085 + 0.206 \cdot [p_{2004} \cdot (1 - p_{2004})]\}. \tag{7}$$

As usual, (7) is modified with α to get the scorecard sample-size formula:

$$n = \alpha \cdot 2 \cdot \left(\frac{z}{c}\right)^2 \cdot \{0.0085 + 0.206 \cdot [p_{2004} \cdot (1 - p_{2004})]\}. \tag{8}$$

In Peru (the only other country for which there is an estimate), the average α across years and poverty lines is about 1.8 (Schreiner, 2008a).

To illustrate the use of (8), 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, and the panel will be scored in 2004. The before-baseline poverty rate is 0.125 percent (p_{2004} =0.125, Figure 2), and suppose $\alpha = 1.8$. Then baseline sample size is $n = 1.8 \cdot 2 \cdot \left(\frac{1.64}{0.02}\right)^2 \cdot \left\{0.0085 + 0.206 \cdot \left[0.125 \cdot (1 - 0.125)\right]\right\} = 751$. Of course, $n = 751$ for the follow-up sample as well.

9. Targeting

When a program uses the scorecard for targeting, people 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. People with higher scores are *non-targeted* and treated—for program purposes—as if they are above a given poverty line.

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

Targeting is successful when people truly below a poverty line are targeted (*inclusion*) and people truly above a poverty line are not targeted (*exclusion*). Of course, no scorecard is perfect, and targeting is unsuccessful when people truly below a poverty line are not targeted (*undercoverage*) or people truly above a poverty line are targeted (*leakage*). Figure 13 illustrates these four possible targeting outcomes. Targeting accuracy varies by cut-off; a higher cut-off has better inclusion (but worse leakage), while a lower cut-off has better exclusion (but worse undercoverage).

A program should weigh these trade-offs when setting a cut-off. A formal way to do this is to assign net benefits—based on a program's values and mission—to each of the four possible targeting outcomes and then to choose the cut-off that maximizes total net benefits (Adams and Hand, 2000; Hoadley and Oliver, 1998).

Figure 14 show the percentage of people by targeting outcome for the scorecard applied to the validation sample. Given an example cut-off of 15–19, outcomes for the national poverty line applied to the validation sample are:

- Inclusion: 6.6 percent are below the line and correctly targeted
- Undercoverage: 4.2 percent are below the line and mistakenly not targeted
- Leakage: 7.1 percent are above the line and mistakenly targeted
- Exclusion: 82.1 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: 8.3 percent are below the line and correctly targeted
- Undercoverage: 2.6 percent are below the line and mistakenly not targeted
- Leakage: 12.0 percent are above the line and mistakenly targeted
- Exclusion: 77.2 percent are above the line and correctly not targeted

Which cut-off is preferred depends on total net benefit. Suppose each targeting outcome has a per-person benefit or cost. Then total net benefit for a given cut-off is:

Benefit per person correctly included	x	People correctly included	+
Cost per person mistakenly not covered	x	People mistakenly not covered	+
Cost per person mistakenly leaked	x	People mistakenly leaked	+
Benefit per person correctly excluded	x	People 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 14 for a 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, 2005a).[22] With this, total net benefit is the number of people correctly included or excluded:

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

Figures 14 shows "Total Accuracy" for all cut-offs for the Vietnam scorecard applied to the validation sample. For the national line, total net benefit is greatest (91.6) for a cut-off of 10–14; that cut-off would correctly classify about nine in ten Vietnamese.

"Total Accuracy" weighs successful inclusion of those below the poverty line equally with successful exclusion of those above the poverty 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 People correctly included) + (1 x People correctly excluded).

Beyond "Total Accuracy", IRIS (2005a) proposes a new yardstick called BPAC. As discussed earlier, BPAC is maximized by minimizing the difference between the estimated poverty rate and the true value. Thus, discussing accuracy in terms of BPAC is just an obscure way of discussing accuracy in terms of estimated poverty rates.

[22] Grootaert and Braithwaite (1998) use this criterion with poverty-assessment tools.

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. Figure 15 shows, for the Vietnam scorecard applied to the validation sample, the expected poverty rate among people who score at or below all possible cut-offs. For the example of the national poverty line in 2004, targeting people who score 25–29 or less would lead to a poverty rate of 32.9 percent and would mean targeting 28.9 percent of all Vietnamese.

10. Conclusion

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

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

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

The scorecard is built with a sub-sample of data from the 2004 VLSS, tested with a different sub-sample, and calibrated to eight poverty lines (national, food, USAID "extreme", \$1/day, \$2/day, \$3/day, Glewwe national, and Glewwe food).

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.

When the scorecard is applied to the validation sample, the difference between estimated versus true group poverty rates at a point in time is always less than 1.1 percentage points and averages—across the eight poverty lines—about 0.6 percentage points. For $n = 16{,}384$ and 90-percent confidence, the precision of these differences is are ± 0.5 percentage points or less, and for $n = 1{,}024$, precision is ± 2.0 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 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 observe and verify. Indicator weights are all zeros or positive integers, and scores range from 0 (most likely to be below a poverty line) to 100 (least likely to be 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 can compute scores in the field.

In sum, the scorecard is a practical, objective way for pro-poor programs in Vietnam 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 consumption survey.

References

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

Baesens, B.; Van Gestel, T.; Viaene, S.; Stepanova, M.; Suykens, J.; and J. Vanthienen. (2003) "Benchmarking State-of-the-Art Classification Algorithms for Credit Scoring", *Journal of the Operational Research Society*, Vol. 54, pp. 627–635.

Baulch, Bob. (2002) "Poverty Monitoring and Targeting Using ROC Curves: Examples from Vietnam", IDS Working Paper No. 161, `www.ids.ac.uk/ids/bookshop/wp/wp161.pdf`, accessed Aug. 10, 2008.

Caire, Dean. (2004) "Building Credit Scorecards for Small Business Lending in Developing Markets", `microfinance.com/English/Papers/Scoring_SMEs_Hybrid.pdf`, accessed July 25, 2008.

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.*

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

Fuller, Rob. (2006) "Measuring Poverty of Microfinance Clients in Haiti", `microfinance.com/English/Papers/Scoring_Poverty_Haiti_Fuller.pdf`, accessed July 27, 2008.

Glewwe, Paul. (2005) "Mission Report for Trip to Vietnam".

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 July 25, 2008.

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. (2008) "Accuracy Results for 17 Poverty Assessment Tool Countries", povertytools.org/other_documents/PAT_17_country_accur.pdf, accessed July 25, 2008.

_____. (2007a) "Client Assessment Survey—Vietnam", povertytools.org/USAID_documents/Tools/Current_Tools/USAID_PAT_VIET_7-2007.xls, accessed August 10, 2008.

_____. (2007b) "Manual for the Implementation of USAID Poverty Assessment Tools", povertytools.org/training_documents/Manuals/USAID_PAT_Manual_Eng.pdf, accessed July 25, 2008.

_____. (2007c) "Introduction to Sampling for the Implementation of PATs", povertytools.org/training_documents/Sampling/Introduction_Sampling.ppt, accessed July 25, 2008.

_____. (2005a) "Notes on Assessment and Improvement of Tool Accuracy", povertytools.org/other_documents/AssessingImproving_Accuracy.pdf, accessed July 25, 2008.

Johnson, Glenn. (2007) "Lesson 3: Two-Way Tables—Dependent Samples", www.stat.psu.edu/online/development/stat504/03_2way/53_2way_compare.htm, accessed July 25, 2008.

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.

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

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 July 25, 2008.

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

Minot, Nicholas. (2000) "Generating Disaggregated Poverty Maps: An Application to Vietnam", *World Development,* Vol. 28, No. 2, pp. 319–331.

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

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.

Sahn, David E.; and David Stifel. (2003) "Exploring Alternative Measures of Welfare in the Absence of Expenditure Data", *Review of Income and Wealth*, Vol. 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. (2013) "Simple Poverty Scorecard Poverty-Assessment Tool: Bangladesh", `SimplePovertyScorecard.com/BGD_2010_ENG.pdf`, accessed 1 July 2016.

_____. (2008a) "Simple Poverty Scorecard Poverty-Assessment Tool: Peru", SimplePovertyScorecard.com/PER_2003_ENG.pdf, accessed 1 July 2016.

_____. (2008b) "Simple Poverty Scorecard Poverty-Assessment Tool: India", SimplePovertyScorecard.com/IND_2005_ENG.pdf, accessed 1 July 2016.

_____. (2006) "Is One Simple Poverty Scorecard Poverty-Assessment Tool Enough for India?", microfinance.com/English/Papers/Scoring_Poverty_India_Segments.pdf, accessed July 25, 2008.

_____. (2005a) "La Herramienta del Índice de Calificacíon de la Pobreza™: México", SimplePovertyScorecard.com/MEX_2002_SPA.pdf, accessed July 25, 2008.

_____. (2005b) "IRIS Questions on the Simple Poverty Scorecard Poverty-Assessment Tool", microfinance.com/English/Papers/Scoring_Poverty_Response_to_IRIS.pdf, accessed July 25, 2008.

_____. (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 July 25, 2008.

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

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

Toohig, Jeff. (2007) "Progress Out of Poverty Index®: Training Guide", progressoutofpoverty.org/toolkit, accessed July 25, 2008.

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

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

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

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

			% with expenditure below a poverty line							
			National		USAID	International			Glewwe	
Sub-sample	**Year of Survey**	**Households**	**National**	**Food**	**'Extreme'**	**$1/day**	**$2/day**	**$3/day**	**National**	**Food**
Construction										
Selecting indicators and weights	'04	2,991	12.4	9.6	6.4	0.7	12.9	37.5	18.9	2.3
Calibration										
Associating scores with likelihoods	'04	3,142	12.6	9.6	6.3	0.7	12.9	36.5	19.2	2.4
Validation										
Applying scorecards	'04	3,056	10.9	8.6	5.2	0.6	11.5	35.5	18.2	1.7
Change in overall poverty rate (percentage points)										
Between samples	'04 to '04	3,056 and 6,133	1.6	1.0	1.1	0.1	1.4	1.5	0.8	0.7

Data Source: 2004 VLSS

Figure 3: Average poverty lines and poverty rates by region

		Poverty line (1000 dong/person/day) and poverty rate (%)																							
	Line or	National			National Food			USAID 'Extreme'			International $1/day			International $2/day			International $3/day			Glewwe National			Glewwe Food		
Region	rate	Urban	Rural	All	Urban	Rural	All	Urban	Rural	All	Urban	Rural	All	Urban	Rural	All	Urban	Rural	All	Urban	Rural	All	Urban	Rural	All
Red River Delta	Line	5.60	5.46	5.49	5.17	5.04	5.07	4.91	4.84	4.85	2.84	2.76	2.78	5.67	5.53	5.56	8.51	8.29	8.34	6.47	6.30	6.34	3.61	3.51	3.53
	Rate	0.1	6.7	5.2	0.1	4.0	3.1	0.0	3.3	2.6	0.0	0.0	0.0	0.1	7.1	5.6	3.9	35.2	28.4	1.0	13.4	10.7	0.0	0.6	0.4
North East	Line	5.66	5.52	5.54	5.22	5.09	5.12	5.42	4.65	4.79	2.86	2.79	2.80	5.73	5.58	5.61	8.59	8.37	8.41	6.53	6.37	6.40	3.64	3.55	3.57
	Rate	3.2	20.1	16.9	0.7	15.0	12.3	1.4	10.1	8.4	0.0	0.6	0.5	3.2	20.9	17.6	11.4	58.8	50.0	4.9	31.6	26.6	0.0	2.0	1.7
North West	Line	5.97	5.82	5.84	5.51	5.37	5.39	4.30	4.42	4.40	3.02	2.94	2.95	6.04	5.88	5.90	9.06	8.83	8.86	6.89	6.71	6.73	3.84	3.74	3.75
	Rate	5.1	50.1	44.3	2.7	43.6	38.3	1.7	25.1	22.0	0.0	4.2	3.7	5.1	51.5	45.5	24.5	83.9	76.2	12.6	64.0	57.4	0.0	13.2	11.5
North Central Coast	Line	5.64	5.50	5.52	5.21	5.07	5.09	5.35	4.41	4.53	2.85	2.78	2.79	5.71	5.56	5.58	8.56	8.34	8.37	6.51	6.34	6.36	3.63	3.54	3.55
	Rate	0.6	25.0	21.7	0.0	20.5	17.7	0.6	12.4	10.8	0.0	1.1	0.9	0.6	26.0	22.6	12.3	62.1	55.5	1.3	36.6	31.9	0.0	4.1	3.5
South Central Coast	Line	5.74	5.60	5.64	5.30	5.17	5.21	4.50	4.50	4.50	2.91	2.83	2.85	5.81	5.66	5.71	8.72	8.50	8.56	6.63	6.46	6.51	3.69	3.60	3.63
	Rate	2.0	17.1	12.8	1.8	12.7	9.6	1.0	8.4	6.3	0.0	1.9	1.4	2.0	17.3	13.0	8.2	47.5	36.3	3.9	23.5	17.9	0.5	4.2	3.2
Central Highlands	Line	6.01	5.85	5.90	5.54	5.40	5.44	4.12	4.48	4.38	3.04	2.96	2.98	6.08	5.92	5.97	9.11	8.88	8.95	6.93	6.76	6.80	3.86	3.77	3.79
	Rate	8.0	34.8	27.4	8.0	29.9	23.8	3.7	17.3	13.5	0.0	4.1	3.0	8.0	35.4	27.8	17.2	63.2	50.5	11.3	44.1	35.1	2.6	11.2	8.8
South East	Line	5.91	5.76	5.84	5.46	5.32	5.39	4.42	4.80	4.59	2.99	2.92	2.96	5.98	5.83	5.91	8.97	8.75	8.87	6.82	6.65	6.74	3.80	3.71	3.76
	Rate	0.3	5.5	2.7	0.3	3.9	1.9	0.0	2.8	1.3	0.0	0.7	0.3	0.3	6.0	2.9	1.4	22.6	11.0	0.8	10.3	5.1	0.0	1.8	0.8
Mekong River Delta	Line	5.82	5.68	5.71	5.38	5.24	5.27	5.10	4.77	4.83	2.95	2.87	2.89	5.89	5.74	5.77	8.84	8.62	8.66	6.72	6.55	6.58	3.75	3.65	3.67
	Rate	1.5	9.9	8.2	1.5	7.4	6.2	0.9	5.0	4.2	0.0	0.1	0.1	1.5	10.5	8.7	18.0	40.3	35.9	6.8	18.1	15.8	0.0	0.9	0.7
All Vietnam	Line	5.79	5.60	5.65	5.35	5.17	5.21	4.75	4.67	4.69	2.93	2.83	2.86	5.86	5.66	5.71	8.79	8.49	8.57	6.68	6.46	6.52	3.73	3.60	3.63
	Rate	1.4	15.6	12.0	1.1	12.1	9.3	0.7	7.8	6.0	0.0	0.9	0.6	1.4	16.3	12.4	8.0	46.4	36.5	3.3	24.2	18.8	0.2	2.8	2.1

Data Source: 2004 VLSS

Figure 4: Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (Possible responses ordered starting with those most strongly indicative of poverty)
241	What kind of cooker does the household have? (None; Electric cooker, rice cooker, or pressure cooker (no gas cooker); Gas cooker)
229	What type of toilet does the household have? (None, or other; Double-vault compost latrine, or toilet directly over the water; *Suilabh*, or flush toilet with septic tank or sewage pipes)
146	Does the household have any telephone sets? (No; Yes)
140	Does the household have a motorcycle? (No; Yes)
138	What is the main source of water for daily use in the household? (Bought water (in tank, bottle), filtered spring water, water from hand-dug, reinforced wells, water from hand-dug, non-reinforced, covered wells, water from hand-dug, non-reinforced, uncovered wells, river, lake, pond, or others; Water pumped from deep drilled wells; Private or public tap water, or rain water)
137	What is the main source of cooking/drinking water for the household? (Filtered spring water, water from hand-dug, reinforced wells, or water from hand-dug, non-reinforced, covered wells, or water from hand-dug, non-reinforced, uncovered wells, or water from river, lake or pond, or other; Rain water, water pumped from deep drilled wells; Public tap water, or bought water (in tank or bottle))
134	What is the highest grade passed by any family member? (No degree, primary school; Lower secondary school; Upper secondary school; Short-term technical worker, long-term technical worker, professional/secondary school; Junior-college diploma, Bachelor's, Master's, or Doctorate)
130	Does the household have any color TV sets? (No; Yes)
126	How many household members participate in the production or services of planting, breeding, forestry or aquaculture? (Four or more; Three; Two; One; None)
115	Does the household have a video player? (No; Yes)
114	Does the household have any refrigerators? (No; Yes)
106	What type of residence does the household have? (No data, temporary, or other; Semi-permanent house; House with a shared kitchen or shared bathroom/toilet; Villa, or house with a private kitchen and private bathroom/toilet)
105	How many household members are 14-years-old or younger? (Three or more; Two; One; None)
81	Do you use a filter or chemicals to purify water for daily consumption? (No; Yes)
73	Does the household have a wardrobe of any kind? (No; Yes)
71	Do you use a filter or chemicals to purify water for cooking or drinking? (No; Yes)
65	Does the male head/spouse participate in the production or services of planting, breeding, forestry, or aquaculture? (No; Yes)

Source: 2004 VLSS; $3/day poverty line.

Figure 4 (continued): Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (Possible responses ordered starting with those most strongly indicative of poverty)
64	What is the highest grade passed by the female head/spouse? (No data or no female head/spouse; No degree, primary school, or lower secondary school; Upper secondary school; Short-term technical worker, long-term technical worker, professional/secondary school, junior-college diploma, Bachelor's, Master's, or Doctorate)
62	How many male members of the household are 17-years-old or younger? (Three or more; Two; One; None)
61	How many children 16-years-old or younger in the household are attending school? (Not all; All; No children in the age range)
60	What is the highest grade passed by the male head/spouse? (No data or no male head/spouse, no degree, or primary school; Lower secondary school; Upper secondary school, short-term technical worker, long-term technical worker, professional/secondary school, junior-college diploma, Bachelor's, Master's, or Doctorate)
60	Does the household have any electric fans? (No; Yes)
57	How many male children 15-years-old or younger in the household are attending school? (Not all; All; No male children in the age range)
52	Does the female head/spouse participate in the production or services of planting, breeding, forestry, or aquaculture? (No; Yes)
50	Does the household have any computers? (No; Yes)
49	Does the household have any washing machines and dryers? (No; Yes)
48	Does the household have any drawing, plowing, or breeding cattle? (No; Yes)
45	What is the main source of lighting? (Electricity, battery lamp, or resin torch; Gas, oil, kerosene lamps, or other)
42	Does the household have any water heaters? (No; Yes)
42	How many female members in the household are 12-years-old or younger? (Two or more; One; None)
41	Does the household have any fruit-grinding and pressurizing machines? (No; Yes)
39	Does the household have any tables, chairs, or sofas? (No; Yes)
38	Can the female head/spouse read and write? (No; Yes)
35	Does the household have a pump? (No; Yes)
35	What is the total living area of the household in square meters? (54 or less; 55 to 99; 100 or more)
34	How many female children between 6 and 12-years-old are attending school? (Not all; All; No female children in the age range)
33	Does the household have any multi-tier stereos? (No; Yes)
27	Does the male head/spouse work for a state-owned enterprise? (No; Yes)
26	Does the female head/spouse work for a state-owned enterprise? (No; Yes)
22	Can the male head/spouse read and write? (No; Yes)

Source: 2004 VLSS; $3/day poverty line.

National Poverty Line Tables

(and tables pertaining to all eight poverty lines)

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

If an individual's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	90.9
5–9	71.0
10–14	65.6
15–19	34.7
20–24	23.7
25–29	15.7
30–34	9.7
35–39	5.4
40–44	1.4
45–49	1.5
50–54	0.8
55–59	0.0
60–64	0.0
65–69	0.0
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Surveyed cases weighted to represent Vietnam's population.

Based on the 2004 VLSS

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

Score	People below poverty line		All people at score		Poverty likelihood (estimated, %)
0–4	5,799	÷	6,381	=	90.9
5–9	3,321	÷	4,677	=	71.0
10–14	4,277	÷	6,519	=	65.6
15–19	1,928	÷	5,557	=	34.7
20–24	1,746	÷	7,374	=	23.7
25–29	1,264	÷	8,070	=	15.7
30–34	635	÷	6,585	=	9.7
35–39	314	÷	5,842	=	5.4
40–44	94	÷	6,788	=	1.4
45–49	80	÷	5,294	=	1.5
50–54	36	÷	4,806	=	0.8
55–59	0	÷	5,057	=	0.0
60–64	0	÷	4,900	=	0.0
65–69	0	÷	4,669	=	0.0
70–74	0	÷	3,907	=	0.0
75–79	0	÷	2,650	=	0.0
80–84	0	÷	3,807	=	0.0
85–89	0	÷	4,168	=	0.0
90–94	0	÷	420	=	0.0
95–100	0	÷	2,526	=	0.0

Number of all people normalized to sum to 100,000.
Based on the 2004 VLSS

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

	Likelihood of having expenditure in range demarcated by daily per capita poverty lines (1000 dong)					
	<USAID	≥USAID and <Food	≥Food and <National	≥National and <Glewwe National	≥Glewwe National and <$3/day	≥$3/day
Score	<DONG4.69	≥DONG4.69 and <DONG5.21	≥DONG5.21 and <DONG5.65	≥DONG5.65 and <DONG6.52	≥DONG6.52 and <DONG8.57	≥DONG8.57
0–4	73.0	14.8	3.1	2.1	5.3	1.6
5–9	51.5	12.1	7.4	14.5	13.1	1.4
10–14	30.8	25.3	9.5	11.1	17.5	5.8
15–19	15.2	8.8	10.8	23.8	26.2	15.3
20–24	11.0	4.5	8.2	15.0	43.0	18.4
25–29	4.6	5.2	5.9	16.1	37.9	30.3
30–34	3.6	2.7	3.4	6.8	35.3	48.2
35–39	1.5	2.6	1.4	7.1	27.8	59.7
40–44	0.0	0.5	0.9	2.7	17.5	78.5
45–49	0.0	0.0	1.5	4.4	14.4	79.7
50–54	0.0	0.0	0.8	1.6	9.1	88.5
55–59	0.0	0.0	0.0	2.8	2.2	95.0
60–64	0.0	0.0	0.0	0.0	0.9	99.1
65–69	0.0	0.0	0.0	0.0	1.1	98.9
70–74	0.0	0.0	0.0	0.0	0.8	99.2
75–79	0.0	0.0	0.0	0.0	0.0	100.0
80–84	0.0	0.0	0.0	0.0	0.0	100.0
85–89	0.0	0.0	0.0	0.0	0.0	100.0
90–94	0.0	0.0	0.0	0.0	0.0	100.0
95–100	0.0	0.0	0.0	0.0	0.0	100.0

All poverty likelihoods in percentage units.

The $2 line is omitted because it is close to the national line.

The $1 line and Glewwe's food line are omitted because they imply poverty rates of almost zero.

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

	Scorecard applied to validation sample difference between estimate and true value			
		Confidence interval (+/- percentage points)		
Score	**Bias**	**90-percent**	**95-percent**	**99-percent**
0–4	2.9	3.8	4.7	6.0
5–9	-7.9	5.9	6.2	6.9
10–14	12.3	3.8	4.6	5.9
15–19	6.6	2.4	2.9	3.7
20–24	-1.8	2.4	2.9	3.9
25–29	2.0	1.6	1.9	2.5
30–34	0.5	1.3	1.6	2.1
35–39	1.4	1.1	1.3	1.8
40–44	-0.9	0.9	1.0	1.2
45–49	1.4	0.1	0.2	0.2
50–54	0.8	0.0	0.0	0.0
55–59	0.0	0.0	0.0	0.0
60–64	0.0	0.0	0.0	0.0
65–69	0.0	0.0	0.0	0.0
70–74	0.0	0.0	0.0	0.0
75–79	0.0	0.0	0.0	0.0
80–84	0.0	0.0	0.0	0.0
85–89	0.0	0.0	0.0	0.0
90–94	0.0	0.0	0.0	0.0
95–100	0.0	0.0	0.0	0.0

Based on scorecard applied to validation sample.

Figure 9 (All poverty lines): Differences and precision for bootstrapped estimates of individuals' poverty likelihoods for the validation sample

	Poverty line							
		National	USAID	International			Glewwe	
Year scorecard applied	National	Food	'Extreme'	$1/day	$2/day	$3/day	National	Food
Estimate minus true value	0.8	0.8	1.0	0.2	0.4	0.1	0.2	1.3
Precision of difference	0.4	0.4	0.4	0.3	0.4	0.3	0.4	0.4

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.

Figure 10 (National poverty line): Differences and precision for bootstrapped estimates of individuals' poverty likelihoods, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/- percentage points)		
Sample size (n)	Bias	90-percent	95-percent	99-percent
2	0.9	31.2	41.8	61.4
4	0.4	24.0	30.8	42.3
8	0.8	16.9	21.1	28.5
16	1.0	11.8	14.1	18.3
32	1.2	8.5	10.3	13.4
64	1.0	6.6	7.7	10.6
128	1.0	5.1	6.0	7.7
256	0.8	3.4	4.0	5.7
512	0.8	2.2	2.7	4.0
1,024	0.8	1.6	1.9	2.5
2,048	0.8	1.1	1.3	1.7
4,096	0.8	0.8	0.9	1.2
8,192	0.8	0.6	0.7	0.8
16,384	0.8	0.4	0.5	0.6

Figure 11 (All poverty lines): Differences, precision, and sample-size α for bootstrapped estimates of groups' poverty rates at a point in time for the scorecard applied to the validation sample

	Poverty line							
		National	USAID	International			Glewwe	
Year scorecard applied	National	Food	'Extreme'	$1/day	$2/day	$3/day	National	Food
Estimate minus true value								
Estimate minus true value	1.1	0.4	0.6	0.0	0.8	0.7	0.4	0.4
Precision of difference								
Estimate minus true value	0.3	0.3	0.3	0.1	0.4	0.5	0.4	0.2
α for sample size								
Estimate minus true value	0.70	0.81	0.75	0.88	0.69	0.65	0.73	0.75

Precision is measured as 90-percent confidence intervals in units of +/- percentage points.
Scorecard is based on construction and calibration samples and is applied to the validation sample in the 2004 VHLSS.
Difference 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 12 (National poverty line): Differences and precision for bootstrapped estimates of groups' poverty rates 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)	Bias	90-percent	95-percent	99-percent
2	0.9	31.2	41.8	61.4
4	0.5	23.9	31.4	42.7
8	0.8	16.2	20.1	28.1
16	1.0	11.0	13.6	18.2
32	1.1	7.8	9.4	12.6
64	1.2	5.7	6.8	9.2
128	1.1	4.2	4.9	6.2
256	1.1	2.9	3.3	4.4
512	1.1	2.0	2.4	3.4
1,024	1.1	1.5	1.7	2.3
2,048	1.1	1.0	1.2	1.5
4,096	1.1	0.7	0.8	1.1
8,192	1.1	0.5	0.6	0.7
16,384	1.1	0.3	0.4	0.5

Figure 13 (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 14 (National poverty line): People 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	1.3	9.5	0.2	88.9	90.3	-73.5
5–9	2.9	8.0	0.6	88.5	91.4	-42.0
10–14	4.8	6.1	2.3	86.8	91.6	8.9
15–19	6.6	4.2	7.1	82.1	88.7	35.1
20–24	8.3	2.6	12.0	77.2	85.5	-10.0
25–29	9.5	1.4	19.4	69.7	79.2	-78.2
30–34	10.4	0.5	28.3	60.8	71.2	-160.2
35–39	10.7	0.2	35.8	53.3	64.0	-229.0
40–44	10.9	0.0	42.8	46.3	57.2	-293.5
45–49	10.9	0.0	50.8	38.4	49.2	-366.8
50–54	10.9	0.0	55.9	33.2	44.1	-413.9
55–59	10.9	0.0	61.4	27.7	38.6	-464.8
60–64	10.9	0.0	66.1	23.1	33.9	-507.3
65–69	10.9	0.0	70.9	18.2	29.1	-552.4
70–74	10.9	0.0	75.2	13.9	24.8	-591.3
75–79	10.9	0.0	79.0	10.1	21.0	-626.6
80–84	10.9	0.0	83.4	5.7	16.6	-666.9
85–89	10.9	0.0	86.5	2.6	13.5	-695.5
90–94	10.9	0.0	87.9	1.2	12.1	-708.2
95–100	10.9	0.0	89.1	0.0	10.9	-719.5

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

Figure 15 (National poverty line): People below the poverty line and all people, at a given score or at or below a given score cut-off, scorecard applied to validation sample

	People below poverty line (%)		All people (%)	
Score	**At score**	**At or below score**	**At score**	**At or below score**
0–4	88.0	88.0	1.5	1.5
5–9	78.8	82.9	1.9	3.5
10–14	53.3	67.8	3.6	7.1
15–19	28.0	48.5	6.6	13.7
20–24	25.4	41.0	6.6	20.3
25–29	13.6	32.9	8.6	28.9
30–34	9.2	26.9	9.8	38.7
35–39	4.0	23.0	7.8	46.5
40–44	2.2	20.2	7.2	53.7
45–49	0.2	17.6	8.0	61.6
50–54	0.0	16.3	5.1	66.8
55–59	0.0	15.0	5.5	72.3
60–64	0.0	14.1	4.6	76.9
65–69	0.0	13.3	4.9	81.8
70–74	0.0	12.6	4.2	86.1
75–79	0.0	12.1	3.8	89.9
80–84	0.0	11.5	4.4	94.3
85–89	0.0	11.2	3.1	97.4
90–94	0.0	11.0	1.4	98.8
95–100	0.0	10.9	1.2	100.0

National Food Poverty Line Tables

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

If an individual's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	87.8
5–9	63.6
10–14	56.1
15–19	23.9
20–24	15.5
25–29	9.8
30–34	6.3
35–39	4.0
40–44	0.5
45–49	0.0
50–54	0.0
55–59	0.0
60–64	0.0
65–69	0.0
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Surveyed cases weighted to represent Vietnam's population.

Based on the 2004 VLSS

Figure 6 (National food line): Illustration of derivation of estimated poverty likelihoods associated with scores

Score	People below poverty line		All people at score		Poverty likelihood (estimated, %)
0–4	5,602	÷	6,381	=	87.8
5–9	2,974	÷	4,677	=	63.6
10–14	3,657	÷	6,519	=	56.1
15–19	1,328	÷	5,557	=	23.9
20–24	1,140	÷	7,374	=	15.5
25–29	791	÷	8,070	=	9.8
30–34	412	÷	6,585	=	6.3
35–39	234	÷	5,842	=	4.0
40–44	33	÷	6,788	=	0.5
45–49	0	÷	5,294	=	0.0
50–54	0	÷	4,806	=	0.0
55–59	0	÷	5,057	=	0.0
60–64	0	÷	4,900	=	0.0
65–69	0	÷	4,669	=	0.0
70–74	0	÷	3,907	=	0.0
75–79	0	÷	2,650	=	0.0
80–84	0	÷	3,807	=	0.0
85–89	0	÷	4,168	=	0.0
90–94	0	÷	420	=	0.0
95–100	0	÷	2,526	=	0.0

Surveyed cases weighted to represent Vietnam's population.
Based on the 2004 VLSS

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

	Scorecard applied to validation sample difference between estimate and true value			
		Confidence interval (+/- percentage points)		
Score	**Bias**	**90-percent**	**95-percent**	**99-percent**
0–4	5.3	4.5	5.5	7.4
5–9	-1.0	4.9	6.0	8.2
10–14	15.7	3.6	4.3	5.9
15–19	0.2	2.3	2.6	3.6
20–24	-3.6	2.9	3.1	3.5
25–29	0.8	1.3	1.5	2.0
30–34	-0.6	1.2	1.5	2.1
35–39	0.6	1.0	1.2	1.6
40–44	-1.0	0.8	0.9	1.1
45–49	-0.2	0.2	0.2	0.2
50–54	0.0	0.0	0.0	0.0
55–59	0.0	0.0	0.0	0.0
60–64	0.0	0.0	0.0	0.0
65–69	0.0	0.0	0.0	0.0
70–74	0.0	0.0	0.0	0.0
75–79	0.0	0.0	0.0	0.0
80–84	0.0	0.0	0.0	0.0
85–89	0.0	0.0	0.0	0.0
90–94	0.0	0.0	0.0	0.0
95–100	0.0	0.0	0.0	0.0

Based on scorecard applied to validation sample.

Figure 10 (National food line): Differences and precision for bootstrapped estimates of individuals' poverty likelihoods, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/- percentage points)		
Sample size (n)	Bias	90-percent	95-percent	99-percent
2	-0.1	27.2	42.3	60.9
4	-0.3	22.4	29.9	41.9
8	0.0	16.8	20.4	29.3
16	0.3	11.1	13.9	19.0
32	0.8	8.6	10.1	14.6
64	0.7	6.7	8.2	10.8
128	0.9	5.3	6.3	8.0
256	0.8	3.7	4.5	5.7
512	0.8	2.4	2.9	4.3
1,024	0.9	1.8	2.1	2.8
2,048	0.9	1.2	1.4	1.8
4,096	0.8	0.9	1.0	1.3
8,192	0.8	0.6	0.7	1.0
16,384	0.8	0.4	0.5	0.7

Figure 12 (National food line): Differences and precision for bootstrapped estimates of groups' poverty rates 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)	Bias	90-percent	95-percent	99-percent
2	-0.1	27.2	42.3	60.9
4	-0.2	21.5	29.7	42.6
8	0.0	15.3	19.1	28.5
16	0.2	10.6	12.5	18.1
32	0.5	7.5	8.8	12.3
64	0.4	5.4	6.7	8.6
128	0.4	3.8	4.5	6.2
256	0.4	2.6	3.2	4.2
512	0.4	1.9	2.2	3.2
1,024	0.4	1.4	1.7	2.1
2,048	0.4	1.0	1.1	1.5
4,096	0.4	0.7	0.8	1.1
8,192	0.4	0.5	0.6	0.7
16,384	0.4	0.3	0.4	0.5

Figure 14 (National food line): People 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	1.3	7.4	0.3	91.1	92.4	-67.6
5–9	2.5	6.1	1.0	90.4	92.9	-30.9
10–14	4.0	4.7	3.1	88.3	92.2	27.9
15–19	5.5	3.1	8.2	83.2	88.7	5.1
20–24	6.8	1.8	13.5	77.9	84.7	-56.7
25–29	7.6	1.1	21.3	70.1	77.6	-147.3
30–34	8.2	0.4	30.5	60.9	69.2	-253.5
35–39	8.5	0.1	38.0	53.4	61.9	-340.9
40–44	8.6	0.0	45.1	46.3	54.9	-423.0
45–49	8.6	0.0	53.0	38.4	47.0	-515.5
50–54	8.6	0.0	58.1	33.2	41.9	-574.9
55–59	8.6	0.0	63.7	27.7	36.3	-639.2
60–64	8.6	0.0	68.3	23.1	31.7	-692.8
65–69	8.6	0.0	73.2	18.2	26.8	-749.7
70–74	8.6	0.0	77.4	13.9	22.6	-798.9
75–79	8.6	0.0	81.3	10.1	18.7	-843.4
80–84	8.6	0.0	85.7	5.7	14.3	-894.2
85–89	8.6	0.0	88.8	2.6	11.2	-930.4
90–94	8.6	0.0	90.2	1.2	9.8	-946.5
95–100	8.6	0.0	91.4	0.0	8.6	-960.7

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

Figure 15 (National food line): People below the poverty line and all people, at a given score or at or below a given score cut-off, scorecard applied to validation sample

	People below poverty line (%)		All people (%)	
Score	At score	At or below score	At score	At or below score
0–4	82.5	82.5	1.5	1.5
5–9	64.4	72.4	1.9	3.5
10–14	40.4	56.1	3.6	7.1
15–19	23.6	40.3	6.6	13.7
20–24	19.1	33.5	6.6	20.3
25–29	9.0	26.2	8.6	28.9
30–34	6.8	21.3	9.8	38.7
35–39	3.4	18.3	7.8	46.5
40–44	1.5	16.0	7.2	53.7
45–49	0.2	14.0	8.0	61.6
50–54	0.0	12.9	5.1	66.8
55–59	0.0	11.9	5.5	72.3
60–64	0.0	11.2	4.6	76.9
65–69	0.0	10.5	4.9	81.8
70–74	0.0	10.0	4.2	86.1
75–79	0.0	9.6	3.8	89.9
80–84	0.0	9.1	4.4	94.3
85–89	0.0	8.8	3.1	97.4
90–94	0.0	8.7	1.4	98.8
95–100	0.0	8.6	1.2	100.0

USAID "Extreme" Poverty Line Tables

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

If an individual's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	73.0
5–9	51.5
10–14	30.8
15–19	15.2
20–24	11.0
25–29	4.6
30–34	3.6
35–39	1.5
40–44	0.0
45–49	0.0
50–54	0.0
55–59	0.0
60–64	0.0
65–69	0.0
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Surveyed cases weighted to represent Vietnam's population.
Based on the 2004 VLSS

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

Score	People below poverty line		All people at score		Poverty likelihood (estimated, %)
0–4	4,661	÷	6,381	=	73.0
5–9	2,411	÷	4,677	=	51.5
10–14	2,009	÷	6,519	=	30.8
15–19	842	÷	5,557	=	15.2
20–24	808	÷	7,374	=	11.0
25–29	374	÷	8,070	=	4.6
30–34	236	÷	6,585	=	3.6
35–39	85	÷	5,842	=	1.5
40–44	0	÷	6,788	=	0.0
45–49	0	÷	5,294	=	0.0
50–54	0	÷	4,806	=	0.0
55–59	0	÷	5,057	=	0.0
60–64	0	÷	4,900	=	0.0
65–69	0	÷	4,669	=	0.0
70–74	0	÷	3,907	=	0.0
75–79	0	÷	2,650	=	0.0
80–84	0	÷	3,807	=	0.0
85–89	0	÷	4,168	=	0.0
90–94	0	÷	420	=	0.0
95–100	0	÷	2,526	=	0.0

Surveyed cases weighted to represent Vietnam's population.
Based on the 2004 VLSS

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

	Scorecard applied to validation sample difference between estimate and true value			
		Confidence interval (+/- percentage points)		
Score	**Bias**	**90-percent**	**95-percent**	**99-percent**
0–4	5.1	5.3	6.2	8.9
5–9	8.7	5.3	6.1	8.1
10–14	2.8	3.5	4.0	5.4
15–19	3.4	1.7	2.0	2.6
20–24	1.1	1.6	1.9	2.4
25–29	-0.3	1.0	1.2	1.5
30–34	0.4	0.8	1.0	1.3
35–39	0.2	0.7	0.8	1.1
40–44	-0.9	0.7	0.7	0.9
45–49	-0.2	0.2	0.2	0.2
50–54	0.0	0.0	0.0	0.0
55–59	0.0	0.0	0.0	0.0
60–64	0.0	0.0	0.0	0.0
65–69	0.0	0.0	0.0	0.0
70–74	0.0	0.0	0.0	0.0
75–79	0.0	0.0	0.0	0.0
80–84	0.0	0.0	0.0	0.0
85–89	0.0	0.0	0.0	0.0
90–94	0.0	0.0	0.0	0.0
95–100	0.0	0.0	0.0	0.0

Based on scorecard applied to validation sample.

Figure 10 (USAID "extreme" line): Differences and precision for bootstrapped estimates of individuals' poverty likelihoods, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/- percentage points)		
Sample size (n)	Bias	90-percent	95-percent	99-percent
2	-0.7	25.6	37.4	55.8
4	-0.3	18.7	24.2	39.3
8	0.0	14.0	18.1	25.7
16	0.3	9.8	12.1	16.4
32	0.6	7.5	8.9	12.9
64	0.7	6.1	7.5	10.1
128	1.0	5.2	6.2	7.9
256	1.0	3.8	4.5	6.0
512	1.0	2.5	3.0	4.2
1,024	1.0	1.8	2.1	2.9
2,048	1.0	1.2	1.5	1.9
4,096	1.0	0.9	1.1	1.4
8,192	1.0	0.7	0.8	1.0
16,384	1.0	0.4	0.5	0.7

Figure 12 (USAID "extreme" line): Differences and precision for bootstrapped estimates of groups' poverty rates 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)	Bias	90-percent	95-percent	99-percent
2	-0.7	25.6	37.4	55.8
4	-0.1	18.0	23.0	37.0
8	0.0	12.9	16.3	22.7
16	0.2	8.8	10.7	15.8
32	0.4	6.2	7.4	10.3
64	0.5	4.3	5.2	6.9
128	0.6	3.1	3.7	5.1
256	0.6	2.1	2.5	3.7
512	0.6	1.5	1.8	2.3
1,024	0.6	1.1	1.3	1.7
2,048	0.6	0.8	1.0	1.2
4,096	0.6	0.5	0.6	0.8
8,192	0.6	0.4	0.5	0.6
16,384	0.6	0.3	0.3	0.4

Figure 14 (USAID "extreme" line): People 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	1.0	4.2	0.5	94.3	95.3	-50.6
5–9	1.9	3.3	1.6	93.2	95.1	2.0
10–14	2.9	2.3	4.2	90.6	93.5	19.7
15–19	3.7	1.6	10.1	84.7	88.4	-92.9
20–24	4.3	0.9	16.0	78.8	83.1	-206.6
25–29	4.7	0.5	24.1	70.7	75.4	-363.2
30–34	5.0	0.2	33.7	61.1	66.2	-545.9
35–39	5.1	0.1	41.3	53.4	58.6	-693.5
40–44	5.2	0.0	48.5	46.3	51.5	-830.1
45–49	5.2	0.0	56.4	38.4	43.6	-983.2
50–54	5.2	0.0	61.6	33.2	38.4	-1,081.3
55–59	5.2	0.0	67.1	27.7	32.9	-1,187.6
60–64	5.2	0.0	71.7	23.1	28.3	-1,276.4
65–69	5.2	0.0	76.6	18.2	23.4	-1,370.4
70–74	5.2	0.0	80.9	13.9	19.1	-1,451.7
75–79	5.2	0.0	84.7	10.1	15.3	-1,525.2
80–84	5.2	0.0	89.1	5.7	10.9	-1,609.4
85–89	5.2	0.0	92.2	2.6	7.8	-1,669.2
90–94	5.2	0.0	93.6	1.2	6.4	-1,695.7
95–100	5.2	0.0	94.8	0.0	5.2	-1,719.2

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

Figure 15 (USAID "extreme" line): People below the poverty line and all people, at a given score or at or below a given score cut-off, scorecard applied to validation sample

	People below poverty line (%)		All people (%)	
Score	At score	At or below score	At score	At or below score
0–4	68.0	68.0	1.5	1.5
5–9	42.7	53.9	1.9	3.5
10–14	28.1	40.7	3.6	7.1
15–19	11.8	26.7	6.6	13.7
20–24	9.9	21.2	6.6	20.3
25–29	4.9	16.4	8.6	28.9
30–34	3.2	13.0	9.8	38.7
35–39	1.2	11.0	7.8	46.5
40–44	0.9	9.7	7.2	53.7
45–49	0.2	8.5	8.0	61.6
50–54	0.0	7.8	5.1	66.8
55–59	0.0	7.2	5.5	72.3
60–64	0.0	6.8	4.6	76.9
65–69	0.0	6.4	4.9	81.8
70–74	0.0	6.1	4.2	86.1
75–79	0.0	5.8	3.8	89.9
80–84	0.0	5.5	4.4	94.3
85–89	0.0	5.3	3.1	97.4
90–94	0.0	5.3	1.4	98.8
95–100	0.0	5.2	1.2	100.0

$1/Day Poverty Line Tables

Figure 5 ($1/day line): Estimated poverty likelihoods associated with scores

If an individual's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	24.7
5–9	2.8
10–14	3.1
15–19	0.0
20–24	0.7
25–29	0.0
30–34	0.0
35–39	0.0
40–44	0.0
45–49	0.0
50–54	0.0
55–59	0.0
60–64	0.0
65–69	0.0
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Surveyed cases weighted to represent Vietnam's population.
Based on the 2004 VLSS

Figure 6 ($1/day line): Illustration of derivation of estimated poverty likelihoods associated with scores

Score	People below poverty line		All people at score		Poverty likelihood (estimated, %)
0–4	1,577	÷	6,381	=	24.7
5–9	130	÷	4,677	=	2.8
10–14	200	÷	6,519	=	3.1
15–19	0	÷	5,557	=	0.0
20–24	52	÷	7,374	=	0.7
25–29	0	÷	8,070	=	0.0
30–34	0	÷	6,585	=	0.0
35–39	0	÷	5,842	=	0.0
40–44	0	÷	6,788	=	0.0
45–49	0	÷	5,294	=	0.0
50–54	0	÷	4,806	=	0.0
55–59	0	÷	5,057	=	0.0
60–64	0	÷	4,900	=	0.0
65–69	0	÷	4,669	=	0.0
70–74	0	÷	3,907	=	0.0
75–79	0	÷	2,650	=	0.0
80–84	0	÷	3,807	=	0.0
85–89	0	÷	4,168	=	0.0
90–94	0	÷	420	=	0.0
95–100	0	÷	2,526	=	0.0

Surveyed cases weighted to represent Vietnam's population.
Based on the 2004 VLSS

Figure 8 ($1/day line): Bootstrapped differences between estimated and true poverty likelihoods for individuals in a large sample (n=16,384) from the validation sample, with confidence intervals

	Scorecard applied to validation sample difference between estimate and true value			
		Confidence interval (+/- percentage points)		
Score	**Bias**	**90-percent**	**95-percent**	**99-percent**
0–4	8.5	4.5	5.4	6.9
5–9	-4.1	3.4	3.7	4.3
10–14	0.4	1.1	1.3	1.7
15–19	-0.3	0.3	0.3	0.4
20–24	0.1	0.4	0.4	0.6
25–29	-0.3	0.2	0.2	0.3
30–34	0.0	0.0	0.0	0.0
35–39	0.0	0.0	0.0	0.0
40–44	0.0	0.0	0.0	0.0
45–49	-0.2	0.2	0.2	0.2
50–54	0.0	0.0	0.0	0.0
55–59	0.0	0.0	0.0	0.0
60–64	0.0	0.0	0.0	0.0
65–69	0.0	0.0	0.0	0.0
70–74	0.0	0.0	0.0	0.0
75–79	0.0	0.0	0.0	0.0
80–84	0.0	0.0	0.0	0.0
85–89	0.0	0.0	0.0	0.0
90–94	0.0	0.0	0.0	0.0
95–100	0.0	0.0	0.0	0.0

Based on scorecard applied to validation sample.

Figure 10 ($1/day line): Differences and precision for bootstrapped estimates of individuals' poverty likelihoods, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/- percentage points)		
Sample size (n)	Bias	90-percent	95-percent	99-percent
2	-0.1	1.0	3.2	31.3
4	-0.2	0.8	9.0	22.6
8	0.0	3.6	6.9	13.0
16	0.0	3.8	5.5	9.8
32	0.0	3.5	4.9	6.3
64	0.1	3.3	4.2	5.8
128	0.1	2.8	3.6	4.7
256	0.2	2.2	2.8	4.0
512	0.2	1.6	1.9	2.5
1,024	0.2	1.1	1.3	1.7
2,048	0.2	0.8	0.9	1.1
4,096	0.2	0.6	0.7	0.8
8,192	0.2	0.4	0.5	0.6
16,384	0.2	0.3	0.3	0.4

Figure 12 ($1/day line): Differences and precision for bootstrapped estimates of groups' poverty rates 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)	Bias	90-percent	95-percent	99-percent
2	-0.1	1.0	3.2	31.3
4	-0.2	0.8	9.0	21.1
8	0.0	3.4	6.3	11.5
16	0.0	3.1	4.4	6.6
32	0.0	2.4	3.1	3.9
64	0.0	1.6	2.0	2.9
128	0.0	1.1	1.4	1.9
256	0.0	0.8	0.9	1.4
512	0.0	0.6	0.6	0.9
1,024	0.0	0.4	0.5	0.6
2,048	0.0	0.3	0.3	0.4
4,096	0.0	0.2	0.2	0.3
8,192	0.0	0.1	0.2	0.2
16,384	0.0	0.1	0.1	0.2

Figure 14 ($1/day line): People 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	0.2	0.3	1.3	98.1	98.4	-122.9
5–9	0.4	0.2	3.1	96.4	96.7	-433.8
10–14	0.5	0.1	6.6	92.8	93.3	-1,044.3
15–19	0.5	0.1	13.2	86.2	86.7	-2,196.1
20–24	0.5	0.0	19.7	79.7	80.2	-3,331.7
25–29	0.6	0.0	28.3	71.1	71.7	-4,818.6
30–34	0.6	0.0	38.1	61.3	61.9	-6,526.6
35–39	0.6	0.0	45.9	53.5	54.1	-7,880.3
40–44	0.6	0.0	53.1	46.3	46.9	-9,128.1
45–49	0.6	0.0	61.1	38.4	38.9	-10,514.0
50–54	0.6	0.0	66.2	33.2	33.8	-11,402.8
55–59	0.6	0.0	71.7	27.7	28.3	-12,365.4
60–64	0.6	0.0	76.4	23.1	23.6	-13,169.1
65–69	0.6	0.0	81.2	18.2	18.8	-14,020.4
70–74	0.6	0.0	85.5	13.9	14.5	-14,757.3
75–79	0.6	0.0	89.3	10.1	10.7	-15,422.8
80–84	0.6	0.0	93.7	5.7	6.3	-16,184.8
85–89	0.6	0.0	96.8	2.6	3.2	-16,726.1
90–94	0.6	0.0	98.2	1.2	1.8	-16,966.8
95–100	0.6	0.0	99.4	0.0	0.6	-17,179.1

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

Figure 15 ($1/day line): People below the poverty line and all people, at a given score or at or below a given score cut-off, scorecard applied to validation sample

	People below poverty line (%)		All people (%)	
Score	**At score**	**At or below score**	**At score**	**At or below score**
0–4	16.2	16.2	1.5	1.5
5–9	6.9	11.0	1.9	3.5
10–14	2.7	6.8	3.6	7.1
15–19	0.3	3.6	6.6	13.7
20–24	0.6	2.7	6.6	20.3
25–29	0.2	1.9	8.6	28.9
30–34	0.0	1.5	9.8	38.7
35–39	0.0	1.2	7.8	46.5
40–44	0.0	1.0	7.2	53.7
45–49	0.2	0.9	8.0	61.6
50–54	0.0	0.9	5.1	66.8
55–59	0.0	0.8	5.5	72.3
60–64	0.0	0.7	4.6	76.9
65–69	0.0	0.7	4.9	81.8
70–74	0.0	0.7	4.2	86.1
75–79	0.0	0.6	3.8	89.9
80–84	0.0	0.6	4.4	94.3
85–89	0.0	0.6	3.1	97.4
90–94	0.0	0.6	1.4	98.8
95–100	0.0	0.6	1.2	100.0

$2/Day Poverty Line Tables

Figure 5 ($2/day line): Estimated poverty likelihoods associated with scores

If an individual's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	90.9
5–9	73.4
10–14	65.6
15–19	34.7
20–24	25.3
25–29	17.5
30–34	9.7
35–39	5.5
40–44	1.6
45–49	1.5
50–54	0.8
55–59	0.0
60–64	0.0
65–69	0.0
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Surveyed cases weighted to represent Vietnam's population.

Based on the 2004 VLSS

Figure 6 ($2/day line): Illustration of derivation of estimated poverty likelihoods associated with scores

Score	People below poverty line		All people at score		Poverty likelihood (estimated, %)
0–4	5,799	÷	6,381	=	90.9
5–9	3,431	÷	4,677	=	73.4
10–14	4,277	÷	6,519	=	65.6
15–19	1,928	÷	5,557	=	34.7
20–24	1,863	÷	7,374	=	25.3
25–29	1,414	÷	8,070	=	17.5
30–34	635	÷	6,585	=	9.7
35–39	321	÷	5,842	=	5.5
40–44	107	÷	6,788	=	1.6
45–49	80	÷	5,294	=	1.5
50–54	36	÷	4,806	=	0.8
55–59	0	÷	5,057	=	0.0
60–64	0	÷	4,900	=	0.0
65–69	0	÷	4,669	=	0.0
70–74	0	÷	3,907	=	0.0
75–79	0	÷	2,650	=	0.0
80–84	0	÷	3,807	=	0.0
85–89	0	÷	4,168	=	0.0
90–94	0	÷	420	=	0.0
95–100	0	÷	2,526	=	0.0

Surveyed cases weighted to represent Vietnam's population.
Based on the 2004 VLSS

Figure 8 ($2/day line): Bootstrapped differences between estimated and true poverty likelihoods for individuals in a large sample (n=16,384) from the validation sample, with confidence intervals

	Scorecard applied to validation sample difference between estimate and true value			
		Confidence interval (+/- percentage points)		
Score	**Bias**	**90-percent**	**95-percent**	**99-percent**
0–4	-0.8	2.8	3.5	4.7
5–9	-7.1	5.5	5.9	6.3
10–14	8.5	3.7	4.4	5.7
15–19	5.4	2.4	2.9	3.9
20–24	-2.1	2.6	3.0	4.1
25–29	2.7	1.7	2.0	2.6
30–34	-0.1	1.4	1.6	2.3
35–39	1.0	1.1	1.3	1.8
40–44	-0.7	0.8	1.0	1.2
45–49	1.4	0.1	0.2	0.2
50–54	0.8	0.0	0.0	0.0
55–59	0.0	0.0	0.0	0.0
60–64	0.0	0.0	0.0	0.0
65–69	0.0	0.0	0.0	0.0
70–74	0.0	0.0	0.0	0.0
75–79	0.0	0.0	0.0	0.0
80–84	0.0	0.0	0.0	0.0
85–89	0.0	0.0	0.0	0.0
90–94	0.0	0.0	0.0	0.0
95–100	0.0	0.0	0.0	0.0

Based on scorecard applied to validation sample.

Figure 10 ($2/day line): Differences and precision for bootstrapped estimates of individuals' poverty likelihoods, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/- percentage points)		
Sample size (n)	Bias	90-percent	95-percent	99-percent
2	0.9	31.3	41.1	60.4
4	0.1	23.8	31.5	42.0
8	0.4	16.9	21.1	28.1
16	0.7	11.6	13.8	18.1
32	0.7	8.5	10.2	13.8
64	0.6	6.3	7.2	9.6
128	0.5	4.8	5.6	7.4
256	0.4	3.2	3.7	5.0
512	0.4	2.1	2.6	3.5
1,024	0.4	1.5	1.8	2.4
2,048	0.4	1.1	1.3	1.6
4,096	0.4	0.7	0.9	1.2
8,192	0.4	0.5	0.6	0.8
16,384	0.4	0.4	0.4	0.6

Figure 12 ($2/day line): Differences and precision for bootstrapped estimates of groups' poverty rates 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)	Bias	90-percent	95-percent	99-percent
2	0.9	31.3	41.1	60.4
4	0.2	23.8	31.7	41.8
8	0.5	16.4	20.0	28.1
16	0.7	11.0	13.7	18.3
32	0.8	7.8	9.6	13.6
64	0.8	5.7	7.1	9.2
128	0.8	4.2	4.9	6.7
256	0.8	2.9	3.5	4.4
512	0.8	2.0	2.5	3.2
1,024	0.8	1.5	1.7	2.2
2,048	0.8	1.1	1.2	1.5
4,096	0.8	0.7	0.8	1.1
8,192	0.8	0.5	0.6	0.8
16,384	0.8	0.4	0.4	0.5

Figure 14 ($2/day line): People 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	1.4	10.1	0.1	88.4	89.8	-74.5
5–9	2.9	8.6	0.5	88.0	90.9	-44.4
10–14	5.0	6.5	2.1	86.4	91.5	4.9
15–19	7.0	4.6	6.8	81.7	88.7	41.3
20–24	8.8	2.8	11.5	77.0	85.7	-0.2
25–29	10.0	1.5	18.8	69.7	79.7	-63.7
30–34	11.0	0.5	27.7	60.8	71.8	-140.7
35–39	11.3	0.2	35.1	53.3	64.7	-205.4
40–44	11.5	0.0	42.2	46.3	57.8	-266.3
45–49	11.5	0.0	50.1	38.4	49.9	-335.6
50–54	11.5	0.0	55.3	33.2	44.7	-380.1
55–59	11.5	0.0	60.8	27.7	39.2	-428.2
60–64	11.5	0.0	65.4	23.1	34.6	-468.4
65–69	11.5	0.0	70.3	18.2	29.7	-510.9
70–74	11.5	0.0	74.6	13.9	25.4	-547.8
75–79	11.5	0.0	78.4	10.1	21.6	-581.0
80–84	11.5	0.0	82.8	5.7	17.2	-619.1
85–89	11.5	0.0	85.9	2.6	14.1	-646.2
90–94	11.5	0.0	87.3	1.2	12.7	-658.2
95–100	11.5	0.0	88.5	0.0	11.5	-668.8

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

Figure 15 ($2/day line): People below the poverty line and all people, at a given score or at or below a given score cut-off, scorecard applied to validation sample

	People below poverty line (%)		All people (%)	
Score	**At score**	**At or below score**	**At score**	**At or below score**
0–4	91.7	91.7	1.5	1.5
5–9	80.4	85.4	1.9	3.5
10–14	57.2	71.0	3.6	7.1
15–19	29.2	50.7	6.6	13.7
20–24	27.4	43.2	6.6	20.3
25–29	14.8	34.7	8.6	28.9
30–34	9.8	28.4	9.8	38.7
35–39	4.5	24.4	7.8	46.5
40–44	2.2	21.4	7.2	53.7
45–49	0.2	18.7	8.0	61.6
50–54	0.0	17.2	5.1	66.8
55–59	0.0	15.9	5.5	72.3
60–64	0.0	15.0	4.6	76.9
65–69	0.0	14.1	4.9	81.8
70–74	0.0	13.4	4.2	86.1
75–79	0.0	12.8	3.8	89.9
80–84	0.0	12.2	4.4	94.3
85–89	0.0	11.8	3.1	97.4
90–94	0.0	11.7	1.4	98.8
95–100	0.0	11.5	1.2	100.0

$3/Day Poverty Line Tables

Figure 5 ($3/day line): Estimated poverty likelihoods associated with scores

If an individual's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	98.4
5–9	98.6
10–14	94.2
15–19	84.7
20–24	81.7
25–29	69.7
30–34	51.8
35–39	40.3
40–44	21.5
45–49	20.3
50–54	11.5
55–59	5.0
60–64	0.9
65–69	1.1
70–74	0.8
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Surveyed cases weighted to represent Vietnam's population.
Based on the 2004 VLSS

Figure 6 ($3/day line): Illustration of derivation of estimated poverty likelihoods associated with scores

Score	People below poverty line		All people at score		Poverty likelihood (estimated, %)
0–4	6,276	÷	6,381	=	98.4
5–9	4,611	÷	4,677	=	98.6
10–14	6,141	÷	6,519	=	94.2
15–19	4,708	÷	5,557	=	84.7
20–24	6,021	÷	7,374	=	81.7
25–29	5,622	÷	8,070	=	69.7
30–34	3,411	÷	6,585	=	51.8
35–39	2,356	÷	5,842	=	40.3
40–44	1,460	÷	6,788	=	21.5
45–49	1,075	÷	5,294	=	20.3
50–54	553	÷	4,806	=	11.5
55–59	254	÷	5,057	=	5.0
60–64	45	÷	4,900	=	0.9
65–69	50	÷	4,669	=	1.1
70–74	32	÷	3,907	=	0.8
75–79	0	÷	2,650	=	0.0
80–84	0	÷	3,807	=	0.0
85–89	0	÷	4,168	=	0.0
90–94	0	÷	420	=	0.0
95–100	0	÷	2,526	=	0.0

Surveyed cases weighted to represent Vietnam's population.
Based on the 2004 VLSS

Figure 8 ($3/day line): Bootstrapped differences between estimated and true poverty likelihoods for individuals in a large sample (n=16,384) from the validation sample, with confidence intervals

	Scorecard applied to validation sample difference between estimate and true value			
		Confidence interval (+/- percentage points)		
Score	**Bias**	**90-percent**	**95-percent**	**99-percent**
0–4	-1.6	0.8	0.8	0.8
5–9	-1.4	0.7	0.7	0.7
10–14	1.7	2.2	2.7	3.3
15–19	-2.9	2.4	2.6	3.0
20–24	5.0	2.3	2.7	3.5
25–29	3.5	2.2	2.6	3.6
30–34	3.8	2.2	2.7	3.5
35–39	3.3	2.2	2.7	3.5
40–44	-7.2	4.8	5.0	5.5
45–49	3.8	1.9	2.3	3.0
50–54	2.9	1.8	2.2	3.0
55–59	-1.1	1.4	1.7	2.3
60–64	-5.6	3.8	4.0	4.6
65–69	-0.9	1.0	1.1	1.5
70–74	0.8	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 validation sample.

Figure 10 ($3/day line): Differences and precision for bootstrapped estimates of individuals' poverty likelihoods, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/- percentage points)		
Sample size (n)	Bias	90-percent	95-percent	99-percent
2	0.5	44.2	54.2	69.0
4	-0.1	31.0	39.5	54.5
8	0.2	22.5	27.9	37.9
16	0.4	15.2	18.7	26.6
32	0.4	10.4	12.9	16.7
64	0.3	7.0	8.5	11.5
128	0.3	4.2	4.9	7.1
256	0.1	2.8	3.4	4.6
512	0.2	1.9	2.3	3.0
1,024	0.2	1.4	1.6	2.1
2,048	0.2	1.0	1.2	1.5
4,096	0.2	0.7	0.8	1.0
8,192	0.1	0.5	0.5	0.7
16,384	0.1	0.3	0.4	0.5

Figure 12 ($3/day line): Differences and precision for bootstrapped estimates of groups' poverty rates 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)	Bias	90-percent	95-percent	99-percent
2	0.5	44.2	54.2	69.0
4	-0.1	31.5	38.9	53.3
8	0.4	22.3	27.7	37.0
16	0.4	15.5	19.0	26.5
32	0.6	10.7	13.1	17.1
64	0.7	7.8	9.6	12.7
128	0.7	5.3	6.8	9.3
256	0.6	3.9	4.9	6.7
512	0.7	2.8	3.4	4.6
1,024	0.7	2.0	2.5	3.3
2,048	0.8	1.4	1.7	2.3
4,096	0.7	1.0	1.2	1.6
8,192	0.7	0.7	0.8	1.2
16,384	0.7	0.5	0.6	0.8

Figure 14 ($3/day line): People 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	1.5	34.0	0.0	64.5	66.0	-91.4
5–9	3.5	32.0	0.0	64.5	68.0	-80.6
10–14	6.8	28.7	0.3	64.2	71.0	-61.0
15–19	12.6	22.9	1.1	63.4	76.0	-25.8
20–24	17.7	17.8	2.6	61.9	79.5	6.9
25–29	23.3	12.2	5.5	59.0	82.3	47.0
30–34	28.1	7.4	10.6	53.9	81.9	70.0
35–39	30.9	4.6	15.5	49.0	79.9	56.2
40–44	33.0	2.5	20.7	43.8	76.9	41.8
45–49	34.3	1.2	27.3	37.2	71.5	23.0
50–54	34.8	0.7	32.0	32.5	67.3	9.8
55–59	35.1	0.4	37.2	27.3	62.4	-4.8
60–64	35.4	0.1	41.5	23.0	58.4	-17.0
65–69	35.5	0.0	46.3	18.2	53.7	-30.5
70–74	35.5	0.0	50.6	13.9	49.4	-42.5
75–79	35.5	0.0	54.4	10.1	45.6	-53.2
80–84	35.5	0.0	58.8	5.7	41.2	-65.6
85–89	35.5	0.0	61.9	2.6	38.1	-74.4
90–94	35.5	0.0	63.3	1.2	36.7	-78.3
95–100	35.5	0.0	64.5	0.0	35.5	-81.7

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

Figure 15 ($3/day line): People below the poverty line and all people, at a given score or at or below a given score cut-off, scorecard applied to validation sample

	People below poverty line (%)		All people (%)	
Score	**At score**	**At or below score**	**At score**	**At or below score**
0–4	100.0	100.0	1.5	1.5
5–9	100.0	100.0	1.9	3.5
10–14	92.5	96.2	3.6	7.1
15–19	87.6	92.0	6.6	13.7
20–24	76.7	87.1	6.6	20.3
25–29	66.2	80.8	8.6	28.9
30–34	48.0	72.5	9.8	38.7
35–39	37.1	66.6	7.8	46.5
40–44	28.8	61.5	7.2	53.7
45–49	16.5	55.7	8.0	61.6
50–54	8.6	52.1	5.1	66.8
55–59	6.1	48.5	5.5	72.3
60–64	6.6	46.0	4.6	76.9
65–69	1.9	43.4	4.9	81.8
70–74	0.0	41.2	4.2	86.1
75–79	0.0	39.5	3.8	89.9
80–84	0.0	37.7	4.4	94.3
85–89	0.0	36.4	3.1	97.4
90–94	0.0	35.9	1.4	98.8
95–100	0.0	35.5	1.2	100.0

Glewwe National Poverty Line Tables

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

If an individual's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	93.0
5–9	85.5
10–14	76.7
15–19	58.5
20–24	38.7
25–29	31.8
30–34	16.5
35–39	12.5
40–44	4.1
45–49	5.9
50–54	2.4
55–59	2.8
60–64	0.0
65–69	0.0
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Surveyed cases weighted to represent Vietnam's population.
Based on the 2004 VLSS

Figure 6 (Glewwe national poverty line): Illustration of derivation of estimated poverty likelihoods associated with scores

Score	People below poverty line		All people at score		Poverty likelihood (estimated, %)
0–4	5,935	÷	6,381	=	93.0
5–9	3,999	÷	4,677	=	85.5
10–14	5,000	÷	6,519	=	76.7
15–19	3,251	÷	5,557	=	58.5
20–24	2,853	÷	7,374	=	38.7
25–29	2,565	÷	8,070	=	31.8
30–34	1,084	÷	6,585	=	16.5
35–39	730	÷	5,842	=	12.5
40–44	276	÷	6,788	=	4.1
45–49	312	÷	5,294	=	5.9
50–54	114	÷	4,806	=	2.4
55–59	141	÷	5,057	=	2.8
60–64	0	÷	4,900	=	0.0
65–69	0	÷	4,669	=	0.0
70–74	0	÷	3,907	=	0.0
75–79	0	÷	2,650	=	0.0
80–84	0	÷	3,807	=	0.0
85–89	0	÷	4,168	=	0.0
90–94	0	÷	420	=	0.0
95–100	0	÷	2,526	=	0.0

Surveyed cases weighted to represent Vietnam's population.
Based on the 2004 VLSS

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

	Scorecard applied to validation sample difference between estimate and true value			
		Confidence interval (+/- percentage points)		
Score	**Bias**	**90-percent**	**95-percent**	**99-percent**
0–4	-0.9	2.4	2.8	4.0
5–9	-2.9	3.0	3.6	5.0
10–14	2.5	3.2	4.1	5.5
15–19	3.1	2.7	3.3	4.1
20–24	-5.8	4.4	4.6	5.0
25–29	2.4	2.3	2.6	3.4
30–34	-0.5	1.6	2.0	2.6
35–39	1.4	1.5	1.8	2.6
40–44	-2.7	2.1	2.2	2.5
45–49	4.1	0.6	0.7	0.9
50–54	1.1	0.9	1.1	1.3
55–59	1.3	0.6	0.8	1.0
60–64	0.0	0.0	0.0	0.0
65–69	0.0	0.0	0.0	0.0
70–74	0.0	0.0	0.0	0.0
75–79	0.0	0.0	0.0	0.0
80–84	0.0	0.0	0.0	0.0
85–89	0.0	0.0	0.0	0.0
90–94	0.0	0.0	0.0	0.0
95–100	0.0	0.0	0.0	0.0

Based on scorecard applied to validation sample.

Figure 10 (Glewwe national poverty line): Differences and precision for bootstrapped estimates of individuals' poverty likelihoods, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/- percentage points)		
Sample size (n)	Bias	90-percent	95-percent	99-percent
2	0.4	38.2	46.4	64.5
4	0.2	27.9	33.7	44.8
8	0.4	20.2	24.9	31.0
16	0.3	13.4	16.7	22.3
32	0.4	9.8	11.7	16.1
64	0.3	6.6	7.8	10.9
128	0.2	4.4	5.4	7.7
256	0.1	3.1	3.7	5.0
512	0.2	2.1	2.4	3.2
1,024	0.2	1.5	1.8	2.3
2,048	0.2	1.1	1.3	1.7
4,096	0.2	0.7	0.9	1.2
8,192	0.2	0.5	0.6	0.8
16,384	0.2	0.4	0.4	0.6

Figure 12 (Glewwe national poverty line): Differences and precision for bootstrapped estimates of groups' poverty rates 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)	Bias	90-percent	95-percent	99-percent
2	0.4	38.2	46.4	64.5
4	0.1	28.1	33.8	44.1
8	0.4	19.5	23.0	31.6
16	0.4	13.7	16.3	22.4
32	0.6	9.8	11.6	16.4
64	0.6	6.7	8.6	11.1
128	0.5	4.8	5.9	7.6
256	0.4	3.6	4.2	5.4
512	0.5	2.4	2.8	3.9
1,024	0.4	1.7	2.1	2.8
2,048	0.4	1.3	1.5	2.2
4,096	0.4	0.9	1.0	1.4
8,192	0.4	0.6	0.7	0.9
16,384	0.4	0.4	0.5	0.7

Figure 14 (Glewwe national poverty line): People 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	1.4	16.8	0.1	81.7	83.1	-83.7
5–9	3.1	15.1	0.3	81.4	84.6	-63.9
10–14	5.8	12.4	1.3	80.5	86.3	-29.4
15–19	9.5	8.7	4.2	77.5	87.0	27.2
20–24	12.4	5.8	7.9	73.9	86.3	56.9
25–29	14.9	3.3	13.9	67.8	82.8	23.6
30–34	16.6	1.6	22.1	59.7	76.3	-21.1
35–39	17.5	0.8	29.0	52.7	70.2	-59.1
40–44	18.0	0.3	35.7	46.1	64.0	-95.8
45–49	18.1	0.1	43.6	38.2	56.3	-138.8
50–54	18.2	0.1	48.6	33.2	51.3	-166.5
55–59	18.2	0.0	54.1	27.7	45.9	-196.4
60–64	18.2	0.0	58.7	23.1	41.3	-221.8
65–69	18.2	0.0	63.6	18.2	36.4	-248.6
70–74	18.2	0.0	67.8	13.9	32.2	-271.9
75–79	18.2	0.0	71.7	10.1	28.3	-292.9
80–84	18.2	0.0	76.0	5.7	24.0	-316.9
85–89	18.2	0.0	79.2	2.6	20.8	-334.0
90–94	18.2	0.0	80.5	1.2	19.5	-341.6
95–100	18.2	0.0	81.8	0.0	18.2	-348.3

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

Figure 15 (Glewwe national poverty line): People below the poverty line and all people, at a given score or at or below a given score cut-off, scorecard applied to validation sample

	People below poverty line (%)		All people (%)	
Score	**At score**	**At or below score**	**At score**	**At or below score**
0–4	94.0	94.0	1.5	1.5
5–9	88.3	90.8	1.9	3.5
10–14	74.1	82.3	3.6	7.1
15–19	55.4	69.3	6.6	13.7
20–24	44.5	61.2	6.6	20.3
25–29	29.3	51.7	8.6	28.9
30–34	16.9	42.9	9.8	38.7
35–39	11.1	37.6	7.8	46.5
40–44	6.8	33.5	7.2	53.7
45–49	1.8	29.3	8.0	61.6
50–54	1.3	27.2	5.1	66.8
55–59	1.4	25.2	5.5	72.3
60–64	0.0	23.7	4.6	76.9
65–69	0.0	22.3	4.9	81.8
70–74	0.0	21.2	4.2	86.1
75–79	0.0	20.3	3.8	89.9
80–84	0.0	19.3	4.4	94.3
85–89	0.0	18.7	3.1	97.4
90–94	0.0	18.5	1.4	98.8
95–100	0.0	18.2	1.2	100.0

Glewwe Food Poverty Line Tables

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

If an individual's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	47.1
5–9	21.0
10–14	13.8
15–19	2.2
20–24	1.5
25–29	0.9
30–34	0.6
35–39	0.8
40–44	0.0
45–49	0.0
50–54	0.0
55–59	0.0
60–64	0.0
65–69	0.0
70–74	0.0
75–79	0.0
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Surveyed cases weighted to represent Vietnam's population.
Based on the 2004 VLSS

Figure 6 (Glewwe food poverty line): Illustration of derivation of estimated poverty likelihoods associated with scores

Score	People below poverty line		All people at score		Poverty likelihood (estimated, %)
0–4	3,002	÷	6,381	=	47.1
5–9	983	÷	4,677	=	21.0
10–14	901	÷	6,519	=	13.8
15–19	123	÷	5,557	=	2.2
20–24	113	÷	7,374	=	1.5
25–29	74	÷	8,070	=	0.9
30–34	41	÷	6,585	=	0.6
35–39	46	÷	5,842	=	0.8
40–44	0	÷	6,788	=	0.0
45–49	0	÷	5,294	=	0.0
50–54	0	÷	4,806	=	0.0
55–59	0	÷	5,057	=	0.0
60–64	0	÷	4,900	=	0.0
65–69	0	÷	4,669	=	0.0
70–74	0	÷	3,907	=	0.0
75–79	0	÷	2,650	=	0.0
80–84	0	÷	3,807	=	0.0
85–89	0	÷	4,168	=	0.0
90–94	0	÷	420	=	0.0
95–100	0	÷	2,526	=	0.0

Surveyed cases weighted to represent Vietnam's population.
Based on the 2004 VLSS

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

	Scorecard applied to validation sample difference between estimate and true value			
		Confidence interval (+/- percentage points)		
Score	**Bias**	**90-percent**	**95-percent**	**99-percent**
0–4	15.8	5.4	6.4	9.2
5–9	9.1	3.5	4.0	5.3
10–14	2.2	2.3	2.8	3.5
15–19	-1.3	1.2	1.3	1.5
20–24	-0.7	0.8	0.9	1.2
25–29	0.3	0.2	0.3	0.4
30–34	-0.2	0.4	0.5	0.7
35–39	0.8	0.0	0.0	0.0
40–44	-0.4	0.3	0.4	0.5
45–49	-0.2	0.2	0.2	0.2
50–54	0.0	0.0	0.0	0.0
55–59	0.0	0.0	0.0	0.0
60–64	0.0	0.0	0.0	0.0
65–69	0.0	0.0	0.0	0.0
70–74	0.0	0.0	0.0	0.0
75–79	0.0	0.0	0.0	0.0
80–84	0.0	0.0	0.0	0.0
85–89	0.0	0.0	0.0	0.0
90–94	0.0	0.0	0.0	0.0
95–100	0.0	0.0	0.0	0.0

Based on scorecard applied to validation sample.

Figure 10 (Glewwe food poverty line): Differences and precision for bootstrapped estimates of individuals' poverty likelihoods, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/- percentage points)		
Sample size (n)	Bias	90-percent	95-percent	99-percent
2	-0.1	5.1	24.3	47.2
4	-0.3	11.7	17.2	31.0
8	-0.1	9.3	12.2	19.0
16	0.1	6.8	9.3	13.7
32	0.4	5.4	7.0	10.1
64	0.7	4.9	5.9	8.0
128	1.1	4.1	5.0	6.9
256	1.3	3.1	3.7	4.7
512	1.3	2.1	2.5	3.2
1,024	1.3	1.5	1.7	2.2
2,048	1.3	1.0	1.2	1.7
4,096	1.3	0.8	0.9	1.1
8,192	1.3	0.5	0.6	0.8
16,384	1.3	0.4	0.4	0.6

Figure 12 (Glewwe food poverty line): Differences and precision for bootstrapped estimates of groups' poverty rates 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)	Bias	90-percent	95-percent	99-percent
2	-0.1	5.1	24.3	47.2
4	-0.2	11.7	15.9	27.8
8	-0.1	8.8	11.3	16.7
16	0.1	5.7	7.1	10.6
32	0.2	4.0	4.9	6.9
64	0.3	2.8	3.3	4.9
128	0.4	2.0	2.4	3.4
256	0.4	1.3	1.6	2.1
512	0.4	1.0	1.1	1.5
1,024	0.4	0.7	0.8	1.1
2,048	0.4	0.5	0.6	0.7
4,096	0.4	0.3	0.4	0.5
8,192	0.4	0.2	0.3	0.4
16,384	0.4	0.2	0.2	0.3

Figure 14 (Glewwe food poverty line): People 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	0.5	1.2	1.1	97.3	97.7	19.5
5–9	0.7	1.0	2.7	95.6	96.3	-63.2
10–14	1.1	0.6	5.9	92.4	93.5	-252.8
15–19	1.4	0.3	12.4	86.0	87.3	-634.3
20–24	1.5	0.2	18.8	79.5	81.0	-1,016.4
25–29	1.6	0.1	27.3	71.0	72.6	-1,523.3
30–34	1.6	0.0	37.1	61.3	62.9	-2,102.8
35–39	1.6	0.0	44.8	53.5	55.1	-2,565.9
40–44	1.7	0.0	52.0	46.3	48.0	-2,991.2
45–49	1.7	0.0	60.0	38.4	40.0	-3,465.3
50–54	1.7	0.0	65.1	33.2	34.9	-3,769.4
55–59	1.7	0.0	70.6	27.7	29.4	-4,098.6
60–64	1.7	0.0	75.2	23.1	24.8	-4,373.6
65–69	1.7	0.0	80.1	18.2	19.9	-4,664.8
70–74	1.7	0.0	84.4	13.9	15.6	-4,916.9
75–79	1.7	0.0	88.2	10.1	11.8	-5,144.6
80–84	1.7	0.0	92.6	5.7	7.4	-5,405.3
85–89	1.7	0.0	95.7	2.6	4.3	-5,590.4
90–94	1.7	0.0	97.1	1.2	2.9	-5,672.8
95–100	1.7	0.0	98.3	0.0	1.7	-5,745.4

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

Figure 15 (Glewwe food poverty line): People below the poverty line and all people, at a given score or at or below a given score cut-off, scorecard applied to validation sample

	People below poverty line (%)		All people (%)	
Score	At score	At or below score	At score	At or below score
0–4	31.3	31.3	1.5	1.5
5–9	11.9	20.5	1.9	3.5
10–14	11.7	16.0	3.6	7.1
15–19	3.5	9.9	6.6	13.7
20–24	2.3	7.4	6.6	20.3
25–29	0.6	5.4	8.6	28.9
30–34	0.8	4.2	9.8	38.7
35–39	0.0	3.5	7.8	46.5
40–44	0.4	3.1	7.2	53.7
45–49	0.2	2.7	8.0	61.6
50–54	0.0	2.5	5.1	66.8
55–59	0.0	2.3	5.5	72.3
60–64	0.0	2.2	4.6	76.9
65–69	0.0	2.1	4.9	81.8
70–74	0.0	2.0	4.2	86.1
75–79	0.0	1.9	3.8	89.9
80–84	0.0	1.8	4.4	94.3
85–89	0.0	1.7	3.1	97.4
90–94	0.0	1.7	1.4	98.8
95–100	0.0	1.7	1.2	100.0