Simple Poverty Scorecard® Poverty-Assessment Tool Nepal

Dean Caire, Mark Schreiner, Shiyuan Chen, and Gary Woller

24 February 2009

This document is at `SimplePovertyScorecard.com`

Abstract

The Simple Poverty Scorecard-brand poverty-assessment tool uses 10 low-cost indicators from Nepal's 2003/4 Nepal 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 Nepal 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 CGAP/Ford Social Indicators Project. The data come from Nepal's Central Bureau of Statistics. Thanks go to Malika Anand, Shaohua Chen, Frank DeGiovanni, Syed Hashemi, Vimala Palaniswamy, Prem Sangraula, Shib Nandan Prasad Shah, Tony Sheldon, and Don Sillers. "Simple Poverty Scorecard" is a Registered Trademark of Microfinance Risk Management, L.L.C. for its brand of poverty-assessment tools.

Simple Poverty Scorecard® Poverty-Assessment Tool

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

Indicator	Response	Points	Score
1. How many household members are 12-years-old or younger?	A. Four or more	0	
	B. Three	6	
	C. Two	12	
	D. One	16	
	E. None	28	
2. How many children ages 5 to 12 attend school?	A. Not all	0	
	B. No children ages 5 to 12	2	
	C. All	5	
3. Do any household members attend a private school?	A. No	0	
	B. Yes	8	
4. Does the female head/spouse know how to read a letter?	A. No	0	
	B. No female head/spouse	5	
	C. Yes	8	
5. What is the main material of the roof?	A. Straw/thatch, wood/planks, earth/mud, or other	0	
	B. Tiles/slate	4	
	C. Galvanized iron, or concrete/cement	10	
6. What type of toilet is used in your household?	A. No toilet	0	
	B. Household non-flush, communal latrine, household flush (connected to municipal sewer), or household flush (connected to septic tank)	7	
7. What type of stove does your household mainly use for cooking?	A. Open fireplace, other, or no data	0	
	B. Mud stove, smokeless stove, or kerosene/gas stove	5	
8. How many radio/tape/CD players does the household own?	A. None	0	
	B. One	6	
	C. Two or more	13	
9. Does the household own any bicycles, motorcycles/scooters, motor cars, etc.	A. No	0	
	B. Yes	5	
10. Does the household own any televisions/decks?	A. No	0	
	B. Yes	11	

Score:

Simple Poverty Scorecard® Poverty-Assessment Tool Nepal

1. Introduction

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

The direct approach to poverty measurement via consumption surveys is difficult and costly, asking households about a lengthy list of items such as "Over the past 12 months, did you purchase or receive in-kind any wood (bundled wood, logwood, or sawdust)? If yes, what is the money value of the amount purchased or received in-kind by your household during the past 10 days? During the past 12 months? Now then, over the past 12 months, did you purchase or receive in-kind any kerosene oil? . . ."

In contrast, the indirect approach via the scorecard is simple, quick, and inexpensive. It uses 10 verifiable indicators (such as "What type of stove does your household mainly use for cooking?" or "What is the main material of the roof?") to get a score that is highly correlated with poverty status as measured by the exhaustive consumption survey.

The scorecard here differs from "proxy means tests" (Coady, Grosh, and Hoddinott, 2002) in that it is tailored to the capabilities and purposes not of national governments but rather of local, pro-poor organizations. The feasible poverty-measurement options for these organizations are typically subjective and relative (such as participatory wealth ranking by skilled field workers) or blunt (such as rules based on land-ownership or housing quality). Results from these approaches are not comparable across organizations nor across countries, they may be costly, and their accuracy is unknown.

If an organization wants to know what share of its participants are below a poverty line (say, $1.25/day 2005 PPP for the Millennium Development Goals, or the poorest half below the national poverty line as required of USAID microenterprise grantees), or if it wants to measure movement across a poverty line (for example, to report to the Microcredit Summit Campaign), then it needs a consumption-based, objective tool with known accuracy. Even though consumption surveys are costly even for governments, many small, local organizations can still implement an inexpensive poverty-assessment tools 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, many decimal places, and standard errors). Thanks to the predictive-modeling phenomenon known as the "flat max", simple, transparent scorecards can be almost as accurate as complex, opaque ones.

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

The scorecard is based on the 2003/4 Nepal Living Standards Survey (NLSS).[1] Indicators are selected to be:

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

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

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

[1] The NLSS was planned for 2002/3, but it took place in 2003/4.

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

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

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

This paper presents a single scorecard whose indicators and points are derived from Nepal household consumption data and Nepal's national poverty line. Scores from this scorecard are calibrated to poverty likelihoods for nine poverty lines.

The scorecard is constructed using a sub-sample of the NLSS data. Its accuracy is validated on a different sub-sample from the NLSS. 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.[2] 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 0.8 percentage points for the national line, and averages about ±0.9 percentage points across all nine lines. This difference is due to sampling variation and not bias; the average difference would be zero if the whole NLSS were to be repeatedly redrawn and divided into sub-samples before repeating the entire process of scorecard-building, calibration, and validation.

For sample sizes of n = 16,384, the 90-percent confidence intervals for these estimates are ±0.5 percentage points or less. For n = 1,024, the 90-percent intervals are ±2.2 percentage points or less.

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

[2] Bias may also result from changes in the form or quality of data collection, from imperfect adjustment of poverty lines across time or geographic regions, from sampling variation across consumption surveys, or, in general, from applying the scorecard to a group of households that is not representative of all of Nepal.

2. Data and poverty lines

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

2.1 Data

The scorecard is based on data from the 2003/4 NLSS. Households are randomly divided into three sub-samples (Figure 2):

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

2.2 Poverty rates and poverty lines

2.2.1 Rates

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

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

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

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

If an organization has only one "participant" per household, however, then the household-level rate is relevant. For example, if a microlender has only one borrower in a household, then it could report household-level poverty rates.

This paper reports poverty rates and poverty lines at both the household-level and the person-level for each of Nepal's six regions for 2003/4 (Figure 3).[3] The

[3] The NLSS was in the field from April 2003 to April 2004 (Central Bureau of Statistics, 2004). In calculating the $1.25/day 2005 PPP poverty line, the World Bank (Prem Sangraula, personal communication) used as a price deflator the average Consumer Price Index for 2003 and 2004, effectively assuming that the NLSS was fielded from January 2003 to December 2004, and that the monetary data in the NLSS database is

scorecard is constructed using the 2003/4 NLSS and household-level lines, scores are calibrated to household-level poverty likelihoods, and accuracy is measured for household-level rates. This focus on the household level reflects the belief that it is the relevant level for most pro-poor organizations.

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

2.2.2 Lines

The national poverty line is 21.08 Nepalese Rupees (NRP) per person per day (Central Bureau of Statistics, 2005). This is the person-weighted average of the cost-of-living-adjusted lines for Nepal's six regions (Figure 3).

The national line is defined as the food line plus a non-food component. The food line of NRP13.59 is defined as the cost of 2,124 calories (Central Bureau of Statistics, 2005). The non-food component (NRP7.49) is the average consumption in the NLSS for "households whose food consumption was around the requisite food poverty line" (Central Bureau of Statistics, 2005, p. 48).

The scorecard here is constructed using the national poverty line. Because local pro-poor organizations may want to use different or various poverty lines, this paper

---

delivered in nominal units. This paper maintains that assumption, in part because monthly CPI data for Nepal are not available.

calibrates scores from its single scorecard to poverty likelihoods for nine lines (figures in parentheses below are per-capita per day poverty lines at the person level, followed by household-level and person-level poverty rates):

- National line (21.08, 25.9 percent, 30.8 percent)
- Food line (13.61, 6.8 percent, 8.3 percent)
- USAID "extreme" line (16.40, 12.3 percent, 15.3 percent)
- $1.08/day 1993 PPP (18.16, 17.1 percent, 20.5 percent)
- $2.16/day 1993 PPP (36.33, 58.4 percent, 63.6 percent)
- $3.24/day 1993 PPP (54.49, 78.4 percent, 82.1 percent)
- $1.25/day 2005 PPP (30.30, 47.8 percent, 53.6 percent)
- $2.50/day 2005 PPP (60.60, 81.8 percent, 85.1 percent)
- $3.75/day 2005 PPP (90.90, 90.6 percent, 92.2 percent)

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

The person-weighted $1.08/day 1993 PPP line for all-Nepal is NRP18.16.[4] This paper adjusts this for regional differences in cost-of-living using:

- L, the all-Nepal $1.08/day 1993 PPP line of NRP18.16
- p_i, population proportions by region ($i = 1$ to 6)
- π_i, regional price deflators (national poverty lines for each region from Central Bureau of Statistics, 2005)

L_i is then the $1.08/day 1993 PPP line adjusted for the cost-of-living in region i:

$$L_i = \frac{L \cdot \pi_i}{\sum_{j=1}^{6} p_j \cdot \pi_j}.$$

[4] We thank Prem Sangraula for this figure and for the $1.25/day 2005 PPP line for Nepal. Sillers (2006) and Ravallion, Chen, and Sangraula (2008) explain the concepts and calculations behind the 1993 and 2005 PPP lines.

The all-Nepal line L is the person-weighted average of regional lines L_i. The \$2.16/day and \$3.24/day 1993 PPP lines are multiples of the \$1.08/day 1993 PPP lines.

The \$1.25/day 2005 PPP line is converted into regional cost-of-living-adjusted lines in the same way, and of course the \$2.50/day and \$3.75/day lines are multiples of the regional \$1.25/day 2005 PPP lines.

3. Context of poverty-assessment tools for Nepal

This section reviews two poverty-assessment tools Nepal. The main aspects of interest are the goals, methods, indicators, and relative/absolute poverty estimation.

3.1 Central Bureau of Statistics

Central Bureau of Statistics (2005, "CBS")[5] uses ordinary least-squares regression on the logarithm of per-capita consumption to associate poverty with these indicators:

- Household size:
  - All members of any age
  - Members ages 0 to 6
- Characteristics of the household head:
  - Sex
  - Age
  - Caste/ethnicity
  - Education
  - Employment:
    - Sector
    - Type (wage-employed or self-employed)
- Land ownership

CBS' goal is to "simulate the effect of a change in characteristics on the probability of being poor" (p. 22), reporting, for example, the average percentage change in the national poverty rate if all households went from having no small children to having one small child.

[5] See also World Bank (2006).

CBS finds that greater poverty is associated with more small children, larger households, female heads, younger heads, lower castes, lower education, agriculture, non-wage employment, and less land ownership. Other than to reinforce awareness of poverty's incidence among various groups, the policy usefulness of this "poverty profile" is unclear; the results are mostly common sense, and governments usually cannot create or destroy wage jobs or small children, except perhaps in the long term by doing things (such as stabilizing the macroeconomy) that they should already be doing anyway.

3.2 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 1995 NLSS 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
  - Television
  - Refrigerator
  - Bicycle
  - Motorized transport
- Household quality:
  - Source of drinking water
  - Type of toilet facility
  - Quality of construction material of floor
- Human capital (education of the household head)

To check coherency between the asset index and reported consumption[6] (or child nutrition), Sahn and Stifel rank Nepalese 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.[7] 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 consumption worse than does predicted consumption from a least-squares regression on household demographics, education, housing quality, and access to public services.

Sahn and Stifel's asset index is close kin to the scorecard here except that it is built without directly taking consumption into account.[8] Other examples of the asset-index approach are Stifel and Christiaensen (2007), Zeller *et al.* (2006), Montgomery *et al.* (2000), Sahn and Stifle (2000), and Filmer and Pritchett (2001). In general, asset indices rank households by poverty status about as well as the scorecard here, largely because of the predictive modeling phenomenon known as a the "flat max".

A strength of asset indices is that, because they do not require consumption data, they can be applied to a wide array of "light" surveys such as censuses, Demographic and Health Surveys, Welfare Monitoring Surveys, and Core Welfare

[6] They check the index against consumption because it is a common proxy for living standards, not because they believe consumption should be the benchmark.
[7] For measuring targeting accuracy, an analysis akin Figure 15 is more useful.
[8] Still, because their indicators are so similar, the PCA-based index and income-based poverty-assessment tools probably pick up the same underlying construct (such as "permanent income", see Bollen, Glanville, and Stecklov, 2007) and rank households much the same.

Indicator Questionnaires. Of course, the flip side is that, without consumption data, they can only rank households and thus provide only relative—not absolute—measures of poverty. Thus, while Sahn and Stifel's asset index can be used for targeting, it cannot estimate households' poverty likelihoods or groups' poverty rates.

4. Scorecard construction

About 70 potential indicators are initially prepared in the areas of:

- Family composition (such as household size and female headship)
- Education (such as school attendance of children)
- Housing (such as main material of the roof)
- Ownership of durable goods (such as bicycles, televisions, and CD players)

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 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 bicycle or television is probably more likely to change in response to changes in poverty than is the education of the male head/spouse.

The scorecard itself is built using Logit regression on the construction sub-sample (Figure 2). Indicator selection uses both judgment and statistics (forward stepwise). 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 Nepal. 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 maximize accuracy but rather to improve the chances that scoring is actually used (Schreiner, 2005b). When scoring projects fail, the reason is not usually technical inaccuracy but rather the failure of an organization to decide to do what is needed to integrate scoring in its processes and to learn to use it properly (Schreiner, 2002). After all, most reasonable scorecards predict tolerably well, thanks to the empirical phenomenon known as the "flat max" (Hand, 2006; Baesens *et al.*, 2003; Lovie and Lovie, 1986; Kolesar and Showers, 1985; Stillwell, Hutton, and Edwards, 1983; Dawes, 1979; Wainer, 1976; Myers and Forgy, 1963). The bottleneck is less technical and more human, not statistics but organizational change management. Accuracy is easier to achieve than adoption.

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

To this end, the scorecard here fits on one page. The construction process, indicators, and points are simple and transparent. "Extra" work is minimized; non-specialists can compute scores 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 in the far-right column
- Add up the points to get the total score
- Implement targeting policy (if any)
- Deliver the paper scorecard to a central office for filing or data entry

Of course, field workers must be trained. Quality outputs depend on quality inputs. If organizations or field workers gather their own data and have an incentive to exaggerate poverty rates (for example, if they are rewarded for higher poverty rates), then it is wise to do on-going quality control via data review and random audits (Matul and Kline, 2003).[9] 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 particular, while collecting scorecard indicators is relatively easier than alternatives, it is still absolutely difficult. Training and careful communication of the definitions of terms and concepts in the scorecard is essential. For the example of Nigeria, Onwujekwe, Hanson, and Fox-Rushby (2006) find distressingly low inter-rater and test-retest correlations for indicators as seemingly simple as whether the household owns an automobile. In contrast, Martinelli and Parker (2006) find that interviewer

[9] If an organization does not want field workers to know the points associated with indicators, then they can use the version of Figure 1 without points and apply the points later in a spreadsheet or database at the central office.

errors and respondent lies had negligible effects on targeting accuracy in a Mexican program. For now, it is unknown whether these results are universal or country-specific.

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

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

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

- Employees of the organization
- Third-party contractors

Responses, scores, and poverty likelihoods can be recorded:

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

The subjects to be scored can be:

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

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

Frequency of application can be:

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

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

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

An example set of choices were made by BRAC and ASA, two microlenders in Bangladesh (each with 7 million participants) who are applying the scorecard (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). Responses are recorded on paper in the field before being sent to a central office to be entered into a database. ASA's and BRAC's sampling plans cover 50,000–100,000 participants each.

6. Estimates of household poverty likelihoods

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

To get absolute units, scores must be converted to *poverty likelihoods*, that is, probabilities of being below a poverty line. This is done via simple look-up tables. For the example of the national line, scores of 10–14 have a poverty likelihood of 62.0 percent, and scores of 40–44 have a poverty likelihood of 12.9 percent (Figure 5).

The poverty likelihood associated with a score varies by poverty line. For example, scores of 40–44 are associated with a poverty likelihood of 12.9 percent for the national line but 3.4 percent for the food line.[10]

6.1 Calibrating scores with poverty likelihoods

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

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

For the example of the national line (Figure 6), there are 8,719 households in the calibration sub-sample with a score of 20–24, of whom 3,744 are below the poverty line. The estimated poverty likelihood associated with a score of 20–24 is then 42.9 percent, because 3,744 ÷ 8,719 = 42.9 percent.

To illustrate with the national line and a score of 40–44, there are 8,176 households in the calibration sample, of whom 1,056 are below the line (Figure 6). Thus, the poverty likelihood for this score is 1,056 ÷ 8,176 = 12.9 percent.

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

Figure 7 shows, for all scores, the likelihood that 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:

- 3.5 percent below the food line
- 2.8 percent between the food and USAID line
- 4.5 percent between the USAID and $1.08/day 1993 PPP line
- 13.3 percent between the $1.08/day 1993 PPP line and the national line
- 22.9 percent between the national line and $1.25/day 2005 PPP line
- 14.4 percent between the $1.25/day 2005 PPP line and $2.16/day 1993 PPP line
- 28.7 percent between the $2.16/day 1993 PPP line and $3.24/day 1993 PPP line
- 3.7 percent between the $3.24/day 1993 PPP line and $2.50/day 2005 PPP line
- 2.7 percent between the $2.50/day 2005 PPP line and $3.75/day 2005 PPP line
- 3.4 percent greater than the $3.75/day 2005 PPP line

Even though the scorecard is constructed partly based on judgment, the calibration process produces poverty likelihoods that are objective, that is, derived from data on consumption-based poverty lines. The poverty likelihoods would be objective even if indicators and/or points were selected without any data at all. In fact, objective

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

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

6.2 Accuracy of estimates of poverty likelihoods

As long as the relationship between indicators and poverty does not change, 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.[11]

Of course, the relationship between indicators and poverty changes with time, so the scorecard applied after the end of the 2003/4 NLSS fieldwork in April 2004 (as it must be in practice) will generally be biased.

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

- 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 also 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 two-sided interval containing the central 900, 950, or 990 differences between estimated and true poverty likelihoods

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

For the national line, the average poverty likelihood across bootstrap samples for scores of 20–24 in the validation sample is too low by 0.8 percentage points (Figure 8). For scores of 25–29, the estimate is too high by 9.1 percentage points.[12]

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

For the validation sample, the 90-percent confidence interval for the differences for scores of 20–24 is ±2.2 percentage points (Figure 8).[13] This means that in 900 of 1,000 bootstraps, the difference between the estimate and the true value is between –3.0 and 1.4 percentage points (because –0.8 – 2.2 = –3.0, and –0.8 + 2.2 = 1.4). In 950 of 1,000 bootstraps (95 percent), the difference is –0.8 ±2.6 percentage points, and in 990 of 1,000 bootstraps (99 percent), the difference is –0.8 ±3.4 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 Nepal'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 fact mitigates 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.

---

[12] 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 the population and split into sub-samples before repeating the entire process of scorecard building and validation.

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

Figure 9 (summarizing Figure 10 across poverty lines) shows that the absolute differences, when averaged across score ranges for a given poverty line, are 1.9 percentage points or less for the validation sample. The differences are due to sampling variation.

By construction, the scorecard here is unbiased. It may still, however, be *overfit* when applied after April 2004, the date that fieldwork for the 2003/4 NLSS ended. That is, it may fit the 2003/4 NLSS 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 2003/4 NLSS. Or the scorecard may be overfit in the sense that it becomes biased as the relationship between indicators and poverty changes.

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

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

and quality (which is beyond the scope of the scorecard) or by reducing overfitting (which likely has limited returns, given the scorecard's parsimony).

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

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

To illustrate, suppose a program samples three households on Jan. 1, 2008 and that they have scores of 20, 30, and 40, corresponding to poverty likelihoods of 42.9, 25.8, and 12.9 percent (national line, Figure 5). The group's estimated poverty rate is the households' average poverty likelihood of $(42.9 + 25.8 + 12.9) \div 3 = 27.2$ percent.[14]

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 precision (confidence intervals for the differences) for the scorecard applied to 1,000 bootstrap samples from the validation sample. For the national line, the scorecard is generally too high by about 0.8 percentage points; it estimates a poverty rate of 25.9 percent for the validation sample, but the true value is 25.1 percent (Figure 2). For all poverty lines, absolute differences for the validation sample are 1.6 percentage points or less, with an average of about 0.9 percentage points (Figure 11).[15]

[14] The group's poverty rate is *not* the poverty likelihood associated with the average score. Here, the average score is $(20 + 30 + 40) \div 3 = 30$, and the poverty likelihood associated with the average score is 25.8 percent. This is not the 27.2 percent found as the average of the three poverty likelihoods associated with each of the three scores.
[15] Figure 11 summarizes Figure 12 across all poverty lines.

As before, these differences are due to sampling variation in the validation sample and in the random division of the 2003/4 NLSS into three sub-samples.

In terms of precision, the 90-percent confidence interval for a group's estimated poverty rate at a point in time and $n = 16{,}384$ is 0.6 percentage points or less (Figure 11). This means that in 900 of 1,000 bootstraps of this size, the difference between the estimate and the true value is within 0.6 percentage points of the average difference. In the specific case of the national line and the validation sample, 90 percent of all samples of $n = 16{,}384$ produce estimates that differ from the true value in the range of 0.8 – 0.5 = 0.3 to 0.8 + 0.5 = 1.3 percentage points. (0.8 is the average difference, and ±0.5 is its 90-percent confidence interval.)

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

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

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

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

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

where

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

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

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

The scorecard, however, does not measure poverty directly, so this formula is not applicable. To derive a similar sample-size formula for the Nepal scorecard, consider the scorecard applied to the validation sample. Figure 2 shows that the expected (before measurement) poverty rate $\hat{p}$ for the national line is 26.4 (that is, the average poverty rate in the construction and calibration sub-samples). In turn, a sample size n of 16,384 and a 90-percent confidence level correspond to a confidence interval of ±0.50 percentage points (Figure 12).[17] Plugging these into the direct-measurement sample-size formula (1) above gives not n = 16,384 but rather $n = \left(\frac{1.64}{0.0050}\right)^2 \cdot 0.264 \cdot (1 - 0.264) =$

[17] Due to rounding, Figure 12 displays 0.5, not 0.50.

20,904. The ratio of the sample size for scoring (derived empirically via the bootstrap) to the sample size for direct measurement (derived from theory) is 16,384 ÷ 20,904 = 0.78.

Applying the same method to $n = 8{,}192$ (confidence interval of ±0.735 percentage points) gives $n = \left(\frac{1.64}{0.00735}\right)^2 \cdot 0.264 \cdot (1 - 0.264) = 9{,}674$. This time, the ratio of the sample size using scoring to the sample size using direct measurement is 8,192 ÷ 9,674 = 0.85. This ratio of 0.85 for $n = 8{,}192$ is relatively close to the ratio of 0.78 for $n = 16{,}384$. Indeed, applying this same procedure for all $n \geq 256$ in Figure 12 gives ratios that average to 0.85. This can be used to define a sample-size formula for the scorecard applied to the population in the validation sample:

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

where $\alpha = 0.85$ and z, c, and $\hat{p}$ are defined as in (1) above. It is this α that appears in Figure 11 as "α for sample size".

To illustrate the use of (2), suppose $c = 0.04125$ (confidence interval of ± 4.125 percentage points) and $z = 1.64$ (90-percent confidence). Then (2) gives $n = 0.85 \cdot \left(\frac{1.64}{0.04125}\right)^2 \cdot 0.264 \cdot (1 - 0.264) = 262$, which is reasonably close to the sample size of 256 for these parameters in Figure 12.

When the sample-size factor α is less than 1.0, it means that the scorecard is more precise than direct measurement, as it is for all of nine poverty lines in Figure 11.

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

In practice after April 2004 (the last month that the 2003/4 NLSS was in the field), an organization would select a poverty line (say, the national line), select a desired confidence level (say, 90 percent, or $z = 1.64$), select a desired confidence interval (say, ± 2 percentage points, or $c = 0.02$), make an assumption about $\hat{p}$ (perhaps based on a previous measurement such as the 25.9 percent national average for the 2003/4 NLSS in Figure 2), look up α (here, 0.85 for the national line), assume that the scorecard will still work in the future,[18] and then compute the required sample size. In this illustration, $n = 0.85 \cdot \left(\frac{1.64}{0.02}\right)^2 0.259 \cdot (1 - 0.259) = 1{,}097$.

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

[18] This paper reports accuracy for the scorecard applied to the validation sample, but it cannot test accuracy for later years. Still, performance after the 2003/4 NLSS will probably resemble that in the 2003/4 NLSS, with some deterioration as time passes.

8. Estimates of changes in group poverty rates over time

The change in a group's poverty rate between two points in time is estimated as the change in the average poverty likelihood of the households in the group. With data for 2003/4 only, this paper cannot estimate changes over time, nor can it present sample-size formula. 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 assuming 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, 2009, a program samples three households who score 20, 30, and 40 and so have poverty likelihoods of 42.9, 25.8, and 12.9 percent (national line, Figure 5). The group's baseline estimated poverty rate is the households' average poverty likelihood of (42.9 + 25.8 + 12.9) ÷ 3 = 27.2 percent.

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

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

By way of illustration, suppose that a year later on Jan. 1, 2010, the program samples three additional households who are in the same cohort as the three households originally sampled (or suppose that the program scores the same three original households a second time) and finds that their scores are 30, 40, and 50 (poverty likelihoods of 25.8, 12.9 and 2.7 percent, national line, Figure 5). Their average poverty likelihood at follow-up is now (25.8+ 12.9 + 2.7) ÷ 3 = 13.8 percent, an improvement of 27.2 – 13.8 = 13.4 percentage points.

This suggests that about one of every eight participants crossed the poverty line in 2008.[19] Among those who started below the line, about one in two (13.4 ÷ 27.2 = 49.2 percent) ended up above the line.[20]

[19] This is a net figure; some people start above the line and end below it, and vice versa.
[20] The scorecard does not reveal the reasons for this change.

8.3 Accuracy for estimated change in two independent samples

With data only for 2003/4, it is not possible to measure the accuracy of scorecard estimates of changes in groups' poverty rates over time. In practice, of course, Nepal'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.[21]

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

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

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

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

For 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 Nepal.

To illustrate the use of (4) to determine sample size for estimating changes in poverty rates across two independent samples, suppose the desired confidence level is 90 percent ($z = 1.64$), the desired confidence interval is 2 percentage points ($c = 0.02$), the poverty line is the national line, $\alpha = 1.50$, and $\hat{p} = 0.259$ (from Figure 2). Then the baseline sample size is $n = 1.50 \cdot 2 \cdot \left(\frac{1.64}{0.02}\right)^2 \cdot 0.259 \cdot (1 - 0.259) = 3{,}872$, and the follow-up sample size is also 3,872.

8.4 Accuracy for estimated change for one sample, scored twice

In general, the direct-measurement sample-size formula for this case is:[22]

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

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.

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

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

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

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

Suppose that 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 baseline measurement, but it is reasonable to use as its expected value a previously observed poverty rate. Given this and a poverty line, a sample-size formula for a single sample directly measured twice for Nepal (once after April 2004 and then again later) is:

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

As usual, (7) is multiplied by α to get scoring's sample-size formula:

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

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

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 sample will first be scored in 2005. The before-baseline poverty rate is 25.9 percent (p_{2004} = 0.259, Figure 2), and suppose α = 1.8. Then the baseline sample size is $n = 1.8 \cdot 2 \cdot \left(\frac{1.64}{0.02}\right)^2 \cdot \{0.0085 + 0.206 \cdot [0.259 \cdot (1 - 0.259)]\} = 1{,}163$.

Of course, the same group of 1,163 households is scored at follow-up as well.

For a given confidence level and confidence interval, sample sizes are smaller when one sample is scored twice than when there are two different samples.

9. Targeting

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

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

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

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

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

Figure 14 shows the distribution of households by targeting outcome for the scorecard applied to the validation sample. For an example cut-off of 15–19, outcomes for the national line applied to the validation sample are:

- Inclusion: 9.5 percent are below the line and correctly targeted
- Undercoverage: 15.6 percent are below the line and mistakenly not targeted
- Leakage: 4.6 percent are above the line and mistakenly targeted
- Exclusion: 70.4 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: 13.4 percent are below the line and correctly targeted
- Undercoverage: 11.7 percent are below the line and mistakenly not targeted
- Leakage: 9.4 percent are above the line and mistakenly targeted
- Exclusion: 65.5 percent are above the line and correctly not targeted

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

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

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

- Assign benefits and costs to possible outcomes, based on its values and mission
- Tally total net benefits for each cut-off using Figure 14 for a given poverty line
- Select the cut-off with the highest total net benefit

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

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

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

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

Figure 14 shows "Total Accuracy" for all cut-offs for the Nepal scorecard. For the national line in the validation sample, total net benefit is greatest (79.8) for a cut-off of 15-19, with nearly eight in ten Nepalese households correctly classified.

"Total Accuracy" weighs successful inclusion of households below the line the same as successful exclusion of households above the line. If a program valued inclusion more (say, twice as much) than exclusion, it could reflect this by setting the benefit for inclusion to 2 and the benefit for exclusion to 1. Then the chosen cut-off would maximize (2 x Households correctly included) + (1 x Households correctly excluded).[23]

[23] Beyond "Total Accuracy", IRIS (2005) proposes a new yardstick called the "Balanced Poverty Accuracy Criterion" that is meant to account for inclusion. USAID uses BPAC as its criteria for certifying poverty-assessment tools. After normalizing by the number of people below the poverty line, the BPAC formula is:
BPAC = (Inclusion – |Undercoverage – Leakage|) x [100 ÷ (Inclusion+Undercoverage)].

As an alternative to assigning benefits and costs to targeting outcomes and then choosing a cut-off to maximize total net benefit, a program could set a cut-off to achieve a desired poverty rate among targeted households. Figure 15 shows, for the Nepal scorecard applied to the validation sample, the expected poverty rate among households who score at or below a given cut-off. For the example of the national line, targeting households who score 15–19 or less would target 14.0 percent of all Nepalese households and produce a poverty rate among those targeted of 67.3 percent.[24]

[24] If potential participants are not representative of all of Nepal, then Figure 15 is valid only if selection into potential participation—whether by the program or potential participant—is unrelated with poverty in any way not captured by the scorecard.

10. Conclusion

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

The scorecard is built with a sub-sample of data from the 2003/4 NLSS, tested with a different sub-sample, and calibrated to nine poverty lines (national, food, USAID "extreme", \$1.08/day 1993 PPP, \$2.16/day 1993 PPP, \$3.24/day 1993 PPP, \$1.25/day 2005 PPP, \$2.50/day 2005 PPP, and \$3.75/day 2005 PPP).

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

When the scorecard is applied to the validation sample, the absolute difference between estimates versus true poverty rates for groups of households at a point in time is always 1.6 percentage points or less and averages—across the nine poverty lines—about 0.9 percentage points. For $n = 16{,}384$ and 90-percent confidence, the precision of

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

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

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

In sum, the scorecard is a practical, objective way for pro-poor programs in Nepal 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, Niall M.; and David J. Hand. (2000) "Improving the Practice of Classifier Performance Assessment", *Neural Computation*, Vol. 12, pp. 305–311.

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

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

Caire, Dean. (2004) "Building Credit Scorecards for Small Business Lending in Developing Markets", `microfinance.com/English/Papers/Scoring_SMEs_Hybrid.pdf`, accessed 24 February 2009.

Central Bureau of Statistics. (2004) *Nepal Living Standards Survey, 2003/4, Statistical Report: Volume 1*, `www.cbs.gov.np/Surveys/NLSSII/NLSS%20II%20Report%20Vol%201.pdf`, accessed 24 February 2009.

Central Bureau of Statistics. (2005) "Poverty Trends in Nepal (1995/6 and 2003/4)", `www.cbs.gov.np/Others/Poverty%20Assessment.pdf`, accessed 24 February 2009.

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

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

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

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

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

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 24 February 2009.

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

Grootaert, Christiaan; and Jeanine Braithwaite. (1998) "Poverty Correlates and Indicator-Based Targeting in Eastern Europe and the Former Soviet Union", World Bank Policy Research Working Paper No. 1942, 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, go.worldbank.org/W9OWN57PD0, retrieved 13 May 2016.

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

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

IRIS Center. (2007a) "Manual for the Implementation of USAID Poverty Assessment Tools", povertytools.org/training_documents/Manuals/USAID_PAT_Manual_Eng.pdf, accessed 24 February 2009.

_____. (2007b) "Introduction to Sampling for the Implementation of PATs", povertytools.org/training_documents/Sampling/Introduction_Sampling.ppt, accessed 24 February 2009.

_____. (2005) "Notes on Assessment and Improvement of Tool Accuracy", povertytools.org/other_documents/AssessingImproving_Accuracy.pdf, accessed 24 February 2009.

Johnson, Glenn. (2007) "Lesson 3: Two-Way Tables—Dependent Samples", www.stat.psu.edu/online/development/stat504/03_2way/53_2way_compare.htm, accessed 24 February 2009.

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

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

Martinelli, César; and Susan W. Parker. (2006) "Deception and Misreporting in a Social Program", `ciep.itam.mx/~martinel/lies4.pdf`, accessed 24 February 2009.

Matul, Michal; and Sean Kline. (2003) "Scoring Change: Prizma's Approach to Assessing Poverty", Microfinance Centre for Central and Eastern Europe and the New Independent States Spotlight Note No. 4, `www.mfc.org.pl/doc/Research/ImpAct/SN/MFC_SN04_eng.pdf`, accessed 24 February 2009.

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

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

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

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

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

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

Ravallion, Martin; Chen; Shaohua; and Prem Sangraula. (2008) "Dollar a Day Revisited", World Bank Policy Research Working Paper No. 4620, `www-wds.worldbank.org/external/default/WDSContentServer/IW3P/IB/2008/08/26/000158349_20080826113239/Rendered/PDF/WPS4703.pdf`, accessed 24 February 2009.

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

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

SAS Institute Inc. (2004) "The LOGISTIC Procedure: Rank Correlation of Observed Responses and Predicted Probabilities", in *SAS/STAT User's Guide, Version 9*, `support.sas.com/documentation/cdl/en/statug/59654/HTML/default/statug_logistic_sect035.htm`, accessed 24 February 2009.

Schreiner, Mark. (2013) "Simple Poverty Scorecard Poverty-Assessmeent Tool: Bangladesh", `simplepovertyscorecard.com/BGD_2010_ENG.pdf`, accessed 8 July 2016.

_____. (2008a) "Simple Poverty Scorecard Poverty-Assessmeent Tool: Peru", `simplepovertyscorecard.com/PER_2003_ENG.pdf`, accessed 8 July 2016.

_____. (2008b) "Simple Poverty Scorecard Poverty-Assessmeent Tool: India", `simplepovertyscorecard.com/IND_2005_ENG.pdf`, accessed 8 July 2016.

_____. (2006) "Is One Simple Poverty Scorecard Poverty-Assessmeent Tool Enough for India?", `microfinance.com/English/Papers/Scoring_Poverty_India_Segments.pdf`, accessed 24 February 2009.

_____. (2005a) "La Herramienta del Índice de Calificación de la Pobreza™: México", `simplepovertyscorecard.com/MEX_2002_SPA.pdf`, accessed 8 July 2016.

_____. (2005b) "IRIS Questions on the Simple Poverty Scorecard Poverty-Assessmeent Tool", `microfinance.com/English/Papers/Scoring_Poverty_Response_to_IRIS.pdf`, accessed 24 February 2009.

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

_____; Matul, Michal; Pawlak, Ewa; and Sean Kline. (2004) "Poverty Scoring: Lessons from a Microlender in Bosnia-Herzegovina", `microfinance.com/English/Papers/Scoring_Poverty_in_BiH_Short.pdf`, accessed 24 February 2009.

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

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

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

Toohig, Jeff. (2007) "PPI: Training Guide", progressoutofpoverty.org/toolkit, accessed 24 February 2009.

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

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

World Bank. (2006) *Resilience Amidst Conflict: An Assessment of Poverty in Nepal, 1995–96 and 2003–04*, Report No. 34834–NP, www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2006/07/13/000090341_20060713084841/Rendered/PDF/348340NP.pdf, accessed 24 February 2009.

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

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

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

		% with expenditure below a poverty line								
			National	USAID	1993 PPP			2005 PPP		
Sub-sample	**Households**	**National**	**Food**	**'Extreme'**	**$1.08/day**	**$2.16/day**	**$3.24/day**	**$1.25/day**	**$2.50/day**	**$3.75/day**
All Nepal	3,912	25.9	6.8	12.3	17.1	58.4	78.4	47.8	81.8	90.6
Construction										
Selecting indicators and weights	1,332	26.4	7.2	12.5	17.3	58.9	79.0	48.8	82.4	91.1
Calibration										
Associating scores with likelihoods	1,270	26.4	7.1	12.9	18.0	57.5	77.2	47.0	81.2	90.3
Validation										
Measuring accuracy	1,310	25.1	6.0	11.5	16.0	58.8	79.0	47.6	81.9	90.5
Change between construction and calibration to validation (percentage points)										
		1.3	1.2	1.2	1.7	-0.6	-1.0	0.3	-0.1	0.1

Source: NLSS 2003/4

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

Region	Line or rate	Poverty line (NRP/person/day) and poverty rate (%)								
			National	USAID	1993 PPP			2005 PPP		
		National	Food	'Extreme'	$1.08/day	$2.16/day	$3.24/day	$1.25/day	$2.50/day	$3.75/day
Katmandu	Line	30.29	18.42	25.16	26.10	52.20	78.30	43.53	87.07	130.60
	Rate	2.8	0.4	1.1	1.8	14.8	28.5	7.8	33.4	53.5
Other urban	Line	21.65	13.48	16.70	18.65	37.30	55.95	31.11	62.22	93.33
	Rate	10.0	1.6	5.5	6.7	29.9	50.0	23.0	54.5	72.2
Rural Western Hill	Line	24.39	15.38	18.89	21.01	42.02	63.03	35.05	70.10	105.14
	Rate	31.6	6.7	15.1	20.2	69.4	87.1	57.4	89.5	97.0
Rural Eastern Hill	Line	22.11	14.55	15.44	19.05	38.10	57.14	31.77	63.55	95.32
	Rate	36.6	14.7	17.2	27.6	69.1	86.9	58.9	89.6	95.3
Rural Western Terai	Line	20.32	11.80	16.38	17.51	35.02	52.53	29.21	58.42	87.63
	Rate	31.3	3.4	15.3	19.5	63.4	84.1	54.1	86.4	92.5
Rural Eastern Terai	Line	16.65	11.84	13.74	14.35	28.70	43.05	23.93	47.87	71.80
	Rate	21.6	5.6	10.3	12.5	59.3	83.8	46.5	88.6	96.2
All-Nepal	Line	21.31	13.73	16.55	18.36	36.71	55.07	30.62	61.24	91.86
	Rate	25.9	6.8	12.3	17.1	58.4	78.4	47.8	81.8	90.6

Source: NLSS 2003/4

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

Region	Line or rate	Poverty line (NRP/person/day) and poverty rate (%)								
			National	USAID	1993 PPP			2005 PPP		
		National	Food	'Extreme'	$1.08/day	$2.16/day	$3.24/day	$1.25/day	$2.50/day	$3.75/day
Katmandu	Line	30.29	18.42	25.16	26.10	52.20	78.30	43.53	87.07	130.60
	Rate	3.3	0.4	1.6	2.2	15.2	29.5	7.8	34.5	55.3
Other urban	Line	21.65	13.48	16.70	18.65	37.30	55.95	31.11	62.22	93.33
	Rate	13.0	2.2	6.3	8.9	34.8	55.8	26.9	60.1	75.8
Rural Western Hill	Line	24.39	15.38	18.89	21.01	42.02	63.03	35.05	70.10	105.14
	Rate	37.4	8.5	18.7	24.2	74.9	90.8	63.8	92.9	98.0
Rural Eastern Hill	Line	22.11	14.55	15.44	19.05	38.10	57.14	31.77	63.55	95.32
	Rate	42.9	18.9	21.3	32.7	74.1	89.7	64.6	91.7	96.1
Rural Western Terai	Line	20.32	11.80	16.38	17.51	35.02	52.53	29.21	58.42	87.63
	Rate	38.1	3.9	18.9	23.9	69.6	88.0	61.9	89.7	94.1
Rural Eastern Terai	Line	16.65	11.84	13.74	14.35	28.70	43.05	23.93	47.87	71.80
	Rate	24.9	6.5	12.2	14.8	63.7	86.2	51.7	90.4	96.7
All-Nepal	Line	21.08	13.61	16.40	18.16	36.33	54.49	30.30	60.60	90.90
	Rate	30.8	8.3	15.3	20.5	63.6	82.1	53.6	85.1	92.2

Source: NLSS 2003/4

Figure 4: Poverty indicators by uncertainty coefficient

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly indicative of poverty)
128	How many household members are 12-years-old or younger? (Four or more; Three; Two; One; None)
127	How many children ages 5 to 12 attend school? (Not all; No children ages 5 to 12; All)
123	How many children ages 5 to 11 attend school? (Not all; No children ages 5 to 11; All)
122	How many household members are 11-years-old or younger? (Four or more; Three; Two; One; None)
111	What is the main material of the roof? (Straw/thatch, wood/planks, earth/mud, or other; Tiles/slate; Galvanized iron or concrete/cement)
106	What type of toilet is used in your household? (No toilet; Household non-flush, communal latrine, household flush (connected to municipal sewer), or household flush (connected to septic tank))
88	What is the main source of lighting for your dwelling? (Electricity; Gas/oil/kerosene, other, or no data)
83	Does the household own any televisions/decks? (No; Yes)
78	How many members does the household have? (Eight or more; Seven; Six; Five; Four; One, two, or three)
76	How many radio/tape/CD players does the household own? (None; One; Two or more)
75	What kind of fuel is most often used in your household for cooking? (Wood/firewood, dung, or leaves/rubbish/straw/thatch; Cylinder gas, kerosene, bio-gas, other, or no data)
72	How many bedrooms are in your house? (None; One; Two; Three; Four or more)
69	What is the highest grade completed by any household member? (Third or lower; Fourth to ninth; Tenth or higher)
69	How are the windows? (No windows/no covering, or other; Shutters; Screen/glass)
69	What is the main construction material of the outside walls? (Wood/branches; Mud-bonded bricks/stones; Concrete, unbaked bricks, other material, or no outside walls; Cement-bonded bricks/stones)

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

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly associated with poverty)
65	What was the highest grade of education completed by the male head/spouse? (None; First to ninth; Tenth or more; No male head/spouse)
65	During the past 12 months, how many household members were employed in agriculture (self-employed or wage)? (Five or more; Four; Three; Two; One; None)
61	What is the main flooring material? (Earth; Wood, stone/brick, cement/tile, or other)
57	Does the male head/spouse know how to read a letter? (No; No male head/spouse; Yes)
56	How many household members receive wages on a daily basis? (Two or more; One; None)
54	Does the female head/spouse know how to read a letter? (No; No female head/spouse; Yes)
54	What was the highest grade of education completed by the female head/spouse? (None; No female head/spouse; First to fifth; Sixth or more)
54	During the past 12 months, how many household members had wage employment in agriculture? (Two or more; One; None)
53	If any household members are employed in agriculture, how many poultry/ducks/pigeons do you own? (One; Two or more; None; No household members are employed in agriculture)
52	Does the household own any fans? (No; Yes)
52	How many household members work in agriculture and receive wages on a daily basis? (Two or more; One; None)
51	During the past 12 months, how many household members were employed as agricultural, fishery or related laborers? (Two or more; One; None)
49	How many rooms does your household occupy? (One; Two; Three; Four; Five or more)
48	If any household members are employed in agriculture, how many he/she buffaloes do you own? (One; None; Two or more; No household members are employed in agriculture)
47	If any household members are employed in agriculture, how many goats/castrated goats do you own? (Two or more; One; None; No household members are employed in agriculture)

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

Uncertainty coefficient	Indicator (Answers ordered starting with those most strongly associated with poverty)
47	If any household members are employed in agriculture, does the household own any agricultural land? (No; Yes; No household members are employed in agriculture)
47	If any household members are employed in agriculture, how many grain-storage bins does the household own? (None; One or more; No household members are employed in agriculture)
46	What is the source of your drinking water? (Other or no data; Open well; Covered well/hand pump; Piped water, but not in the house; Water piped to the house)
45	If any household members are employed in agriculture, how many ploughs does the household own? (One; Two or more; None; No household members are employed in agriculture)
44	If any household members are employed in agriculture, how many bullocks/cows do you own? (Two or more; One; None; No household members are employed in agriculture)
44	During the past 12 months, how many household members were self-employed in agriculture? (Five or more; Four; Three; Two; One; None)
44	Do you have a land-line telephone and/or mobile phone in your dwelling unit? (No phones or no data; Mobile only; Land-line only; Both land-line and mobile)
43	If any household members are employed in agriculture, do they cultivate land owned by someone else or that is mortgaged in? (No; Yes; No household members are employed in agriculture)
41	Do any household members attend a private school? (No; Yes)
37	How big is the inside of the dwelling (in meters squared)? (Less than 150; 151 to 475; 476 to 750; 751 or more)
36	During the past 12 months, how many household members worked in the VDC/NP? (Five or more; Four; Three; Two; None or one)
32	What kind of sewerage facility does your household have? (None; Underground drains, open drains, or soak pit)
26	Does your household have a kitchen? (No; Yes)
26	What type of stove does your household mainly use for cooking? (Open fireplace, other, or no data; Mud stove, smokeless stove, or kerosene/gas stove)

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

20	Does the household own any bicycles, motorcycles/scooters, motor cars, etc.? (None; One; Two or more)
20	How many household members work in non-agriculture and receive wages on a daily basis? (One or more; None)
20	How many household members receive wages on a long-term basis? (None; One or more)
19	Does any household member know how to read a letter? (No; Yes)
10	How old is the female head/spouse? (25 to 35 years; 36 years or more; 24 years or less; No female head/spouse)
7	What is the structure of household headship? (Both male and female heads/spouses; Female head/spouse only; Male head/spouse only)
7	How old is the male head/spouse? (31 to 45 years; 61 years or more; 30 years or less; 46 to 60 years; No male head/spouse)
6	What is the present marital status of the male head/spouse? (Married; Divorced; Separated; Widow/widower; Never-married; No male head/spouse)
5	Does the male head/spouse work in the VDC/NP? (Yes; No male head/spouse; No)
4	Does the female head/spouse work in the VDC/NP? (Yes; No female head/spouse; No)
3	What is the current marital status of the female head/spouse? (Married; Widow/Widower; No female head/spouse; Separated; Never-married)
3	Is there a kitchen garden in your household? (No; Yes)
3	How many living/dining rooms are in your household? (None; One or more)
2	During the past 12 months, how many household members were self-employed in non-agriculture? (None; One; Two or more)
2	During the past 12 months, how many household members had non-agricultural wage employment? (Two or more; One; None)
2	How many household members suffer from a chronic illness or have had to stop doing their usual activities because of an illness or injury? (None; One; Two or more)
1	How many household members work on a piece-rate basis? (One or more; None)
0	Which year was the house that you are living in built? (Before 1996; 1996 or after)

Source: 2003/4 NLSS, national poverty line.

National Poverty Line Tables

(and tables pertaining to all nine poverty lines)

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

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	90.2
5–9	87.5
10–14	62.0
15–19	57.8
20–24	42.9
25–29	43.5
30–34	25.8
35–39	24.2
40–44	12.9
45–49	13.5
50–54	2.7
55–59	2.1
60–64	0.0
65–69	0.4
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 Nepal's households.
Based on the NLSS 2003/4.

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

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	906	÷	1,005	=	90.2
5–9	2,328	÷	2,661	=	87.5
10–14	2,722	÷	4,394	=	62.0
15–19	3,461	÷	5,989	=	57.8
20–24	3,744	÷	8,719	=	42.9
25–29	4,767	÷	10,969	=	43.5
30–34	2,819	÷	10,911	=	25.8
35–39	2,669	÷	11,042	=	24.2
40–44	1,056	÷	8,176	=	12.9
45–49	931	÷	6,899	=	13.5
50–54	175	÷	6,427	=	2.7
55–59	104	÷	4,964	=	2.1
60–64	0	÷	3,870	=	0.0
65–69	16	÷	4,019	=	0.4
70–74	0	÷	3,769	=	0.0
75–79	0	÷	2,755	=	0.0
80–84	0	÷	2,133	=	0.0
85–89	0	÷	442	=	0.0
90–94	0	÷	540	=	0.0
95–100	0	÷	317	=	0.0

Surveyed cases weighted to represent Nepal's households.
Based on the NLSS 2003/4.

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

	Likelihood (%) of having expenditure in range demarcated by poverty lines in NRP per day per person									
Score	< Food	≥Food and <USAID	≥USAID and <$1.08/day 93	≥$1.08/day 93 and <National	≥National and <$1.25/day 05	≥$1.25/day 05 and <$2.16/day 93	≥$2.16/day 93 and <$3.24/day 93	≥$3.24/day 93 and <$2.50/day 05	≥$2.50/day 05 and <$3.75/day 05	>$3.75/day 05
	<13.61	≥13.61 and <16.40	≥16.40 and <18.16	≥18.16 and <21.08	≥21.08 and <30.30	≥30.30 and <36.33	≥36.33 and <54.49	≥54.49 and <60.60	≥60.60 and <90.90	≥90.90
0–4	62.7	27.5	0.0	0.0	9.8	0.0	0.0	0.0	0.0	0.0
5–9	59.2	16.2	7.5	4.6	11.5	1.0	0.0	0.0	0.0	0.0
10–14	27.2	12.3	9.8	12.7	26.0	7.6	4.5	0.0	0.0	0.0
15–19	16.6	10.0	12.2	19.0	33.6	5.7	3.0	0.0	0.0	0.0
20–24	9.8	10.2	8.6	14.4	33.8	7.2	14.4	0.6	1.1	0.0
25–29	5.5	15.8	9.8	12.3	31.9	13.7	8.2	0.6	1.3	0.9
30–34	1.6	7.5	6.4	10.3	34.9	13.1	20.2	0.0	4.1	1.9
35–39	3.5	2.8	4.5	13.3	22.9	14.4	28.7	3.7	2.7	3.4
40–44	3.4	1.4	0.9	7.3	21.5	20.9	28.5	9.2	5.5	1.4
45–49	1.3	0.0	6.0	6.2	16.5	14.3	33.5	5.1	9.8	7.3
50–54	0.0	0.0	0.0	2.7	14.6	19.6	35.4	7.0	18.1	2.5
55–59	0.0	0.0	0.0	2.1	6.7	3.9	43.4	10.8	28.3	4.8
60–64	0.0	0.0	0.0	0.0	0.0	14.6	39.9	5.6	22.4	17.6
65–69	0.0	0.0	0.4	0.0	0.0	4.0	30.6	12.1	28.5	24.5
70–74	0.0	0.0	0.0	0.0	1.5	1.2	19.6	4.3	29.9	43.6
75–79	0.0	0.0	0.0	0.0	3.5	3.0	5.1	10.1	29.0	49.3
80–84	0.0	0.0	0.0	0.0	0.0	1.3	8.4	9.8	19.1	61.4
85–89	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8.6	2.5	88.9
90–94	0.0	0.0	0.0	0.0	0.0	0.0	0.0	11.0	40.7	48.4
95–100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0

Note: All poverty likelihoods in percentage units.

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

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	–2.6	2.8	3.4	4.1
5–9	+12.4	3.5	4.1	5.4
10–14	–12.5	7.6	7.8	8.2
15–19	+4.8	2.7	3.2	4.2
20–24	–0.8	2.2	2.6	3.4
25–29	+9.1	2.0	2.3	3.0
30–34	–6	3.8	4.0	4.4
35–39	+6.7	1.6	1.9	2.4
40–44	–3.4	2.6	2.8	3.2
45–49	+5.4	1.3	1.6	2.1
50–54	–4.9	3.2	3.3	3.6
55–59	–0.1	0.7	0.8	1.0
60–64	+0	0.0	0.0	0.0
65–69	–1.6	1.3	1.4	1.6
70–74	+0	0.0	0.0	0.0
75–79	–0.4	0.4	0.4	0.5
80–84	+0	0.0	0.0	0.0
85–89	+0	0.0	0.0	0.0
90–94	+0	0.0	0.0	0.0
95–100	+0	0.0	0.0	0.0

Based on scorecard applied to the validation sample.

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

	Poverty line								
		National	USAID	1993 PPP			2005 PPP		
	National	Food	'Extreme'	$1.08/day	$2.16/day	$3.24/day	$1.25/day	$2.50/day	$3.75/day
Estimate minus true value	0.4	1.0	1.7	1.9	-0.9	-1.3	0.0	0.7	1.5
Precision of difference	0.4	0.4	0.5	0.5	0.4	0.5	0.4	0.5	0.5

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.

Scorecard is based on construction and validation samples and is applied to the validation sample in the NLSS 2003/4.

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

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Sample size (n)	Diff.	90-percent	95-percent	99-percent
2	0.6	44.3	52.5	67.5
4	0.5	33.7	38.2	52.0
8	0.1	24.7	29.3	38.0
16	0.5	17.3	20.8	28.1
32	0.5	12.0	14.4	19.4
64	0.5	8.0	9.7	13.0
128	0.3	5.4	6.4	8.8
256	0.4	3.6	4.2	5.3
512	0.3	2.5	2.9	3.8
1,024	0.4	1.7	2.1	2.6
2,048	0.4	1.2	1.5	1.9
4,096	0.4	0.8	1.0	1.3
8,192	0.4	0.6	0.7	0.9
16,384	0.4	0.4	0.5	0.6

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

	Poverty line								
		National	USAID	1993 PPP			2005 PPP		
	National	Food	'Extreme'	$1.08/day	$2.16/day	$3.24/day	$1.25/day	$2.50/day	$3.75/day
Estimate minus true value	0.8	0.6	0.8	1.4	-1.6	-1.5	-0.9	-0.3	0.4
Precision of difference	0.5	0.3	0.4	0.5	0.5	0.4	0.5	0.4	0.6
α for sample size	0.85	0.96	0.89	0.88	0.70	0.54	0.71	0.52	0.50

Precision is measured as 90-percent confidence intervals in units of +/− percentage points.

Differences and precision estimated from 1,000 bootstraps of size n = 16,384.

α is estimated from 1,000 bootstrap samples of n = 256, 512, 1,024, 2,048, 4,096, 8,192, and 16,384.

Figure 12 (National poverty line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Sample size (n)	Diff.	90-percent	95-percent	99-percent
2	0.6	44.3	52.5	67.5
4	0.5	33.3	37.9	50.7
8	0.2	24.7	28.1	37.9
16	0.5	16.8	19.8	27.8
32	0.8	11.6	14.1	18.3
64	0.7	8.5	9.9	13.4
128	0.8	5.8	7.0	8.9
256	0.8	4.1	4.9	6.2
512	0.8	3.0	3.4	4.7
1,024	0.8	2.1	2.5	3.2
2,048	0.8	1.5	1.8	2.3
4,096	0.8	1.0	1.3	1.7
8,192	0.8	0.7	0.9	1.1
16,384	0.8	0.5	0.6	0.8

Figure 13 (All poverty lines): Possible types of outcomes from targeting by 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): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.9	24.2	0.1	74.8	75.7	-92.4
5–9	2.9	22.1	0.7	74.2	77.1	-73.7
10–14	6.2	18.8	1.8	73.1	79.4	-42.9
15–19	9.5	15.6	4.6	70.4	79.8	-6.2
20–24	13.4	11.7	9.4	65.5	78.9	44.2
25–29	17.1	8.0	16.7	58.3	75.3	33.4
30–34	20.5	4.6	24.2	50.8	71.3	3.5
35–39	22.4	2.6	33.3	41.7	64.1	-32.8
40–44	23.7	1.4	40.2	34.8	58.5	-60.4
45–49	24.3	0.7	46.5	28.5	52.8	-85.4
50–54	24.8	0.3	52.4	22.5	47.3	-109.2
55–59	24.9	0.1	57.2	17.7	42.7	-128.4
60–64	24.9	0.1	61.1	13.9	38.8	-143.8
65–69	25.0	0.0	65.0	9.9	35.0	-159.6
70–74	25.0	0.0	68.8	6.2	31.2	-174.6
75–79	25.1	0.0	71.5	3.4	28.5	-185.5
80–84	25.1	0.0	73.7	1.3	26.3	-194.0
85–89	25.1	0.0	74.1	0.9	25.9	-195.8
90–94	25.1	0.0	74.6	0.3	25.4	-197.9
95–100	25.1	0.0	74.9	0.0	25.1	-199.2

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

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

Score	Households at or below score who are also below the poverty line (%)	Households at or below score (%)
0–4	89.1	1.0
5–9	79.7	3.7
10–14	77.5	8.1
15–19	67.3	14.0
20–24	58.6	22.8
25–29	50.6	33.7
30–34	45.9	44.6
35–39	40.3	55.7
40–44	37.1	63.9
45–49	34.4	70.8
50–54	32.1	77.2
55–59	30.4	82.2
60–64	29.0	86.0
65–69	27.8	90.0
70–74	26.7	93.8
75–79	25.9	96.6
80–84	25.4	98.7
85–89	25.3	99.1
90–94	25.1	99.7
95–100	25.1	100.0

National Food Poverty Line Tables

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

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	62.7
5–9	59.2
10–14	27.2
15–19	16.6
20–24	9.8
25–29	5.5
30–34	1.6
35–39	3.5
40–44	3.4
45–49	1.3
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 Nepal's households.
Based on the NLSS 2003/4.

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

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	630	÷	1,005	=	62.7
5–9	1,574	÷	2,661	=	59.2
10–14	1,193	÷	4,394	=	27.2
15–19	994	÷	5,989	=	16.6
20–24	852	÷	8,719	=	9.8
25–29	604	÷	10,969	=	5.5
30–34	177	÷	10,911	=	1.6
35–39	390	÷	11,042	=	3.5
40–44	275	÷	8,176	=	3.4
45–49	90	÷	6,899	=	1.3
50–54	0	÷	6,427	=	0.0
55–59	0	÷	4,964	=	0.0
60–64	0	÷	3,870	=	0.0
65–69	0	÷	4,019	=	0.0
70–74	0	÷	3,769	=	0.0
75–79	0	÷	2,755	=	0.0
80–84	0	÷	2,133	=	0.0
85–89	0	÷	442	=	0.0
90–94	0	÷	540	=	0.0
95–100	0	÷	317	=	0.0

Surveyed cases weighted to represent Nepal's households.
Based on the NLSS 2003/4.

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

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	–4	6.1	7.1	10.9
5–9	+12.8	3.8	4.7	6.1
10–14	+9.3	2.3	2.7	3.8
15–19	+5.4	1.7	2.0	2.9
20–24	+0.8	1.3	1.5	2.0
25–29	–0.4	1.1	1.4	1.7
30–34	–2.8	1.8	1.9	2.1
35–39	+0.4	0.7	0.8	1.0
40–44	–3.2	2.4	2.5	2.8
45–49	+0.6	0.4	0.4	0.5
50–54	+0	0.0	0.0	0.0
55–59	–0.3	0.3	0.3	0.4
60–64	+0	0.0	0.0	0.0
65–69	+0	0.0	0.0	0.0
70–74	+0	0.0	0.0	0.0
75–79	+0	0.0	0.0	0.0
80–84	+0	0.0	0.0	0.0
85–89	+0	0.0	0.0	0.0
90–94	+0	0.0	0.0	0.0
95–100	+0	0.0	0.0	0.0

Based on scorecard applied to the validation sample.

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

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Sample size (n)	Diff.	90-percent	95-percent	99-percent
2	0.2	30.5	39.1	54.8
4	0.6	19.0	26.0	36.5
8	0.3	16.6	19.9	25.6
16	0.8	11.4	13.4	18.7
32	1.2	8.5	9.8	13.1
64	1.3	6.5	7.7	10.1
128	1.2	5.3	6.2	7.8
256	1.0	4.0	4.7	6.2
512	1.1	2.7	3.2	4.3
1,024	1.0	1.9	2.3	2.8
2,048	1.0	1.2	1.5	1.9
4,096	1.0	0.9	1.1	1.4
8,192	1.0	0.6	0.8	1.0
16,384	1.0	0.4	0.5	0.7

Figure 12 (Food line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Sample size (n)	Diff.	90-percent	95-percent	99-percent
2	0.2	30.5	39.1	54.8
4	0.7	19.3	25.4	35.4
8	0.3	14.8	18.1	24.3
16	0.6	10.2	12.2	16.4
32	0.8	6.9	8.4	10.5
64	0.8	5.1	5.9	7.3
128	0.7	3.6	4.3	5.4
256	0.6	2.6	3.1	4.0
512	0.6	1.9	2.2	2.9
1,024	0.6	1.3	1.6	1.9
2,048	0.6	0.9	1.1	1.4
4,096	0.6	0.7	0.8	1.1
8,192	0.6	0.4	0.6	0.7
16,384	0.6	0.3	0.4	0.5

Figure 14 (Food line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.6	5.3	0.4	93.7	94.3	-72.4
5–9	1.9	4.0	1.7	92.3	94.2	-6.3
10–14	2.7	3.2	5.3	88.7	91.4	11.3
15–19	3.4	2.6	10.7	83.3	86.7	-78.2
20–24	4.1	1.8	18.6	75.4	79.5	-210.8
25–29	4.7	1.3	29.1	64.9	69.6	-385.4
30–34	5.1	0.9	39.5	54.5	59.6	-559.8
35–39	5.5	0.5	50.2	43.8	49.3	-738.2
40–44	5.9	0.1	58.0	36.0	41.9	-867.8
45–49	6.0	0.0	64.8	29.2	35.2	-981.9
50–54	6.0	0.0	71.2	22.8	28.7	-1,089.2
55–59	6.0	0.0	76.2	17.8	23.8	-1,171.4
60–64	6.0	0.0	80.0	14.0	20.0	-1,236.0
65–69	6.0	0.0	84.1	10.0	15.9	-1,303.1
70–74	6.0	0.0	87.8	6.2	12.2	-1,366.0
75–79	6.0	0.0	90.6	3.4	9.4	-1,412.0
80–84	6.0	0.0	92.7	1.3	7.3	-1,447.6
85–89	6.0	0.0	93.2	0.9	6.8	-1,455.0
90–94	6.0	0.0	93.7	0.3	6.3	-1,464.0
95–100	6.0	0.0	94.0	0.0	6.0	-1,469.3

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

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

Score	Households at or below score who are also below the	Households at or below score (%)
0–4	64.5	1.0
5–9	53.1	3.7
10–14	34.1	8.1
15–19	24.0	14.0
20–24	18.2	22.8
25–29	13.8	33.7
30–34	11.5	44.6
35–39	9.8	55.7
40–44	9.2	63.9
45–49	8.4	70.8
50–54	7.7	77.2
55–59	7.3	82.2
60–64	7.0	86.0
65–69	6.7	90.0
70–74	6.4	93.8
75–79	6.2	96.6
80–84	6.1	98.7
85–89	6.0	99.1
90–94	6.0	99.7
95–100	6.0	100.0

USAID "Extreme" Poverty Line Tables

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

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	90.2
5–9	75.4
10–14	39.4
15–19	26.6
20–24	19.9
25–29	21.3
30–34	9.1
35–39	6.3
40–44	4.7
45–49	1.3
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 Nepal's households.
Based on the NLSS 2003/4.

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

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	906	÷	1,005	=	90.2
5–9	2,006	÷	2,661	=	75.4
10–14	1,732	÷	4,394	=	39.4
15–19	1,591	÷	5,989	=	26.6
20–24	1,737	÷	8,719	=	19.9
25–29	2,339	÷	10,969	=	21.3
30–34	994	÷	10,911	=	9.1
35–39	698	÷	11,042	=	6.3
40–44	385	÷	8,176	=	4.7
45–49	90	÷	6,899	=	1.3
50–54	0	÷	6,427	=	0.0
55–59	0	÷	4,964	=	0.0
60–64	0	÷	3,870	=	0.0
65–69	0	÷	4,019	=	0.0
70–74	0	÷	3,769	=	0.0
75–79	0	÷	2,755	=	0.0
80–84	0	÷	2,133	=	0.0
85–89	0	÷	442	=	0.0
90–94	0	÷	540	=	0.0
95–100	0	÷	317	=	0.0

Surveyed cases weighted to represent Nepal's households.
Based on the NLSS 2003/4.

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

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+13.4	5.5	6.7	8.8
5–9	+18.2	3.9	4.5	5.9
10–14	+3.2	3.1	3.7	4.9
15–19	+0	2.3	2.8	3.8
20–24	–1.8	1.8	2.2	3.0
25–29	+5	1.6	1.9	2.5
30–34	–1.1	1.3	1.4	2.0
35–39	+0.9	0.9	1.0	1.4
40–44	–1.9	1.7	1.8	2.2
45–49	–1.7	1.3	1.3	1.5
50–54	+0	0.0	0.0	0.0
55–59	–1.3	0.9	1.0	1.1
60–64	+0	0.0	0.0	0.0
65–69	+0	0.0	0.0	0.0
70–74	+0	0.0	0.0	0.0
75–79	+0	0.0	0.0	0.0
80–84	+0	0.0	0.0	0.0
85–89	+0	0.0	0.0	0.0
90–94	+0	0.0	0.0	0.0
95–100	+0	0.0	0.0	0.0

Based on scorecard applied to the validation sample.

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

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Sample size (n)	Diff.	90-percent	95-percent	99-percent
2	0.1	34.4	44.0	63.3
4	0.6	25.3	30.8	42.5
8	0.2	19.2	22.9	29.3
16	0.9	14.1	16.7	22.8
32	1.1	10.2	11.7	16.4
64	1.4	7.5	8.7	11.3
128	1.6	5.4	6.8	8.6
256	1.6	3.9	4.6	6.0
512	1.7	2.6	3.1	4.0
1,024	1.7	1.8	2.2	2.9
2,048	1.7	1.3	1.6	2.0
4,096	1.7	0.9	1.1	1.5
8,192	1.7	0.6	0.8	1.0
16,384	1.7	0.5	0.5	0.7

Figure 12 (USAID "extreme" line): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Sample size (n)	Diff.	90-percent	95-percent	99-percent
2	0.1	34.4	44.0	63.3
4	0.8	24.3	30.7	42.2
8	0.3	18.1	21.7	26.8
16	0.6	12.8	15.5	19.3
32	0.8	8.8	10.4	13.4
64	0.8	6.4	7.6	10.2
128	0.8	4.6	5.4	7.1
256	0.8	3.2	3.8	5.2
512	0.8	2.2	2.7	3.6
1,024	0.9	1.6	2.0	2.6
2,048	0.8	1.1	1.4	1.8
4,096	0.8	0.8	1.0	1.3
8,192	0.8	0.6	0.7	0.9
16,384	0.8	0.4	0.5	0.7

Figure 14 (USAID "extreme" line): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.7	10.8	0.3	88.2	89.0	-84.9
5–9	2.4	9.2	1.3	87.2	89.5	-47.8
10–14	3.9	7.6	4.1	84.4	88.3	4.1
15–19	5.5	6.0	8.5	79.9	85.4	25.8
20–24	7.4	4.2	15.4	73.1	80.4	-33.7
25–29	9.0	2.5	24.7	63.8	72.8	-114.3
30–34	10.1	1.4	34.5	54.0	64.1	-199.7
35–39	10.7	0.8	44.9	43.5	54.3	-290.1
40–44	11.2	0.4	52.7	35.8	46.9	-357.5
45–49	11.4	0.1	59.3	29.1	40.6	-415.1
50–54	11.4	0.1	65.8	22.7	34.1	-470.9
55–59	11.5	0.0	70.6	17.8	29.4	-513.1
60–64	11.5	0.0	74.5	14.0	25.5	-546.7
65–69	11.5	0.0	78.5	10.0	21.5	-581.6
70–74	11.5	0.0	82.3	6.2	17.7	-614.3
75–79	11.5	0.0	85.0	3.4	15.0	-638.2
80–84	11.5	0.0	87.2	1.3	12.8	-656.7
85–89	11.5	0.0	87.6	0.9	12.4	-660.6
90–94	11.5	0.0	88.2	0.3	11.8	-665.2
95–100	11.5	0.0	88.5	0.0	11.5	-668.0

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

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

Score	Households at or below score who are also below the	Households at or below score (%)
0–4	73.6	1.0
5–9	64.2	3.7
10–14	48.9	8.1
15–19	39.1	14.0
20–24	32.4	22.8
25–29	26.8	33.7
30–34	22.7	44.6
35–39	19.3	55.7
40–44	17.5	63.9
45–49	16.1	70.8
50–54	14.8	77.2
55–59	14.0	82.2
60–64	13.4	86.0
65–69	12.8	90.0
70–74	12.3	93.8
75–79	11.9	96.6
80–84	11.7	98.7
85–89	11.6	99.1
90–94	11.6	99.7
95–100	11.5	100.0

$1.08/Day 1993 PPP Poverty Line Tables

Figure 5 ($1.08/day line 1993 PPP): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	90.2
5–9	82.9
10–14	49.3
15–19	38.8
20–24	28.5
25–29	31.2
30–34	15.5
35–39	10.9
40–44	5.6
45–49	7.3
50–54	0.0
55–59	0.0
60–64	0.0
65–69	0.4
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 Nepal's households.
Based on the NLSS 2003/4.

Figure 6 ($1.08/day line 1993 PPP): Derivation of estimated poverty likelihoods associated with scores

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	906	÷	1,005	=	90.2
5–9	2,205	÷	2,661	=	82.9
10–14	2,164	÷	4,394	=	49.3
15–19	2,321	÷	5,989	=	38.8
20–24	2,486	÷	8,719	=	28.5
25–29	3,418	÷	10,969	=	31.2
30–34	1,694	÷	10,911	=	15.5
35–39	1,199	÷	11,042	=	10.9
40–44	461	÷	8,176	=	5.6
45–49	504	÷	6,899	=	7.3
50–54	0	÷	6,427	=	0.0
55–59	0	÷	4,964	=	0.0
60–64	0	÷	3,870	=	0.0
65–69	16	÷	4,019	=	0.4
70–74	0	÷	3,769	=	0.0
75–79	0	÷	2,755	=	0.0
80–84	0	÷	2,133	=	0.0
85–89	0	÷	442	=	0.0
90–94	0	÷	540	=	0.0
95–100	0	÷	317	=	0.0

Surveyed cases weighted to represent Nepal's households.
Based on the NLSS 2003/4.

Figure 8 ($1.08/day line 1993 PPP): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) from the validation sample, with confidence intervals

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Score	Diff.	90-percent	95-percent	99-percent
0–4	+9.8	5.4	6.4	8.4
5–9	+21.5	3.9	4.5	5.7
10–14	–2.4	3.1	3.7	4.8
15–19	+4.1	2.5	3.0	3.8
20–24	–0.3	2.1	2.5	3.3
25–29	+7.8	1.8	2.1	2.8
30–34	–1.8	1.6	1.8	2.3
35–39	+2.8	1.1	1.3	1.7
40–44	–4.7	3.2	3.3	3.7
45–49	+2	1.1	1.2	1.7
50–54	–1.4	1.0	1.0	1.2
55–59	–1.3	0.9	1.0	1.1
60–64	+0	0.0	0.0	0.0
65–69	+0.4	0.0	0.0	0.0
70–74	+0	0.0	0.0	0.0
75–79	–0.4	0.4	0.4	0.5
80–84	+0	0.0	0.0	0.0
85–89	+0	0.0	0.0	0.0
90–94	+0	0.0	0.0	0.0
95–100	+0	0.0	0.0	0.0

Based on scorecard applied to the validation sample.

Figure 10 ($1.08/day line 1993 PPP): Differences and precision of differences for bootstrapped estimates of households' poverty likelihoods, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Sample size (n)	Diff.	90-percent	95-percent	99-percent
2	0.3	39.7	49.3	66.7
4	0.9	28.9	34.1	46.5
8	0.8	21.6	25.2	34.1
16	1.3	15.9	19.1	25.0
32	1.4	11.2	13.1	17.1
64	1.7	8.1	9.7	12.3
128	1.8	5.7	6.9	8.9
256	1.9	4.0	4.7	6.1
512	1.9	2.6	3.2	4.1
1,024	1.9	1.9	2.2	2.8
2,048	1.9	1.3	1.6	2.0
4,096	1.9	0.9	1.1	1.4
8,192	1.9	0.7	0.8	1.0
16,384	1.9	0.5	0.6	0.7

Figure 12 ($1.08/day line 1993 PPP): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Sample size (n)	Diff.	90-percent	95-percent	99-percent
2	0.3	39.7	49.3	66.7
4	1.1	28.2	34.3	43.8
8	1.1	21.2	24.7	30.1
16	1.2	14.9	17.6	22.2
32	1.3	10.2	12.4	15.2
64	1.4	7.4	8.8	11.4
128	1.4	5.3	6.2	8.0
256	1.4	3.7	4.5	5.8
512	1.4	2.7	3.2	4.0
1,024	1.4	1.8	2.2	3.0
2,048	1.4	1.3	1.5	2.0
4,096	1.4	0.9	1.1	1.5
8,192	1.4	0.7	0.8	1.0
16,384	1.4	0.5	0.5	0.8

Figure 14 ($1.08/day line 1993 PPP): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	0.8	15.2	0.2	83.8	84.6	-88.7
5–9	2.5	13.5	1.2	82.9	85.4	-61.3
10–14	4.8	11.2	3.3	80.8	85.6	-19.5
15–19	6.9	9.1	7.2	76.8	83.7	30.9
20–24	9.3	6.6	13.4	70.6	80.0	16.0
25–29	11.8	4.1	21.9	62.1	73.9	-37.2
30–34	13.7	2.3	31.0	53.1	66.8	-93.9
35–39	14.6	1.4	41.1	42.9	57.5	-157.4
40–44	15.3	0.6	48.5	35.5	50.8	-204.0
45–49	15.8	0.2	55.0	29.0	44.8	-244.4
50–54	15.8	0.1	61.3	22.7	38.5	-284.2
55–59	15.9	0.0	66.2	17.8	33.8	-314.6
60–64	15.9	0.0	70.1	13.9	29.9	-338.8
65–69	15.9	0.0	74.1	9.9	25.9	-364.0
70–74	15.9	0.0	77.9	6.2	22.1	-387.6
75–79	16.0	0.0	80.6	3.4	19.4	-404.7
80–84	16.0	0.0	82.7	1.3	17.3	-418.0
85–89	16.0	0.0	83.2	0.9	16.8	-420.8
90–94	16.0	0.0	83.7	0.3	16.3	-424.2
95–100	16.0	0.0	84.0	0.0	16.0	-426.2

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

Figure 15 ($1.08/day line 1993 PPP): Households below the poverty line and all households at a given score or at or below a given score cut-off, scorecard applied to validation sample

Score	Households at or below score who are also below the	Households at or below score (%)
0–4	79.1	1.0
5–9	68.4	3.7
10–14	59.6	8.1
15–19	48.8	14.0
20–24	41.1	22.8
25–29	35.0	33.7
30–34	30.7	44.6
35–39	26.2	55.7
40–44	24.0	63.9
45–49	22.3	70.8
50–54	20.5	77.2
55–59	19.4	82.2
60–64	18.5	86.0
65–69	17.7	90.0
70–74	17.0	93.8
75–79	16.5	96.6
80–84	16.2	98.7
85–89	16.1	99.1
90–94	16.0	99.7
95–100	16.0	100.0

$2.16/Day 1993 PPP Poverty Line Tables

Figure 5 ($2.16/day line 1993 PPP): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	95.6
15–19	97.0
20–24	83.9
25–29	89.0
30–34	73.9
35–39	61.4
40–44	55.3
45–49	44.3
50–54	37.0
55–59	12.8
60–64	14.6
65–69	4.3
70–74	2.6
75–79	6.5
80–84	1.3
85–89	0.0
90–94	0.0
95–100	0.0

Surveyed cases weighted to represent Nepal's households.
Based on the NLSS 2003/4.

Figure 6 ($2.16/day line 1993 PPP): Derivation of estimated poverty likelihoods associated with scores

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	1,005	÷	1,005	=	100.0
5–9	2,661	÷	2,661	=	100.0
10–14	4,198	÷	4,394	=	95.6
15–19	5,812	÷	5,989	=	97.0
20–24	7,316	÷	8,719	=	83.9
25–29	9,766	÷	10,969	=	89.0
30–34	8,058	÷	10,911	=	73.9
35–39	6,782	÷	11,042	=	61.4
40–44	4,521	÷	8,176	=	55.3
45–49	3,057	÷	6,899	=	44.3
50–54	2,377	÷	6,427	=	37.0
55–59	633	÷	4,964	=	12.8
60–64	565	÷	3,870	=	14.6
65–69	174	÷	4,019	=	4.3
70–74	99	÷	3,769	=	2.6
75–79	178	÷	2,755	=	6.5
80–84	28	÷	2,133	=	1.3
85–89	0	÷	442	=	0.0
90–94	0	÷	540	=	0.0
95–100	0	÷	317	=	0.0

Surveyed cases weighted to represent Nepal's households.
Based on the NLSS 2003/4.

Figure 8 ($2.16/day line 1993 PPP): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) from the validation sample, with confidence intervals

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Score	**Diff.**	**90-percent**	**95-percent**	**99-percent**
0–4	+0	0.0	0.0	0.0
5–9	+8.6	2.2	2.6	3.6
10–14	–3.5	2.0	2.0	2.1
15–19	+6.5	1.6	1.9	2.4
20–24	–2.1	1.8	2.0	2.4
25–29	+5	1.4	1.6	2.3
30–34	–3.9	2.8	3.0	3.4
35–39	–5.6	3.7	3.9	4.3
40–44	–10.1	6.1	6.3	6.9
45–49	+6.3	2.3	2.8	3.9
50–54	–3.4	2.9	3.2	3.8
55–59	–8.4	5.5	5.8	6.4
60–64	–6.5	4.7	5.0	5.6
65–69	–5.2	3.5	3.7	4.0
70–74	–1.1	1.1	1.2	1.6
75–79	+6.1	0.3	0.4	0.5
80–84	+1.3	0.0	0.0	0.0
85–89	+0	0.0	0.0	0.0
90–94	+0	0.0	0.0	0.0
95–100	+0	0.0	0.0	0.0

Based on scorecard applied to the validation sample.

Figure 10 ($2.16/day line 1993 PPP): Differences and precision of differences for bootstrapped estimates of households' poverty likelihoods, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Sample size (n)	Diff.	90-percent	95-percent	99-percent
2	0.5	46.4	55.4	73.0
4	-1.1	33.3	39.1	48.2
8	-1.6	23.4	26.9	36.5
16	-1.5	17.4	19.9	25.0
32	-1.6	12.3	14.6	18.9
64	-1.5	8.0	9.9	13.0
128	-1.2	5.6	6.6	8.7
256	-1.0	3.5	4.0	5.4
512	-1.0	2.3	2.8	3.7
1,024	-1.0	1.8	2.1	2.5
2,048	-1.0	1.1	1.4	1.9
4,096	-0.9	0.8	0.9	1.2
8,192	-0.9	0.6	0.7	0.9
16,384	-0.9	0.4	0.5	0.7

Figure 12 ($2.16/day line 1993 PPP): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Sample size (n)	Diff.	90-percent	95-percent	99-percent
2	0.5	46.4	55.4	73.0
4	-1.0	32.6	38.7	46.4
8	-1.6	23.4	26.7	34.6
16	-1.7	16.9	19.0	25.2
32	-1.7	11.9	13.7	17.7
64	-1.8	8.5	10.1	13.3
128	-1.6	5.9	7.1	9.4
256	-1.6	4.2	5.1	6.5
512	-1.6	3.1	3.5	4.6
1,024	-1.6	2.2	2.6	3.4
2,048	-1.6	1.4	1.8	2.5
4,096	-1.6	1.1	1.2	1.6
8,192	-1.6	0.7	0.9	1.1
16,384	-1.6	0.5	0.6	0.8

Figure 14 ($2.16/day line 1993 PPP): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	1.0	57.8	0.0	41.2	42.2	-96.6
5–9	3.4	55.4	0.2	41.0	44.4	-87.9
10–14	7.8	51.0	0.3	40.9	48.7	-73.1
15–19	13.2	45.6	0.9	40.3	53.5	-53.7
20–24	20.8	38.0	2.0	39.2	60.0	-25.9
25–29	29.9	28.8	3.8	37.4	67.4	8.3
30–34	38.4	20.4	6.3	35.0	73.3	41.3
35–39	45.8	13.0	9.9	31.3	77.2	72.7
40–44	51.1	7.6	12.7	28.5	79.6	78.3
45–49	53.9	4.9	16.9	24.4	78.3	71.3
50–54	56.3	2.4	20.9	20.4	76.7	64.5
55–59	57.3	1.5	24.8	16.4	73.7	57.7
60–64	58.1	0.7	27.9	13.3	71.4	52.5
65–69	58.5	0.2	31.5	9.7	68.2	46.4
70–74	58.7	0.0	35.1	6.2	64.9	40.3
75–79	58.8	0.0	37.8	3.4	62.2	35.7
80–84	58.8	0.0	39.9	1.3	60.1	32.1
85–89	58.8	0.0	40.4	0.9	59.6	31.3
90–94	58.8	0.0	40.9	0.3	59.1	30.4
95–100	58.8	0.0	41.2	0.0	58.8	29.9

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

Figure 15 ($2.16/day line 1993 PPP): Households below the poverty line and all households at a given score or at or below a given score cut-off, scorecard applied to validation sample

Score	Households at or below score who are also below the	Households at or below score (%)
0–4	100.0	1.0
5–9	93.4	3.7
10–14	96.2	8.1
15–19	93.7	14.0
20–24	91.2	22.8
25–29	88.8	33.7
30–34	86.0	44.6
35–39	82.3	55.7
40–44	80.1	63.9
45–49	76.2	70.8
50–54	73.0	77.2
55–59	69.8	82.2
60–64	67.5	86.0
65–69	65.0	90.0
70–74	62.6	93.8
75–79	60.9	96.6
80–84	59.6	98.7
85–89	59.3	99.1
90–94	59.0	99.7
95–100	58.8	100.0

$3.24/Day 1993 PPP Poverty Line Tables

Figure 5 ($3.24/day line 1993 PPP): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	100.0
15–19	100.0
20–24	98.3
25–29	97.2
30–34	94.0
35–39	90.1
40–44	83.8
45–49	77.8
50–54	72.4
55–59	56.1
60–64	54.5
65–69	34.9
70–74	22.2
75–79	11.6
80–84	9.7
85–89	0.0
90–94	0.0
95–100	0.0

Surveyed cases weighted to represent Nepal's households.
Based on the NLSS 2003/4.

Figure 6 ($3.24/day line 1993 PPP): Derivation of estimated poverty likelihoods associated with scores

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	1,005	÷	1,005	=	100.0
5–9	2,661	÷	2,661	=	100.0
10–14	4,394	÷	4,394	=	100.0
15–19	5,989	÷	5,989	=	100.0
20–24	8,574	÷	8,719	=	98.3
25–29	10,666	÷	10,969	=	97.2
30–34	10,256	÷	10,911	=	94.0
35–39	9,951	÷	11,042	=	90.1
40–44	6,851	÷	8,176	=	83.8
45–49	5,367	÷	6,899	=	77.8
50–54	4,651	÷	6,427	=	72.4
55–59	2,786	÷	4,964	=	56.1
60–64	2,110	÷	3,870	=	54.5
65–69	1,404	÷	4,019	=	34.9
70–74	835	÷	3,769	=	22.2
75–79	319	÷	2,755	=	11.6
80–84	207	÷	2,133	=	9.7
85–89	0	÷	442	=	0.0
90–94	0	÷	540	=	0.0
95–100	0	÷	317	=	0.0

Surveyed cases weighted to represent Nepal's households.
Based on the NLSS 2003/4.

Figure 8 ($3.24/day line 1993 PPP): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) from the validation sample, with confidence intervals

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Score	**Diff.**	**90-percent**	**95-percent**	**99-percent**
0–4	+0	0.0	0.0	0.0
5–9	+0	0.0	0.0	0.0
10–14	+0	0.0	0.0	0.0
15–19	+0	0.0	0.0	0.0
20–24	+1	0.7	0.9	1.2
25–29	+0.3	0.7	0.8	1.0
30–34	–1.7	1.2	1.3	1.5
35–39	–4.6	2.8	2.8	3.0
40–44	–3.8	2.6	2.7	3.0
45–49	–4.5	3.2	3.4	3.7
50–54	+2.1	2.3	2.8	3.7
55–59	–1.3	3.0	3.7	4.9
60–64	+8.5	3.5	4.1	5.2
65–69	–2.7	3.3	4.0	5.3
70–74	–16.2	9.9	10.2	11.0
75–79	–0.2	2.7	3.2	4.3
80–84	+3.6	2.5	3.0	3.8
85–89	+0	0.0	0.0	0.0
90–94	–5.6	4.9	5.3	6.3
95–100	+0	0.0	0.0	0.0

Based on scorecard applied to the validation sample.

Figure 10 ($3.24/day line 1993 PPP): Differences and precision of differences for bootstrapped estimates of households' poverty likelihoods, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Sample size (n)	Diff.	90-percent	95-percent	99-percent
2	-0.3	36.4	45.8	67.2
4	-1.0	25.3	32.2	42.0
8	-1.5	18.9	22.5	30.0
16	-1.8	14.6	17.4	22.6
32	-1.5	11.4	13.7	17.7
64	-1.4	9.0	10.5	13.9
128	-1.3	5.8	6.9	9.3
256	-1.2	3.7	4.4	6.0
512	-1.3	2.7	3.2	4.6
1,024	-1.3	1.9	2.2	2.9
2,048	-1.3	1.3	1.5	1.9
4,096	-1.3	0.9	1.1	1.3
8,192	-1.3	0.6	0.7	1.0
16,384	-1.3	0.5	0.5	0.7

Figure 12 ($3.24/day line 1993 PPP): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Sample size (n)	Diff.	90-percent	95-percent	99-percent
2	-0.3	36.4	45.8	67.2
4	-0.9	24.9	30.4	40.5
8	-1.4	17.4	20.7	27.2
16	-1.6	12.6	15.3	18.9
32	-1.5	8.9	10.4	13.2
64	-1.5	6.6	7.8	9.9
128	-1.4	4.2	5.0	6.9
256	-1.4	3.1	3.7	5.2
512	-1.5	2.2	2.6	3.8
1,024	-1.5	1.6	1.9	2.6
2,048	-1.5	1.1	1.3	1.6
4,096	-1.5	0.8	0.9	1.2
8,192	-1.5	0.6	0.7	0.8
16,384	-1.5	0.4	0.5	0.6

Figure 14 ($3.24/day line 1993 PPP): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	1.0	78.0	0.0	21.0	22.0	-97.5
5–9	3.7	75.3	0.0	21.0	24.7	-90.7
10–14	8.1	70.9	0.0	21.0	29.1	-79.6
15–19	14.0	65.0	0.0	21.0	35.0	-64.4
20–24	22.6	56.4	0.2	20.8	43.3	-42.6
25–29	33.2	45.8	0.6	20.4	53.6	-15.3
30–34	43.6	35.4	1.1	19.9	63.5	11.6
35–39	54.0	25.0	1.7	19.3	73.4	38.9
40–44	61.2	17.8	2.7	18.3	79.5	58.3
45–49	66.9	12.1	3.9	17.1	84.0	74.2
50–54	71.3	7.7	5.9	15.1	86.4	87.9
55–59	74.1	4.9	8.1	12.9	87.0	89.8
60–64	76.0	3.1	10.1	10.9	86.9	87.2
65–69	77.4	1.6	12.7	8.3	85.7	84.0
70–74	78.6	0.4	15.2	5.8	84.3	80.7
75–79	78.9	0.1	17.7	3.3	82.2	77.6
80–84	79.0	0.0	19.7	1.3	80.2	75.0
85–89	79.0	0.0	20.2	0.8	79.8	74.5
90–94	79.0	0.0	20.7	0.3	79.3	73.8
95–100	79.0	0.0	21.0	0.0	79.0	73.4

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

Figure 15 ($3.24/day line 1993 PPP): Households below the poverty line and all households at a given score or at or below a given score cut-off, scorecard applied to validation sample

Score	Households at or below score who are also below the	Households at or below score (%)
0–4	100.0	1.0
5–9	100.0	3.7
10–14	100.0	8.1
15–19	100.0	14.0
20–24	99.1	22.8
25–29	98.4	33.7
30–34	97.6	44.6
35–39	97.0	55.7
40–44	95.8	63.9
45–49	94.5	70.8
50–54	92.3	77.2
55–59	90.2	82.2
60–64	88.3	86.0
65–69	85.9	90.0
70–74	83.8	93.8
75–79	81.7	96.6
80–84	80.0	98.7
85–89	79.7	99.1
90–94	79.3	99.7
95–100	79.0	100.0

$1.25/Day 2005 PPP Poverty Line Tables

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

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	99.0
10–14	87.9
15–19	91.4
20–24	76.7
25–29	75.4
30–34	60.7
35–39	47.1
40–44	34.4
45–49	30.0
50–54	17.3
55–59	8.8
60–64	0.0
65–69	0.4
70–74	1.5
75–79	3.5
80–84	0.0
85–89	0.0
90–94	0.0
95–100	0.0

Surveyed cases weighted to represent Nepal's households.
Based on the NLSS 2003/4.

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

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	1,005	÷	1,005	=	100.0
5–9	2,635	÷	2,661	=	99.0
10–14	3,863	÷	4,394	=	87.9
15–19	5,473	÷	5,989	=	91.4
20–24	6,689	÷	8,719	=	76.7
25–29	8,268	÷	10,969	=	75.4
30–34	6,625	÷	10,911	=	60.7
35–39	5,196	÷	11,042	=	47.1
40–44	2,811	÷	8,176	=	34.4
45–49	2,072	÷	6,899	=	30.0
50–54	1,114	÷	6,427	=	17.3
55–59	439	÷	4,964	=	8.8
60–64	0	÷	3,870	=	0.0
65–69	16	÷	4,019	=	0.4
70–74	55	÷	3,769	=	1.5
75–79	95	÷	2,755	=	3.5
80–84	0	÷	2,133	=	0.0
85–89	0	÷	442	=	0.0
90–94	0	÷	540	=	0.0
95–100	0	÷	317	=	0.0

Surveyed cases weighted to represent Nepal's households.
Based on the NLSS 2003/4.

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

	Scorecard applied to the validation sample Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Score	**Diff.**	**90-percent**	**95-percent**	**99-percent**
0–4	+0	0.0	0.0	0.0
5–9	+7.6	2.2	2.7	3.6
10–14	+0.7	2.0	2.5	3.4
15–19	+8.5	2.1	2.5	3.6
20–24	–0.1	1.9	2.2	2.7
25–29	+4.8	1.9	2.2	3.1
30–34	–8.9	5.3	5.5	5.9
35–39	–8.3	5.2	5.4	5.8
40–44	–6.4	4.3	4.6	5.1
45–49	+10.5	1.9	2.3	2.8
50–54	–4.1	3.1	3.4	3.8
55–59	+3.3	1.2	1.4	1.9
60–64	–8.4	5.1	5.3	5.7
65–69	–2.9	2.0	2.1	2.3
70–74	+0.3	0.5	0.6	0.8
75–79	+3.1	0.3	0.4	0.5
80–84	+0	0.0	0.0	0.0
85–89	+0	0.0	0.0	0.0
90–94	+0	0.0	0.0	0.0
95–100	+0	0.0	0.0	0.0

Based on scorecard applied to the validation sample.

Figure 10 ($1.25/day line 2005 PPP): Differences and precision of differences for bootstrapped estimates of households' poverty likelihoods, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Sample size (n)	Diff.	90-percent	95-percent	99-percent
2	1.2	48.4	56.5	73.6
4	0.1	32.3	37.2	50.6
8	-0.9	23.8	28.3	40.4
16	-0.9	17.0	21.0	27.4
32	-0.4	11.7	14.4	20.6
64	-0.3	7.9	10.0	13.1
128	-0.2	4.9	6.1	8.4
256	-0.1	3.2	3.8	4.9
512	-0.1	2.1	2.6	3.3
1,024	-0.1	1.5	1.8	2.3
2,048	0.0	1.1	1.3	1.8
4,096	0.0	0.8	0.9	1.2
8,192	0.0	0.5	0.6	0.8
16,384	0.0	0.4	0.4	0.6

Figure 12 ($1.25/day line 2005 PPP): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Sample size (n)	Diff.	90-percent	95-percent	99-percent
2	1.2	48.4	56.5	73.6
4	0.2	32.5	39.1	47.6
8	-0.6	23.8	28.4	38.2
16	-1.1	16.9	20.7	29.4
32	-1.0	12.2	14.8	20.4
64	-1.1	8.7	10.3	14.8
128	-1.0	6.1	7.3	9.4
256	-1.0	4.2	5.2	6.5
512	-1.0	3.1	3.6	4.9
1,024	-0.9	2.1	2.6	3.3
2,048	-0.9	1.6	1.9	2.6
4,096	-0.9	1.1	1.3	1.6
8,192	-0.9	0.8	0.9	1.2
16,384	-0.9	0.5	0.6	0.9

Figure 14 ($1.25/day line 2005 PPP): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	1.0	46.6	0.0	52.4	53.4	-95.8
5–9	3.4	44.1	0.2	52.2	55.6	-85.1
10–14	7.3	40.3	0.8	51.6	58.9	-67.8
15–19	12.3	35.3	1.8	50.7	62.9	-44.7
20–24	19.0	28.5	3.7	48.7	67.7	-12.1
25–29	26.8	20.8	7.0	45.5	72.2	27.2
30–34	34.3	13.3	10.4	42.0	76.3	65.9
35–39	40.3	7.2	15.4	37.1	77.4	67.7
40–44	43.7	3.9	20.2	32.3	76.0	57.6
45–49	45.2	2.4	25.6	26.9	72.0	46.2
50–54	46.6	1.0	30.6	21.8	68.4	35.6
55–59	46.9	0.6	35.2	17.2	64.1	25.9
60–64	47.3	0.3	38.8	13.7	61.0	18.5
65–69	47.5	0.1	42.6	9.8	57.3	10.5
70–74	47.5	0.0	46.3	6.2	53.7	2.7
75–79	47.6	0.0	49.0	3.4	51.0	-3.0
80–84	47.6	0.0	51.1	1.3	48.9	-7.5
85–89	47.6	0.0	51.6	0.9	48.4	-8.4
90–94	47.6	0.0	52.1	0.3	47.9	-9.6
95–100	47.6	0.0	52.4	0.0	47.6	-10.2

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

Figure 15 ($1.25/day line 2005 PPP): Households below the poverty line and all households at a given score or at or below a given score cut-off, scorecard applied to validation sample

Score	Households at or below score who are also below the	Households at or below score (%)
0–4	100.0	1.0
5–9	93.4	3.7
10–14	90.1	8.1
15–19	87.3	14.0
20–24	83.5	22.8
25–29	79.3	33.7
30–34	76.7	44.6
35–39	72.4	55.7
40–44	68.4	63.9
45–49	63.9	70.8
50–54	60.3	77.2
55–59	57.1	82.2
60–64	55.0	86.0
65–69	52.7	90.0
70–74	50.7	93.8
75–79	49.3	96.6
80–84	48.2	98.7
85–89	48.0	99.1
90–94	47.7	99.7
95–100	47.6	100.0

$2.50/Day 2005 PPP Poverty Line Tables

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

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	100.0
15–19	100.0
20–24	98.9
25–29	97.8
30–34	94.0
35–39	93.8
40–44	93.0
45–49	82.9
50–54	79.4
55–59	66.9
60–64	60.1
65–69	47.0
70–74	26.5
75–79	21.7
80–84	19.5
85–89	8.6
90–94	11.0
95–100	0.0

Surveyed cases weighted to represent Nepal's households.
Based on the NLSS 2003/4.

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

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	1,005	÷	1,005	=	100.0
5–9	2,661	÷	2,661	=	100.0
10–14	4,394	÷	4,394	=	100.0
15–19	5,989	÷	5,989	=	100.0
20–24	8,626	÷	8,719	=	98.9
25–29	10,730	÷	10,969	=	97.8
30–34	10,256	÷	10,911	=	94.0
35–39	10,362	÷	11,042	=	93.8
40–44	7,606	÷	8,176	=	93.0
45–49	5,721	÷	6,899	=	82.9
50–54	5,101	÷	6,427	=	79.4
55–59	3,320	÷	4,964	=	66.9
60–64	2,325	÷	3,870	=	60.1
65–69	1,888	÷	4,019	=	47.0
70–74	997	÷	3,769	=	26.5
75–79	597	÷	2,755	=	21.7
80–84	416	÷	2,133	=	19.5
85–89	38	÷	442	=	8.6
90–94	59	÷	540	=	11.0
95–100	0	÷	317	=	0.0

Surveyed cases weighted to represent Nepal's households.
Based on the NLSS 2003/4.

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

	Scorecard applied to the validation sample Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	**Diff.**	**90-percent**	**95-percent**	**99-percent**
0–4	0.0	0.0	0.0	0.0
5–9	0.0	0.0	0.0	0.0
10–14	0.0	0.0	0.0	0.0
15–19	0.0	0.0	0.0	0.0
20–24	0.1	0.5	0.6	0.7
25–29	0.9	0.7	0.8	1.1
30–34	-3.0	1.8	1.9	2.0
35–39	-0.9	0.9	1.1	1.3
40–44	-0.4	1.1	1.3	1.7
45–49	-3.4	2.5	2.6	3.1
50–54	-3.4	2.7	2.8	3.2
55–59	7.8	2.9	3.7	4.8
60–64	5.9	3.6	4.3	5.6
65–69	3.2	3.5	4.3	5.4
70–74	-15.6	9.6	9.9	10.6
75–79	2.8	3.3	4.0	5.1
80–84	13.4	2.5	3.0	3.8
85–89	8.6	0.0	0.0	0.0
90–94	5.3	3.7	4.4	5.9
95–100	0.0	0.0	0.0	0.0

Based on scorecard applied to the validation sample.

Figure 10 ($2.50/day line 2005 PPP): Differences and precision of differences for bootstrapped estimates of households' poverty likelihoods, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Sample size (n)	Diff.	90-percent	95-percent	99-percent
2	0.5	34.6	43.8	61.1
4	-0.1	24.4	30.4	41.6
8	-0.1	18.5	22.4	27.2
16	-0.4	14.2	16.5	21.2
32	0.0	11.1	13.3	17.3
64	0.2	8.6	10.7	13.5
128	0.4	5.9	6.9	9.3
256	0.6	3.8	4.6	5.7
512	0.6	2.7	3.3	4.3
1,024	0.6	1.8	2.3	3.0
2,048	0.7	1.3	1.5	2.0
4,096	0.7	0.9	1.0	1.3
8,192	0.7	0.6	0.8	1.0
16,384	0.7	0.5	0.5	0.7

Figure 12 ($2.50/day line 2005 PPP): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Sample size (n)	Diff.	90-percent	95-percent	99-percent
2	0.5	34.6	43.8	61.1
4	0.0	22.8	27.9	40.9
8	-0.1	16.7	20.8	24.9
16	-0.3	12.1	14.0	17.4
32	-0.3	8.3	9.8	12.8
64	-0.3	6.0	7.0	9.3
128	-0.3	4.1	4.8	6.5
256	-0.3	2.8	3.5	4.5
512	-0.3	2.1	2.5	3.4
1,024	-0.3	1.5	1.8	2.4
2,048	-0.3	1.0	1.2	1.5
4,096	-0.3	0.7	0.9	1.1
8,192	-0.4	0.5	0.6	0.8
16,384	-0.3	0.4	0.4	0.6

Figure 14 ($2.50/day line 2005 PPP): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to validation sample

Score	< poverty line correctly targeted	< poverty line mistakenly non-targeted	=> poverty line mistakenly targeted	=> poverty line correctly non-targeted	Inclusion + Exclusion	See text
0–4	1.0	80.9	0.0	18.1	19.1	-97.5
5–9	3.7	78.3	0.0	18.1	21.7	-91.1
10–14	8.1	73.9	0.0	18.1	26.1	-80.3
15–19	14.0	67.9	0.0	18.1	32.1	-65.7
20–24	22.7	59.3	0.1	18.0	40.6	-44.5
25–29	33.3	48.6	0.4	17.6	50.9	-18.2
30–34	43.8	38.1	0.8	17.2	61.1	8.0
35–39	54.3	27.6	1.4	16.7	71.0	34.2
40–44	61.9	20.0	1.9	16.1	78.1	53.5
45–49	67.9	14.0	2.9	15.2	83.1	69.2
50–54	73.1	8.8	4.0	14.0	87.2	83.5
55–59	76.1	5.8	6.1	12.0	88.1	92.6
60–64	78.2	3.7	7.8	10.3	88.5	90.5
65–69	79.9	2.0	10.1	8.0	87.9	87.7
70–74	81.3	0.6	12.5	5.6	86.9	84.8
75–79	81.8	0.1	14.8	3.3	85.1	82.0
80–84	81.9	0.0	16.8	1.3	83.2	79.5
85–89	81.9	0.0	17.2	0.8	82.7	79.0
90–94	81.9	0.0	17.7	0.3	82.3	78.3
95–100	81.9	0.0	18.1	0.0	81.9	78.0

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

Figure 15 ($2.50/day line 2005 PPP): Households below the poverty line and all households at a given score or at or below a given score cut-off, scorecard applied to validation sample

Score	Households at or below score who are also below the poverty line (%)	Households at or below score (%)
0–4	100.0	1.0
5–9	100.0	3.7
10–14	100.0	8.1
15–19	100.0	14.0
20–24	99.6	22.8
25–29	98.7	33.7
30–34	98.1	44.6
35–39	97.5	55.7
40–44	97.0	63.9
45–49	96.0	70.8
50–54	94.8	77.2
55–59	92.6	82.2
60–64	90.9	86.0
65–69	88.8	90.0
70–74	86.7	93.8
75–79	84.7	96.6
80–84	83.0	98.7
85–89	82.6	99.1
90–94	82.2	99.7
95–100	81.9	100.0

$3.75/Day 2005 PPP Poverty Line Tables

Figure 5 ($3.75/day line 2005 PPP): Estimated poverty likelihoods associated with scores

If a household's score is . . .	. . . then the likelihood (%) of being below the poverty line is:
0–4	100.0
5–9	100.0
10–14	100.0
15–19	100.0
20–24	100.0
25–29	99.1
30–34	98.1
35–39	96.6
40–44	98.6
45–49	92.7
50–54	97.5
55–59	95.2
60–64	82.4
65–69	75.5
70–74	56.4
75–79	50.7
80–84	38.6
85–89	11.1
90–94	51.6
95–100	0.0

Surveyed cases weighted to represent Nepal's households.
Based on the NLSS 2003/4.

Figure 6 ($3.75/day line 2005 PPP): Derivation of estimated poverty likelihoods associated with scores

Score	Households below poverty line		All households at score		Poverty likelihood (estimated, %)
0–4	1,005	÷	1,005	=	100.0
5–9	2,661	÷	2,661	=	100.0
10–14	4,394	÷	4,394	=	100.0
15–19	5,989	÷	5,989	=	100.0
20–24	8,719	÷	8,719	=	100.0
25–29	10,870	÷	10,969	=	99.1
30–34	10,701	÷	10,911	=	98.1
35–39	10,664	÷	11,042	=	96.6
40–44	8,059	÷	8,176	=	98.6
45–49	6,395	÷	6,899	=	92.7
50–54	6,265	÷	6,427	=	97.5
55–59	4,725	÷	4,964	=	95.2
60–64	3,190	÷	3,870	=	82.4
65–69	3,035	÷	4,019	=	75.5
70–74	2,124	÷	3,769	=	56.4
75–79	1,397	÷	2,755	=	50.7
80–84	824	÷	2,133	=	38.6
85–89	49	÷	442	=	11.1
90–94	279	÷	540	=	51.6
95–100	0	÷	317	=	0.0

Surveyed cases weighted to represent Nepal's households.

Figure 8 ($3.75/day line 2005 PPP): Bootstrapped differences between estimated and true poverty likelihoods for households in a large sample (n = 16,384) from the validation sample, with confidence intervals

	Scorecard applied to the validation sample Difference between estimate and true value			
		Confidence interval (+/− percentage points)		
Score	**Diff.**	**90-percent**	**95-percent**	**99-percent**
0–4	0.0	0.0	0.0	0.0
5–9	0.0	0.0	0.0	0.0
10–14	0.0	0.0	0.0	0.0
15–19	0.0	0.0	0.0	0.0
20–24	0.0	0.0	0.0	0.0
25–29	0.0	0.4	0.5	0.6
30–34	-1.5	0.9	0.9	0.9
35–39	-3.1	1.6	1.6	1.7
40–44	-0.9	0.6	0.6	0.6
45–49	-5.8	3.2	3.3	3.3
50–54	2.0	1.1	1.3	1.6
55–59	12.7	2.3	2.8	3.6
60–64	3.8	3.0	3.6	4.6
65–69	10.6	3.3	3.9	5.1
70–74	-11.8	7.4	7.8	8.2
75–79	10.1	4.3	5.1	6.3
80–84	1.2	4.7	5.5	7.5
85–89	-4.9	6.7	7.9	10.5
90–94	28.9	6.9	8.2	10.5
95–100	0.0	0.0	0.0	0.0

Based on scorecard applied to the validation sample.

Figure 10 ($3.75/day line 2005 PPP): Differences and precision of differences for bootstrapped estimates of households' poverty likelihoods, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Sample size (n)	Diff.	90-percent	95-percent	99-percent
2	0.5	26.8	34.2	46.5
4	0.0	17.0	21.0	28.2
8	0.4	12.9	16.0	23.5
16	0.5	10.7	13.5	17.7
32	0.6	9.0	11.1	14.8
64	0.9	7.3	8.9	11.7
128	1.2	5.8	6.7	8.8
256	1.4	4.0	4.8	6.6
512	1.4	3.2	3.7	4.6
1,024	1.4	2.2	2.6	3.4
2,048	1.4	1.5	1.7	2.2
4,096	1.4	1.0	1.2	1.6
8,192	1.4	0.7	0.9	1.1
16,384	1.5	0.5	0.7	0.9

Figure 12 ($3.75/day line 2005 PPP): Differences and precision of differences for bootstrapped estimates of poverty rates for groups of households at a point in time, by sample size, scorecard applied to validation sample

	Difference between estimate and true value			
		Confidence interval (+/– percentage points)		
Sample size (n)	Diff.	90-percent	95-percent	99-percent
2	0.5	26.8	34.2	46.5
4	0.0	16.4	20.3	27.4
8	0.2	11.5	13.9	19.9
16	0.2	8.5	10.4	13.7
32	0.1	5.9	7.2	9.6
64	0.1	4.1	5.1	7.1
128	0.1	3.0	3.5	4.9
256	0.1	2.1	2.6	3.3
512	0.1	1.6	1.8	2.4
1,024	0.1	1.1	1.3	1.7
2,048	0.1	0.7	0.9	1.2
4,096	0.1	0.5	0.6	0.8
8,192	0.1	0.4	0.4	0.6
16,384	0.1	0.3	0.3	0.4

Figure 14 ($3.75/day line 2005 PPP): Households by targeting classification and score, along with "Total Accuracy" and BPAC, scorecard applied to validation sample

Score	Inclusion: < poverty line correctly targeted	Undercoverage: < poverty line mistakenly non-targeted	Leakage: => poverty line mistakenly targeted	Exclusion: => poverty line correctly non-targeted	Total Accuracy Inclusion + Exclusion	BPAC See text
0–4	1.0	89.5	0.0	9.5	10.5	-97.8
5–9	3.7	86.9	0.0	9.5	13.1	-91.9
10–14	8.1	82.5	0.0	9.5	17.5	-82.2
15–19	14.0	76.5	0.0	9.5	23.5	-69.0
20–24	22.8	67.8	0.0	9.5	32.2	-49.7
25–29	33.6	56.9	0.1	9.4	43.0	-25.6
30–34	44.5	46.1	0.2	9.3	53.8	-1.6
35–39	55.4	35.1	0.2	9.2	64.7	22.7
40–44	63.6	27.0	0.3	9.2	72.7	40.7
45–49	70.4	20.2	0.4	9.1	79.4	55.9
50–54	76.5	14.1	0.7	8.7	85.2	69.7
55–59	80.5	10.0	1.6	7.9	88.4	79.7
60–64	83.6	7.0	2.4	7.0	90.6	87.3
65–69	86.2	4.4	3.9	5.6	91.8	94.6
70–74	88.6	2.0	5.2	4.2	92.8	94.2
75–79	89.6	0.9	6.9	2.5	92.2	92.3
80–84	90.3	0.2	8.4	1.1	91.4	90.7
85–89	90.4	0.1	8.8	0.7	91.1	90.3
90–94	90.5	0.0	9.1	0.3	90.9	89.9
95–100	90.5	0.0	9.5	0.0	90.5	89.5

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

Figure 15 ($3.75/day line 2005 PPP): Households below the poverty line and all households at a given score or at or below a given score cut-off, scorecard applied to validation sample

Score	Households at or below score who are also below the poverty line (%)	Households at or below score (%)
0–4	100.0	1.0
5–9	100.0	3.7
10–14	100.0	8.1
15–19	100.0	14.0
20–24	100.0	22.8
25–29	99.7	33.7
30–34	99.6	44.6
35–39	99.6	55.7
40–44	99.5	63.9
45–49	99.4	70.8
50–54	99.0	77.2
55–59	98.0	82.2
60–64	97.2	86.0
65–69	95.7	90.0
70–74	94.4	93.8
75–79	92.8	96.6
80–84	91.5	98.7
85–89	91.2	99.1
90–94	90.8	99.7
95–100	90.5	100.0