A Simple Poverty Scorecard for Mexico

Mark Schreiner

November 2, 2006

Senior Scholar, Center for Social Development
Washington University in Saint Louis
Campus Box 1196, One Brookings Drive
Saint Louis, MO 63130-4899, U.S.A.
and
Director, Microfinance Risk Management, L.L.C.
6970 Chippewa St. #1W, Saint Louis, MO 63109-3060, U.S.A.
Telephone: +1 (314) 481-9788, `http://www.microfinance.com`

Abstract

How poor are participants in development projects in Mexico? This research uses a national survey to construct an easy-to-use, objective scorecard that estimates the likelihood that a person has income below the national poverty line. The scorecard uses 10 simple indicators that field workers can quickly collect and easily verify. Poverty scores can be computed on paper in the field in real time. With 90-percent confidence, a group's estimated overall poverty rate is accurate within +/–1.0 percentage points. The poverty scorecard can help programs to target services, track changes in poverty over time, and report on poverty rates.

Acknowledgements

This paper was commissioned by Grameen Foundation USA and funded by the Consultative Group to Assist the Poorest under the CGAP-Ford Social Indicators Project. Data was provided by Mexico's Instituto Nacional de Estadística. I am grateful for help from Nigel Bigger, Gretel Guzmán, Sergio Navajas, Luca David Opramolla, Erica Gabriela Rascón Ramírez, and Jeff Toohig. An earlier paper that tested 5–, 10–, and 15-indicator scorecards and that also tested scorecards segments by urban, peri-urban, and rural was circulated in Spanish as "Un índice de pobreza para México".

A Simple Poverty Scorecard for Mexico

1. Introduction

This paper presents an easy-to-use, objective poverty scorecard to help development programs in Mexico to target services, track changes in poverty over time, and report clients' poverty rates.

Rather than asking for hours on end about all possible sources of income ("Did you raise any cattle in the past year? How many were born? How many died? Did you eat them? If you had sold them, what price would they have fetched in the market? Now then, did you raise any pigs in the past year? . . ."), the scorecard uses 10 simple indicators (such as "Does the house have a bathroom?" or "What type of fuel does the household use for cooking?") to produce a score that is highly correlated with poverty status measured by the exhaustive survey.

Indicators in the scorecard were derived from an analysis of 17,167 households surveyed in the 2002 *Encuesta Nacional de Ingresos y Gastos de Hogares* (ENIGH). Indicators were 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 scorecard weights are positive integers, and scores range from 0 (most-likely "poor") to 100 (least-likely "poor"). Field workers can compute scores by hand, on paper, in real time.

A participant's score corresponds to a "poverty likelihood", that is, the probability of being poor. In a group, the share of clients who are poor is defined as their average poverty likelihood. For a given group over time, progress (or regress) is the change in average poverty likelihood.

The scorecard here was constructed for use in all of Mexico. Schreiner (2005a) finds only small returns to segmenting Mexican poverty scorecards by rural and urban.

The scorecard can also be used to classify clients as very poor (poorest half in poverty), poor (top half in poverty), or non-poor.

The scorecard accurately and objectively estimates the likelihood that households have income below the national poverty line. It should qualify for certification for the reporting required of USAID's microenterprise partners. In particular, the scorecard is accurate. With 90-percent confidence, a group's estimated overall poverty rate is accurate within +/–1.0 percentage points.

2. Data and poverty lines

The analysis here uses the 17,167 households in the 2002 ENIGH. This is the best, most recent household survey with income data available for Mexico. This paper divides households into three random samples (Figure 1), with one-half used for constructing the scorecard, one-fourth used for associating scores with estimated poverty likelihoods via a bootstrap method, and the final one-fourth used for measuring the accuracy of estimates derived from the scorecard, again via the bootstrap. The average surveyed household represented about 5,900 people. The 197 households who each represented more than 40,000 people were omitted because they led to the breakdown of some bootstrap estimates (see Singh, 1998).

Poverty in Mexico is wide and deep. The overall poverty rate ("head count index") in 2002 was 51.7 percent (Figure 1). The urban rate was 42.0 percent, and rural was 67.5 percent.

The official poverty line was established by the Comité Técnico para la Medición de la Pobreza (2002), see also Rascón Ramírez (2002). It represents the income required for basic nutrition, health care, education, clothing, housing, and transport. (Mexico does not have an official expenditure-based poverty line.) The poverty line does not make adjustments for household size. For rural areas (less than 15,000 people), the poverty line is 31.12 pesos/person/day ($3.59 at purchase-power parity, see Sillers, 2006). For urban areas (15,000 people or more), the poverty line is 44.95 pesos/person/day ($5.18 at purchase-power parity).

3. Scorecard construction

More than 2,000 potential poverty indicators were prepared, including:

- Household and housing characteristics (such as cooking fuel and type of floor)
- Individual characteristics (such as age and highest grade completed)
- Household consumption (such as milk and apples)
- Household durable goods (such as electric fans and telephones)

How well each indicator predicts poverty was tested first with the entropy-based "uncertainty coefficient" (Goodman and Kruskal, 1979), with about 200 indicators selected for further analysis. Figures 2 and 3 list (in English and Spanish) the top 50, ranked by their uncertainty coefficients. For a given indicator, responses are ordered by the strength of their association with poverty.

Many indicators in Figures 2 and 3 are similar in terms of their association with poverty. For example, most households who have a bathroom also have a toilet supplied by piped water. If a scorecard already includes "has a bathroom", then "toilet supplied by piped water" is superfluous. Thus, many indicators strongly associated with poverty are not needed because similar indicators are already included.

The scorecard also aims to measure *changes* in poverty through time. Thus, some powerful indicators (such as education of the female spouse/head) that are unlikely to change as poverty changes were left out in favor of slightly less-powerful indicators that are more likely to change (such as ownership of a stereo). Some other powerful indicators (such as "In the past six months, did anyone in the household shop at a supermarket or department store") were not selected because they are not verifiable.

The scorecard itself is constructed using Logit regression. Indicator selection combines statistics with the judgment of an analyst with expertise in scoring and development. Starting with a scorecard with no indicators, each candidate indicator is added, one-by-one, to a one-indicator scorecard, using Logit to derive weights. The improvement in accuracy for each indicator is recorded using the "c" statistic.[1]

After all indicators are tested, one is selected based on several factors (Schreiner *et al.*, 2004; Zeller, 2004). These include the improvement in accuracy, the likelihood of acceptance by users (determined by simplicity, cost of collection, and "face validity" in terms of experience, theory, and common sense), the ability of the indicator to change values as poverty status changes, variety vis-à-vis other indicators already in the scorecard, and observability/verifiability.

The selected indicator is then added to the scorecard, and the previous steps are repeated until 10 indicators were selected. Finally, the Logit coefficients are transformed

---

[1] Higher "c" indicates greater ability to rank households by poverty status. For a Logit regression with a categorical outcome (such as poor/not poor), "c" is a general measure of explanatory power, much like R^2 in a least-squares regression on a continuous outcome (such as income). "c" is equal to the Mann-Whitney statistic (also known as the Wilcoxon rank-sum statistic) that indicates how much two distributions overlap (here, the distributions are of the estimated poverty likelihoods for poor and non-poor households). "c" is also equivalent to the area under an ROC curve—discussed in more detail later—that plots the share of poor and non-poor households versus all households ranked by score. Finally, "c" can also be seen as the share of all possible pairs of poor and non-poor households in which the poor household has a lower score. The more often the poor household has the lower score, the better the ranking by poverty status.

into non-negative integers such that the lowest possible score is 0 (most likely poor) and the highest is 100.

The statistical algorithm is the Logit analogue to the stepwise "MAXR" in, for example, Zeller, Alcaraz and Johannsen (2005) and IRIS (2005a and 2005b). The procedure here diverges from naïve stepwise in that expert judgment and non-statistical criteria were used to select from the most-predictive indicators. This improves robustness and, more importantly, helps ensure that the indicators are simple and sensible and so likely to be accepted by users.

4. Scorecard use

As explained in Schreiner (2005b), the main goal is not to maximize accuracy but to maximize the likelihood of programs' using scoring appropriately. When scoring projects fail, the culprit is usually not inaccuracy but rather the failure of users to accept scoring and to use it properly (Schreiner, 2002). The challenge is less technical and more human and organizational, not statistics but change management. Accuracy is easier—and less important—than practicality.

The scorecard here was designed to help users to understand and trust it (and thus use it properly). While accuracy matters, it must be balanced against simplicity, ease-of-use, and "face validity". In particular, programs are more likely to collect data, compute scores, and pay attention to the results if, in their view, scoring avoids creating "extra" work and if the whole process generally seems to make sense.

This “practicality” focus naturally leads to a one-page scorecard (Figures 4 and 5 in English and Spanish) that allows field workers to score households by hand in real time because it features:

- Only 10 indicators
- Only categorical indicators (“flooring material”, not “square meters of floor space”)
- User-friendly weights (non-negative integers, no arithmetic beyond simple addition)

Among other things, this simplicity enables “rapid targeting”, such as determining (in a day) who in a village qualifies for, say, work-for-food, or ration cards.

The scorecard in Figures 4 and 5 can be photocopied for immediate use. It can also serve as a template for data-entry screens with database software that records indicators, indicator values, scores, and poverty likelihoods.

A field agent collecting data and computing scores on paper would:

- Read each question off the scorecard
- Circle the response and the corresponding points
- Write the points in the far-right column
- Add up the points to get the total score
- Implement program policy based on the score

4.1 Scores and poverty likelihoods

A score is not a poverty likelihood (that is, the estimated probability of being poor), but each score is associated with a poverty likelihood via a simple table (Figure 6). For example, scores of 40–44 correspond to a poverty likelihood of 63.8 percent.

Scores (sums of scorecard weights) are associated with poverty likelihoods (estimated probabilities of being poor) via the "bootstrap" (Efron and Tibshirani, 1993):

- From the first one-fourth hold-out sample from the 2002 ENIGH, draw a new sample of the same size *with replacement*
- For each score range, compute the share of people with the score who are poor
- Repeat the previous two steps 10,000 times
- For a given score range, define the poverty likelihood as the average of the shares of people who are poor across the 10,000 samples

These resulting poverty likelihoods are objective, that is, based on data. This process would produce objective poverty likelihoods *even if the scorecards were constructed without data*. In fact, scorecards of objective, proven accuracy are often constructed *only* with qualitative judgment (Fuller, 2006; Schreiner *et al*., 2004). Of course, the scorecard here used data. The fact that the analyst used judgment in some choices during scorecard construction—as is inevitable in any statistical analysis—in no way impunes the objectivity of the poverty likelihoods, which depends on using data to associate scores with poverty likelihoods, not on whether only data was used to construct scorecards.

Figure 7 depicts the precision of estimated poverty likelihoods as point estimates with 90-, 95-, and 99-percent confidence intervals, derived from the bootstrap. Confidence intervals are a standard, widely understood way to measure accuracy.

For example, the average poverty rate across bootstrap samples (the poverty likelihood) for people with scores of 40–44 is 63.8 percent. In 90 percent of the 10,000 samples, this share is between 58.1–69.4 percent. In 95 percent of samples, the share is between 57.0–70.4; in 99 percent of samples, it is between 54.9–72.1.

For estimated and true poverty likelihoods, Figure 8 depicts mean absolute differences and confidence intervals from 10,000 bootstrap samples on the first one-fourth hold-out sample. Weighting by the people in a score range, the mean absolute difference is 8.2 percentage points.

This discussion so far looks at whether estimated poverty likelihoods are close to true poverty likelihoods. There is another aspect of accuracy: how well the poor are concentrated in low scores. A perfect scorecard would assign all the lowest scores to poor people (and all the highest scores to non-poor people). In reality, no scorecard is perfect, so some non-poor have low scores, and vice versa.

ROC curves are standard tools for showing how well scorecards concentrate the poor in lower scores (Baulch, 2003; Wodon, 1997). They plot the share of poor and non-poor households against the share of all households ranked by score.

What does an ROC curve mean? Suppose a program sets a cut-off so as to target the lowest-scoring x percent of potential participants. The ROC curve then shows the share of the poor and non-poor who would be targeted.

In Figure 9, the two northwest (southeast) curves depict accuracy among the poor (non-poor). As a benchmark, the external trapezoid shows the accuracy of a hypothetical perfect scorecard that assigns all of the lowest scores to poor people.

The inner lines represent the actual scorecard. They show, for example, that targeting the 30 percent of cases with the lowest scores would target 49 percent of all the poor and 9 percent of all the non-poor. Greater ability to rank-order—with less leakage and less undercoverage—is signified by curves that are closer to the northwest and southeast corners.

Figure 9 also reports two other common measures of ability to rank-order. The first is the Kolmogorov-Smirnov statistic, the maximum distance between the poor and non-poor curves (here 54.5). The second is the ratio of the area inside the ROC curves to the area inside the trapezoid of a hypothetical perfect scorecard (here 71.3).

Is this scorecard accurate enough for targeting? Errors due to scorecard inaccuracy are probably small relative to errors due to other sources (such as mistakes in data collection or fraud) and relative to the accuracy of other feasible targeting tools.

4.2 Estimates of overall poverty rates

The estimated overall poverty rate is the average of the estimated poverty likelihoods of individuals. For example, suppose a program had three participants on Jan. 1, 2006 who had scores of 20, 30, and 40, corresponding to poverty likelihoods of 89.2, 63.7, and 63.8 percent (Figure 6). The poverty rate is the participants' average poverty likelihood, that is, $(89.2 + 63.7 + 63.8) \div 3 = 72.2$ percent.

As a test, the scorecard was applied to 10,000 bootstrap replicates from the second one-fourth hold-out sample from the 2002 ENIGH, comparing the estimated overall poverty rates with the true values. The mean difference was 3.6 percentage points, with a standard deviation of 0.61. The 90-percent confidence interval around the mean was +/–1.0 percentage points, the 95-percent interval was +/–1.2 percentage points, and the 99-percent interval was +/–1.6 percentage points.

In practice, this means that subtracting 3.6 percentage points from a group's average poverty likelihood would produce an unbiased estimate that, in 99 of 100 cases, would be within +/–1.6 percentage points of the true overall poverty rate.

4.3 Progress out of poverty through time

For a given group, progress out of poverty over time is estimated as the change in average poverty likelihood. Continuing the previous example, suppose that on Jan. 1, 2007, the same three people (some of whom may no longer be participants) have scores of 25, 35, and 60 (poverty likelihoods of 85.5, 65.1, and 21.4 percent). Their average poverty likelihood is now 57.3 percent, an improvement of 72.2 – 57.3 = 14.9 percentage points in one year.

In a large group, this means that about 14.9 of every 100 progressed out of poverty. Given that 72.2 percent were poor in the first place, about one in five (14.9 ÷ 72.2 = 20.6 percent) of those who were poor left poverty.

Of course, this does not mean that program participation *caused* the progress; the scorecard just measures what happened, regardless of cause.

5. Setting targeting cut-offs

Potential participants with scores at or below a targeting cut-off are labeled *targeted* and treated—for program purposes—as if they were poor. Those with higher scores are *non-targeted* and treated—for program purposes—as if they were non-poor.

Poverty status (income below a poverty line) is distinct from *targeting status* (score below a cut-off). Poverty status is a fact whose determination requires an expensive survey. In contrast, targeting status is a policy choice whose determination requires a cut-off and an inexpensive estimate of poverty likelihood. Indeed, the purpose of scoring is to infer poverty status without incurring the cost of direct measurement.

No scorecard is perfect, so some of the truly poor will not be targeted, and some of the truly non-poor will be targeted. Targeting is accurate to the extent that poverty status matches targeting status. Accuracy in turn depends in part on the targeting cut-offs; some cut-offs are more accurate for the poor, others for the non-poor.

Setting a cut-off requires making this trade-off. The standard technique uses a *classification matrix* and a *net-benefit matrix* (Adams and Hand, 2000; Hoadley and Oliver, 1998; Greene, 1993).

5.1 Classification matrix

Given a targeting cut-off, there are four possible classification results:

A.	Truly poor	correctly	targeted	(score at or below the cut-off)
B.	Truly poor	mistakenly	non-targeted	(score above cut-off)
C.	Truly non-poor	mistakenly	targeted	(score at or below cut-off)
D.	Truly non-poor	correctly	non-targeted	(score above cut-off)

These four possibilities can be shown as a general classification matrix (Figure 10). Accuracy improves as there are more cases in A and D and fewer in B and C.

Figure 11 shows the number of people in each classification by score in the 2002 ENIGH. For example, with a cut-off of 40–44, there are:

A.	44.0	truly poor	correctly	targeted
B.	11.5	truly poor	mistakenly	non-targeted
C.	11.3	truly non-poor	mistakenly	targeted
D.	33.2	truly non-poor	correctly	non-targeted

Targeting accuracy (and errors of undercoverage and leakage) depends on the cut-off. For example, if the cut-off were increased to 45–49, more poor (but less non-poor) are correctly targeted:

A. 47.2	truly poor	correctly	targeted
B. 8.3	truly poor	mistakenly	non-targeted
C. 15.0	truly non-poor	mistakenly	targeted
D. 29.5	truly non-poor	correctly	non-targeted

Whether a cut-off of 40–44 is preferred to 45–49 depends on net benefit.

5.2 Net-benefit matrix

Each of the four classification results is associated with a net benefit (Figure 12):

α. Benefit	truly poor	correctly	targeted
β. Cost (negative net benefit)	truly poor	mistakenly	non-targeted
γ. Cost (negative net benefit)	truly non-poor	mistakenly	targeted
δ. Benefit	truly non-poor	correctly	non-targeted

Given a net-benefit matrix and a classification matrix, total net benefit is:

$$\text{Total net benefit} = \alpha \cdot \text{A} + \beta \cdot \text{B} + \gamma \cdot \text{C} + \delta \cdot \text{D}.$$

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

- Select a net-benefit matrix based on its values and mission
- Compute total net benefits for each cut-off with the net-benefit matrix and Figure 11
- Select the cut-off with the highest total net benefit

The only difficult step is selecting a net-benefit matrix. Some common net-benefit matrices are discussed below. In general, however, each program should thoughtfully decide for itself how much it values successful targeting versus errors of undercoverage and leakage. Of course, any program that targets already uses (if only implicitly) a net-benefit matrix. It is healthy to go through a process of thinking explicitly and intentionally about the value of the different targeting outcomes.

5.2.1 "Total Accuracy"

As an example net-benefit matrix, suppose a program selects the one that corresponds to the "Total Accuracy" criterion (Figure 13, IRIS, 2005b). Then total net benefit is the number of people correctly classified:

$$\begin{aligned}\text{Total net benefit} \quad &= 1\cdot\text{A} + 0\cdot\text{B} + 0\cdot\text{C} + 1\cdot\text{D},\\ &= \text{A} + \text{D}.\end{aligned}$$

Grootaert and Braithwaite (1998) and Zeller, Alcaraz, y Johannsen (2005) use "Total Accuracy" to evaluate the accuracy of poverty scorecards.

Figure 14 shows "Total Accuracy" for all cut-offs for the scorecard here. Total net benefit is greatest (77.2) for a cut-off of 40–44; at that point, poverty segment matches poverty status for about three of four people.

"Total Accuracy" weighs correct classifications of the poor and non-poor equally. But if most potential participants are non-poor and/or if a scorecard is more accurate for the non-poor, then "Total Accuracy" might be high even if very few poor people are correctly classified. Most programs targeting the poor, however, probably value correct classification more for the poor than the non-poor.

A simple, transparent way to reflect this valuation is to increase the relative net benefit of correctly classifying the poor. For example, if a program values correctly targeting the poor twice as much as correctly not targeting the non-poor, then α should be set twice as high as δ in the net-benefit matrix. Then the new optimal cut-off is 50–54, the cut-off point where 2.A + D is highest.

5.2.2 "Poverty Accuracy"

A criterion that emphasizes solely the importance of correctly classifying the poor is "Poverty Accuracy" (Figure 15, IRIS, 2005b):

$$\begin{aligned}\text{Total net benefit} \quad &= 1\cdot\text{A} + 0\cdot\text{B} + 0\cdot\text{C} + 0\cdot\text{D},\\ &= \text{A}.\end{aligned}$$

Of course, correctly targeting the poor is rarely the sole criteria. In fact, Figure 14 shows that "Poverty Accuracy" is greatest with a cut-off of 95–100. While targeting everyone does ensure that all poor people are targeted and so minimizes *undercoverage* of the poor (second-to-last column of Figure 14), it also targets all the non-poor and so maximizes *leakage* (the final column of Figure 14). A universal program may or may not be appropriate; the point here is to make explicit the implications of "Poverty Accuracy" as a criterion for choosing a targeting cut-off.

5.2.3 "Non-poverty Accuracy"

"Non-poverty Accuracy" counts only correct classifications of the non-poor (total net benefit is D). This is maximized by setting a cut-off of 0–4 and thus not targeting anyone (minimum leakage but maximum undercoverage).

5.2.4 "BPAC"

IRIS (2005b) proposes a new measure of accuracy called the "Balanced Poverty Accuracy Criterion". BPAC balances two goals:

- Accuracy of the estimated overall poverty rate
- "Poverty Accuracy"

According to IRIS (2005b), the first goal is optimized when undercoverage B is balanced by leakage C, and the second goal is optimized by maximizing A. If B > C, then Figure 16 is BPAC's net-benefit matrix. Thus, BPAC maximizes A while making B as close to C as possible:

$$\begin{aligned}\text{Total net benefit} \quad &= 1\cdot\text{A} + 1\cdot\text{B} + (-1)\cdot\text{C} + 0\cdot\text{D}, \\ &= \text{A} + (\text{B} - \text{C}).\end{aligned}$$

If C > B, then total net benefit under BPAC is A + (C – B).

BPAC was invented because IRIS does not estimate poverty likelihoods. Instead, IRIS estimates expenditure and then labels as poor those households with estimated expenditure less than the poverty line. In this set-up, the estimated overall poverty rate is the share of people targeted, and this estimate is most accurate (that is, it matches the true value) when undercoverage B equals leakage C.

For a scorecard (like the one here) that estimates poverty likelihoods, however, BPAC is not meaningful. This is because the estimated overall poverty rate is the average of participants' estimated poverty likelihoods. These estimates are independent of whatever targeting cut-off a program might set. In contrast, the targeting errors of undercoverage B and leakage C depend directly on the cut-off chosen. Thus, for scorecards that estimate poverty likelihoods, getting B close to C is not related to optimizing the accuracy of the estimated overall poverty rate and so is not related to the goals of BPAC.

6. Training, quality-control, and MIS

The technical aspects of scorecard construction and accuracy just discussed are important, but gaining the trust and acceptance of managers and field workers is even more important (Schreiner, 2002).

In particular, the field workers who collect indicators must be trained. If they put garbage in, the scorecard will put garbage out. To prevent abuse, on-going quality control of data is required.

Programs should record in their MIS at least the poverty likelihood along with an identifier for each client. Ideally, they would also record the score, the indicators, and the values of the indicators. This will allow quick computation of average poverty likelihoods (as well as other analyses), both for a point in time and for changes through time (Matul and Kline, 2003).

7. Calibrating the scorecard for the very poor

The simple poverty scorecard in Figures 4 and 5 can be used to track outreach not only to the poor but also to the *very poor*, that is, the poorest half of the poor.

7.1 Poverty likelihoods

As before, scores are associated with the probability of being very poor by bootstrapping 10,000 samples from the first one-fourth hold-out sample. The poverty likelihood for a given score is then taken as the average of the shares of people with that score who are very poor across the 10,000 samples.

Columns 2–4 in Figure 17 are the poverty likelihoods for the three classes for all scores. For example, if a potential participant has a score of 25–29, the probability of being very poor is 54.1 percent, the probability of being poor is 31.5 percent, and the probability of being non-poor is 14.5 percent.

Columns 5–7 in Figure 17 are the share of targeted participants by poverty status and by cut-off. For example, for a cut-off of 35–39, 47.0 percent of those targeted would be very poor, 34.9 percent would be poor, and 18.0 percent would be non-poor.

Each person is associated with three poverty likelihoods. For example, a person with a score of 25 may be targeted as very poor, but the likelihood of truly being very poor is not 100 percent but rather 54.1 percent (from Figure 17). The same person has a 31.5-percent likelihood of being truly poor, and a 14.5-percent likelihood of being truly non-poor. Each person has one targeting status (for program purposes), one true poverty status (in reality), and three estimated poverty likelihoods (one for each possible poverty status).

As before, these poverty likelihoods are objective, that is, based on data. They are valid even though the scorecard was not constructed originally to predict the likelihood of being very poor. It works because the likelihood of being very poor is highly correlated with having a low score (high likelihood of being poor). A scorecard could be built specifically for the very poor, but it would add cost and complexity.

Figure 18 shows the precision of estimated poverty likelihoods for being very poor as point estimates with 90-, 95-, and 99-percent confidence intervals. For example, the average poverty rate (the poverty likelihood) across bootstrap samples for people with scores of 40–44 was 22.5 percent. In 90 percent of 10,000 samples from the first one-fourth hold-out, the share was between 17.6–26.6 percent. In 95 percent of samples, it was between 16.8–28.7, and in 99 percent of samples, it was between 14.8–30.8.

For estimated and true poverty likelihoods, Figure 19 depicts mean absolute differences and confidence intervals from 10,000 bootstrap samples on the second one-fourth hold-out sample. Weighting by the people in a score range, the mean absolute difference is 9.3 percentage points.

The other aspect of accuracy is how well the very poor are concentrated in low scores. Once again, an ROC curve is a useful way to look at this.

Figure 20 plots the share of the very poor against the share of the not very poor, ranked by score. For example, targeting the 25 percent of cases with the lowest scores would target 55 percent of all the very poor and 14 percent of all the not very poor.

In terms of the Kolmogorov-Smirnov statistic, the maximum distance between the curves is 53.8. In terms of the ratio of the area inside the scorecard curves to the area inside the trapezoid of a hypothetical perfect scorecard, the value is 65.7.

All in all, Figures 18–20 suggest that the estimated likelihoods of being very poor are estimated both accurately and precisely.

7.2 Overall poverty rates for the very poor

The average of estimated poverty likelihoods for a group is their estimated overall (very poor) poverty rate. To measure the accuracy and precision of this estimate, the scorecard was applied to 10,000 bootstrap replicates from the second one-fourth hold-out sample, and then the estimated overall poverty rates were compared with the true values. The mean difference was 1.4 percentage points, with a standard deviation of 0.58. The 90-percent confidence interval around the mean was +/–1.0 percentage points, the 95-percent interval was +/–1.1 percentage points, and the 99-percent interval was +/–1.5 percentage points.

Thus, subtracting 1.4 percentage points to a group's average poverty likelihood would produce an unbiased estimate that, in 99 of 100 cases, would be within +/–1.5

percentage points of the true overall (very poor) poverty rate. This estimate is both accurate and precise.

7.3 Targeting the very poor

As before, targeting involves using a classification matrix and a net-benefit matrix to select a cut-off. The wrinkle is that there are now three poverty statuses:

- Very poor: Poorest half of those with expenditure at or below the poverty line
- Poor: Least-poor half of those with expenditure at or below poverty
- Non-poor: Expenditure above poverty

There are also three targeting segments:

- Very poor: Score at or below the very poor/poor cut-off
- Poor: Score above the very poor/poor cut-off and at or below the poor/non-poor cut-off
- Non-poor: Score above the poor/non-poor cut-off

There are two cut-offs (very poor/poor and poor/non-poor) and 9 classification results (Figure 21):

A. Truly very poor correctly classified as very poor

B. Truly very poor incorrectly classified as poor

C. Truly very poor incorrectly classified as non-poor

D. Truly poor incorrectly classified as very poor

E. Truly poor correctly classified as poor

F.	Truly poor	incorrectly classified as	non-poor
G.	Truly non-poor	incorrectly classified as	very poor
H.	Truly non-poor	incorrectly classified as	poor
I.	Truly non-poor	correctly classified as	non-poor

The general classification matrix (Figure 21) and the net-benefit matrix (Figure 22) are combined to define total net benefit:

Total net benefit = $\alpha \cdot A + \beta \cdot B + \gamma \cdot C + \delta \cdot D + \varepsilon \cdot E + \zeta \cdot F + \eta \cdot G + \theta \cdot H + \iota \cdot I$.

Figure 23 shows classification results for all possible pairs of cut-off scores in the second one-fourth hold-out sample. For example, suppose a program defined:

- Very poor/poor cut-off of 35–39 (so scores of 0–39 are targeted as very poor)
- Poor/non-poor cut-off of 50–54 (so scores of 40–54 are targeted as poor, and scores of 55–100 are targeted as non-poor)

As with any scorecard and cut-offs, there is both successful targeting and errors. For the example cut-offs of 35–39 and 50–54, targeting would be correct for 81 percent of the very poor, 26 percent of the poor, and 59 percent of the non-poor (Figure 24).

The program chooses a set of cut-offs to optimize the benefits of correct classifications, net of the costs (negative benefits) of incorrect classifications. For example, suppose the net-benefit matrix is Figure 25, representing one way to reflect:

- Greater importance of correctly targeting the very poor and poor
- Greater cost of gross errors such as targeting the truly very poor as non-poor

Given the classification results in Figure 24 and net benefits in Figure 25, total net benefit for the cut-off pair of 35–39 and 50–55 is +979 (Figure 26).

Is this the best pair of cut-offs? The answer requires applying the net-benefit matrix to the classification results for all 190 possible pairs (Figure 23). It turns out that total net benefit is indeed highest for cut-offs of 35–39 and 50–55.

8. Conclusion

Mexico has more than 50 million poor people. An easy-to-use, inexpensive tool for identifying the poor could improve targeting and speed progress out of poverty. This paper presents a simple scorecard that estimates the likelihood that a person has income below the official poverty line.

The scorecard is built and tested using data on 17,167 households from the 2002 ENIGH. The scorecard is calibrated to estimate the likelihood of being poor (income below the national poverty line) or very poor (poorest half of the poor).

Out-of-sample bootstrap tests show that the estimates are both accurate and precise. For a group's overall poverty rate (again, whether poor or very poor), estimates are within +/–1.6 percentage points of the true value with 99-percent confidence and within +/–1.0 percentage points with 90-percent confidence.

For targeting, programs can use the classification results reported here to select the best choice of cut-off according to their values and mission.

Accuracy is important, but ease-of-use is even more important; a perfectly accurate scorecard is worthless if programs feel daunted by its complexity and so never even try to use it. For this reason, the scorecard here is kept simple, using 10 indicators that are inexpensive to collect and that are straightforward to observe and verify. Indicator weights are all zeros or positive integers, and scores range from 0 (most likely poor) to 100 (least likely poor). Scores are related to poverty likelihoods via a simple look-up table, and targeting cut-offs are also simple to apply. Thus, field workers not only can understand the scorecard, but they can also use it to compute scores in the field, by hand, in real time.

Overall, the poverty scorecard can help Mexican development programs to target services to the poor, track participants' progress out of poverty through time, and report on participants' overall poverty rate.

References

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

Baulch, Bob. (2003) "Poverty Monitoring and Targeting Using ROC Curves: Examples from Vietnam", IDS Working Paper No. 161, http://www.ids.ac.uk/ids/bookshop/wp/wp161.pdf.

Comité Técnico para la Medición de la Pobreza. (2002) "Medición de la Pobreza: Variantes Metodológicas y Estimación Preliminar", Secretaria de Desarrollo Social, México, D.F., ISBN 968–838–476–3, http://www.oportunidades.gob.mx/cd/docsconsulta/comtecmedpobr.pdf.

Efron, Bradley; and Robert J. Tibshirani. (1993) *An Introduction to the Bootstrap*, New York: Chapman and Hall, ISBN 0–412–04231–2.

Fuller, Rob. (2006) "Measuring Poverty of Microfinance Clients in Haiti", http://www.microfinance.com/English/Papers/Scoring_Poverty_Haiti_Fuller.pdf.

Goodman, L.A. and Kruskal, W.H. (1979) *Measures of Association for Cross Classification*, New York, NY: Springer-Verlag, ISBN 0–38–790443–3.

Greene, William H. (1993) *Econometric Analysis: Second Edition*, New York, NY: MacMillan, ISBN 0–02–346391–0.

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, Washington, D.C., http://www.worldbank.org/html/dec/Publications/Workpapers/WPS1900series/wps1942/wps1942.pdf.

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. (2005a) "Accuracy Results for 12 Poverty Assessment Tool Countries", http://www.povertytools.org/documents/Accuracy%20Results%20for%2012%20Countries.pdf.

IRIS Center. (2005b) "Notes on Assessment and Improvement of Tool Accuracy", http://www.povertytools.org/documents/Assessing%20and%20Improving%20Accuracy.pdf.

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

Rascón Ramírez, Erica Gabriela. (2002a) "Nota Técnica para la medición de la pobreza con base en los resultados de la Encuesta Nacional de Ingresos y Gastos de los Hogares, 2002", Secretaria de Desarrollo Social, México, D.F., http://www.sedesol.gob.mx/subsecretarias/prospectiva/medicion_pobreza/Nota_tecnica_pobreza_2002.pdf.

Schreiner, Mark (2005a) "Un Indice de Pobreza para México", memo for Grameen Foundation U.S.A., http://www.microfinance.com/Castellano/Documentos/Scoring_Pobreza_Mexico.pdf.

Schreiner, Mark. (2005b) "IRIS questions on poverty scorecards", memo for Grameen Foundation U.S.A., http://www.microfinance.com/English/Papers/Scoring_Poverty_Response_to_IRIS.pdf.

Schreiner, Mark. (2002) *Scoring: The Next Breakthrough in Microfinance?* Occasional Paper No. 7, Consultative Group to Assist the Poorest, Washington, D.C., http://www.cgap.org/docs/OccasionalPaper_07.pdf.

Schreiner, Mark; Matul, Michal; Pawlak, Ewa; and Sean Kline. (2004) "Poverty Scorecards: Lessons from a Microlender in Bosnia-Herzegovina", Microfinance Risk Management, http://www.microfinance.com/English/Papers/Scoring_Poverty_in_BiH_Short.pdf.

Sillers, Don. (2006) "National and International Poverty Lines: An Overview", Washington, D.C.: United States Agency for International Development, http://povertytools.org/Project_Documents/Poverty_lines___An_Overview_1_4_06.pdf

Singh, Kesar. (1998) "Breakdown Theory for Bootstrap Quantiles", *Annals of Statistics*, Vol. 26, pp. 1719–1732.

Wodon, Quentin T. (1997) "Targeting the Poor Using ROC Curves", *World Development*, Vol. 25, No. 12, pp. 2083–2092.

Zeller, Manfred. (2004) "Review of Poverty Assessment Tools", Accelerated Microenterprise Advancement Project, http://www.povertytools.org/documents/Review%20of%20Poverty%20Assessment%20Tools.pdf.

Zeller, Manfred; Alcaraz V., Gabriela; and Julia Johannsen. (2005) "Developing and Testing Poverty-Assessment Tools: Results from Accuracy Tests in Peru", Accelerated Microenterprise Advancement Project, http://www.povertytools.org/documents/Peru%20Accuracy%20Report.pdf.

Figure 1: Households surveyed, people represented, and overall poverty rates

Sub-sample	Households	People	% poor
Constructing scorecards	8,528	50,153,473	51.1
Associating scores with likelihoods	4,354	25,278,733	51.3
Testing accuracy	4,285	26,090,208	53.2
Source: 2002 ENIGH.	17,167	101,522,414	51.7

Figure 2: Poverty indicators by uncertainty coefficient

Uncertainty coefficient		Indicator (Answers ordered starting with those most strongly associated with poverty)
1.	72.1	Does the house have a bathroom? (No; Yes)
2.	72.0	Does the house have a shower? (No; Yes)
3.	72.0	Is the toilet supplied by piped water? (No; Yes)
4.	71.9	Does the house have a sink? (No; Yes)
5.	71.7	Does the house have piped water? (No piped water in the house; Piped water in the building or yard; Piped water inside the house)
6.	71.6	What are the floors of the house primarily made out of? (Earth; Cement; Other)
7.	71.4	Does the household own a land-line telephone or a cellular telephone? (No; Yes)
8.	71.3	Does the household own a land-line telephone and a cellular telephone? (No; Yes)
9.	71.1	How many people are there per room, not counting kitchen, bathroom, and hallways? (4 or more; 3 to 3.99; 2 to 2.99; 1 to 1.99; 0 to 0.99)
10.	69.5	How many rooms are in the house, not counting kitchen, bathroom, and hallways? (1; 2; 3 or more)
11.	68.9	How many household members are aged 0 to 11 years? (3 or more; 2; 0 or 1)
12.	68.6	Does the household own a water heater? (No; Yes)
13.	68.1	In the last six months, did the household buy anything in a supermarket or department store? (No; Yes)
14.	67.4	Does the house have a closet? (No; Yes)
15.	67.2	Does the household own an automobile, pick-up, mini-van, cargo truck, etc.? (No; Yes)
16.	66.6	Does the household own a dishwasher? (No; Yes)
17.	66.6	Does the household own a refrigerator? (No; Yes)
18.	66.1	In the past month, did the household buy any fabric softener? (No; Yes)
19.	65.9	Does the household own a microwave oven? (No; Yes)
20.	65.2	Does the household own a VCR? (No; Yes)
21.	64.8	Does the house's roof have some sort of external covering? (No; Yes)
22.	64.7	What type of fuel does the household use for cooking? (Wood; Other)
23.	64.3	Does the household own a food mixer? (No; Yes)
24.	64.2	Does the household own a color television? (No; Yes)
25.	63.9	Does the household own a modular/component stereo system? (No; Yes)

Source: Calulations based on the 2002 ENIGH.

Figure 2 (cont.): Poverty indicators by uncertainty coefficient

Uncertainty coefficient		Indicator (Answers ordered starting with those most strongly associated with poverty)
26.	63.9	Does the household own a coffee pot? (No; Yes)
27.	63.9	Does the household own a juice extractor? (No; Yes)
28.	63.5	What was the highest grade passed by the female head/spouse? (Up to sixth grade of primary; From sixth grade to the third year of vocational prepatory; Seventh semester of Basic Normal or more)
29.	63.4	Do all children ages 6 to 14 go to school? (No; Yes)
30.	62.8	Do all children ages 6 to 11 go to school? (No; Yes)
31.	62.3	In the past week, did the household consume dried beans of any kind? (No; Yes)
32.	61.9	Does the household own a blender? (No; Yes)
33.	61.4	Does the household own a fan? (No; Yes)
34.	61.3	How many household members have blue-collar wage jobs? (0; 1; 2 or more)
35.	61.3	In the past three months, did the household buy any clothes for a member aged 17 or older? (No; Yes)
36.	61.1	In the past week, did the household consume milk in any form? (No; Yes)
37.	61.0	Does the household own a toaster? (No; Yes)
38.	61.0	Do all girls ages 6 to 17 go to school? (No; Yes)
39.	61.0	Does the household own a clothes iron? (No; Yes)
40.	59.2	Does the household own a computer, printer, or any other equipment connected to a computer? (No; Yes)
41.	59.0	In the past month, did anyone in the household pay for a haircut? (No; Yes)
42.	59.0	Do all girls ages 6 to 11 go to school? (No; Yes)
43.	58.9	In the past week, did the household consume meat of any kind?
44.	58.9	Does the household own a television of any kind? (No; Yes)
45.	58.8	Does the household have cable or satellite television? (No; Yes)
46.	58.8	In the past three months, did the household buy shoes for a member aged 17 or older? (No; Yes)
47.	58.7	In the past week, did the household consume any refined sugar, whether white or brown? (No; Yes)
48.	58.7	In the past month, did anyone in the household go to the movies, the theatre, or a concert? (No; Yes)
49.	58.0	Does any child go to a private school? (No; Yes)
50.	57.2	In the past week, did any household member eat a meal outside the home? (No; Yes)

Source: Calulations based on the 2002 ENIGH.

Figura 3: Indicadores de pobreza según coeficiente de inciertadumbre

“c”, la coeficiente de inciertadumbre		Indicador (Respuestas ordenados comenzando con la cual está más estrechamente vinculado con la pobreza)
1.	72.1	¿Cuenta con un lavabo? (No; Sí)
2.	72.0	¿Cuenta con una regadera? (No; Sí)
3.	72.0	¿Tiene conexión de agua para el sanitario? (No; Sí)
4.	71.9	¿Cuenta con un fregadero? (No; Sí)
5.	71.7	¿El agua está dentro de la vivienda? (No hay en la vivienda; Hay en el edificio o terreno; Hay dentro de la vivienda)
6.	71.6	¿De qué material es la mayor parte de los pisos de la vivienda? (Tierra; Cemento o firme; Otro)
7.	71.4	¿Tiene teléfono fijo o teléfono celular? (No; Sí)
8.	71.3	¿Tiene teléfono fijo y teléfono celular? (No; Sí)
9.	71.1	¿Cuántas personas hay por cuarto, sin contar cocina, baño y pasillos? (4 o más; 3 a <4; 2 a <3; 1 a <2; 0 a <1)
10.	69.5	¿Cuántos cuartos hay en la vivienda sin contar cocina, baño y pasillos? (1; 2; 3 o más)
11.	68.9	¿Cuántos miembros del hogar son de edades de 0 a 11 años? (3 o más; 2; 1 o 0)
12.	68.6	¿Cuenta con una calentador de gas? (No; Sí)
13.	68.1	En los últimos seis meses, ¿Compró el hogar algo en un supermercado o en una tienda departamental? (No; Sí)
14.	67.4	¿Cuenta con un closet? (No; Sí)
15.	67.2	¿Cuenta con automóvil, camioneta, combi, camioneta de caja, etc.? (No; Sí)
16.	66.6	¿Cuenta con una lavadora? (No; Sí)
17.	66.6	¿Cuenta con un refrigerador? (No; Sí)
18.	66.1	En el último mes, ¿Compró el hogar suavizantes de telas? (No; Sí)
19.	65.9	¿Cuenta con un horno de microondas? (No; Sí)
20.	65.2	¿Cuenta con una video casetera? (No; Sí)
21.	64.8	¿Los techos de la vivienda tiene algún recubrimiento por la parte externa? (No; Sí)
22.	64.7	¿Qué combustible utiliza para cocinar? (Leña; gas u otro)
23.	64.3	¿Cuenta con una batidora? (No; Sí)
24.	64.2	¿Cuenta con una televisión a color? (No; Sí)
25.	63.9	¿Cuenta con un estéreo modular o consolar? (No; Sí)

Fuente: Cálculos basados en el 2002 ENIGH.

Figura 3 (cont.): Indicadores de pobreza según coeficiente de inciertadumbre

"c", la coeficiente de inciertadumbre		Indicador (Respuestas ordenados comenzando con la cual está más estrechamente vinculado con la pobreza)
26.	63.9	¿Cuenta con una cafetera? (No; Sí)
27.	63.9	¿Cuenta con un exprimidor o extractor de jugos? (No; Sí)
28.	63.5	¿Ultimo grado de estudios que terminó la jefa/esposa? (Hasta 6to. primaria; Hasta 6to. semestre o 3ro. año preparatoria vocacional; 7tmo. semestre de la Normal Básica o más)
29.	63.4	¿Todos los hijos/hijas de edades 6 a 17 estudian? (No; Sí)
30.	62.8	¿Todos los hijos/hijas de edades 6 a 11 estudian? (No; Sí)
31.	62.3	En la semana pasada, ¿Consumió en el hogar frijol, bayo, flor de mayo, negro etc. en cualquier forma? (No; Sí)
32.	61.9	¿Cuenta con una licuadora? (No; Sí)
33.	61.4	¿Cuenta con un ventilador? (No; Sí)
34.	61.3	¿Número de miembros del hogar que son "obreros o empleados"? (0; 1; 2 o más)
35.	61.3	En los últimos tres meses, ¿Compró el hogar prendas de vestir para una persona de 17 años o más? (No; Sí)
36.	61.1	En la semana pasada, ¿Consumió en el hogar leche de cualquier tipo? (No; Sí)
37.	61.0	¿Cuenta con una tostadora? (No; Sí)
38.	61.0	¿Todas las hijas de edades 6 a 17 estudian? (No; Sí)
39.	61.0	¿Cuenta con una plancha para ropa? (No; Sí)
40.	59.2	¿Cuenta con una computadora, impresora, y/o otras aparatos integrados a la computadora? (No; Sí)
41.	59.0	En el último mes, ¿Compró alguien del hogar un servicio de corte de cabello? (No; Sí)
42.	59.0	¿Todas las hijas de edades 6 a 11 estudian? (No; Sí)
43.	58.9	En la semana pasada, ¿Consumió en el hogar carne de res o ternera en cualquier forma?
44.	58.9	¿Cuenta con una televisión, sea a color o de blanco y negro? (No; Sí)
45.	58.8	¿Cuenta con televisión por cable, Sky o Direc-TV o Multivisión? (No; Sí)
46.	58.8	En los últimos tres meses, ¿Compró el hogar calzado para una persona de 17 años o más? (No; Sí)
47.	58.7	En la semana pasada, ¿Consumió en el hogar azúcar blanca o morena? (No; Sí)
48.	58.7	En el último mes, ¿Asistió alguien del hogar a un cine, teatro o concierto? (No; Sí)
49.	58.0	¿Asiste algún niño a una escuela privada? (No; Sí)
50.	57.2	En la semana pasada, ¿Comió alguien del hogar en una lonchería, fonda, tortería, taquería, cocina económica o cenaduría? (No; Sí)

Fuente: Cálculos basados en el 2002 ENIGH.

Figure 4: A simple poverty scorecard for Mexico

Indicator						Points
1.	Does the house have a bathroom?			No 0	Yes 7	
2.	How many children ages 6 to 17 attend school?	Three or more, or not all go 0	Two, and both go 6	One, and he/she goes 11	No children this age 19	
3.	Does the household own a land-line telephone or a cellular telephone?			No 0	Yes 10	
4.	What are the floors of the house primarily made out of?		Earth 0	Cement 3	Other 7	
5.	Does the house have a closet?			No 0	Yes 9	
6.	What type of fuel does the household use for cooking?			Wood 0	Other 10	
7.	Do the household own a coffee pot?			No 0	Yes 9	
8.	Does any child go to a private school			No 0	Yes 15	
9.	Does the household own a modular/component stereo system?			No 0	Yes 6	
10.	How many rooms does the house have, not counting kitchen, bathroom, or hallways?		One 0	Two 4	Three or more 7	
					Total:	

Source: Calculations based on the 2002 ENIGH.

Figura 5: Una indíce sencillo de pobreza en México

Indicador					Puntaje
1. ¿Esta vivienda cuenta con lavabo?			No 0	Sí 7	
2. ¿Cuántos niños de 6 a 17 años de edad asisten a la escuela?	Tres o más, o no todos asisten 0	Dos, y ambos asisten 6	Uno, y asiste 11	No hay niños 19	
3. ¿Cuenta con teléfono fijo o teléfono celular?			No 0	Sí 10	
4. ¿De qué material es la mayor parte de los pisos de la vivienda?		Tierra 0	Cemento o firme 3	Otro 7	
5. ¿Cuenta la vivienda con un closet?			No 0	Sí 9	
6. ¿Qué combustible utiliza para cocinar?			Leña 0	Otro 10	
7. ¿Cuenta con una cafetera?			No 0	Sí 9	
8. ¿Asiste algún niño a una escuela privada?			No 0	Sí 15	
9. ¿Cuenta con un estéreo modular o consolar?			No 0	Sí 6	
10. ¿Cuántos cuartos tiene la vivienda sin contar cocina, baño y pasillos?		Uno 0	Dos 4	Tres o más 7	
Fuente: Cálculos basado en el 2002 ENIGH.				Total:	

Figure 6: Scores and poverty likelihoods

Score	Poverty likelihood for people with score in range (%)	% of people <=score who are poor	% of people >score who are non-poor
0-4	99.6	99.6	45.9
5-9	98.2	98.9	47.2
10-14	97.1	98.1	49.7
15-19	85.5	93.5	52.3
20-24	89.2	92.2	56.4
25-29	85.5	90.6	61.2
30-34	63.7	85.6	64.2
35-39	65.1	82.0	69.0
40-44	63.8	79.6	74.3
45-49	46.0	75.8	78.0
50-54	48.9	73.5	83.1
55-59	32.5	69.4	87.9
60-64	21.4	66.6	90.2
65-69	19.3	63.6	93.9
70-74	2.6	61.3	92.9
75-79	12.1	58.2	99.0
80-84	2.0	57.6	99.3
85-89	0.8	55.7	100.0
90-94	0.0	55.5	0.0
95-100	0.0	55.5	0.0

Surveyed cases weighted to represent the total population.

Source: Based on the 2002 ENIGH.

Figure 7: Confidence intervals for estimated poverty likelihoods

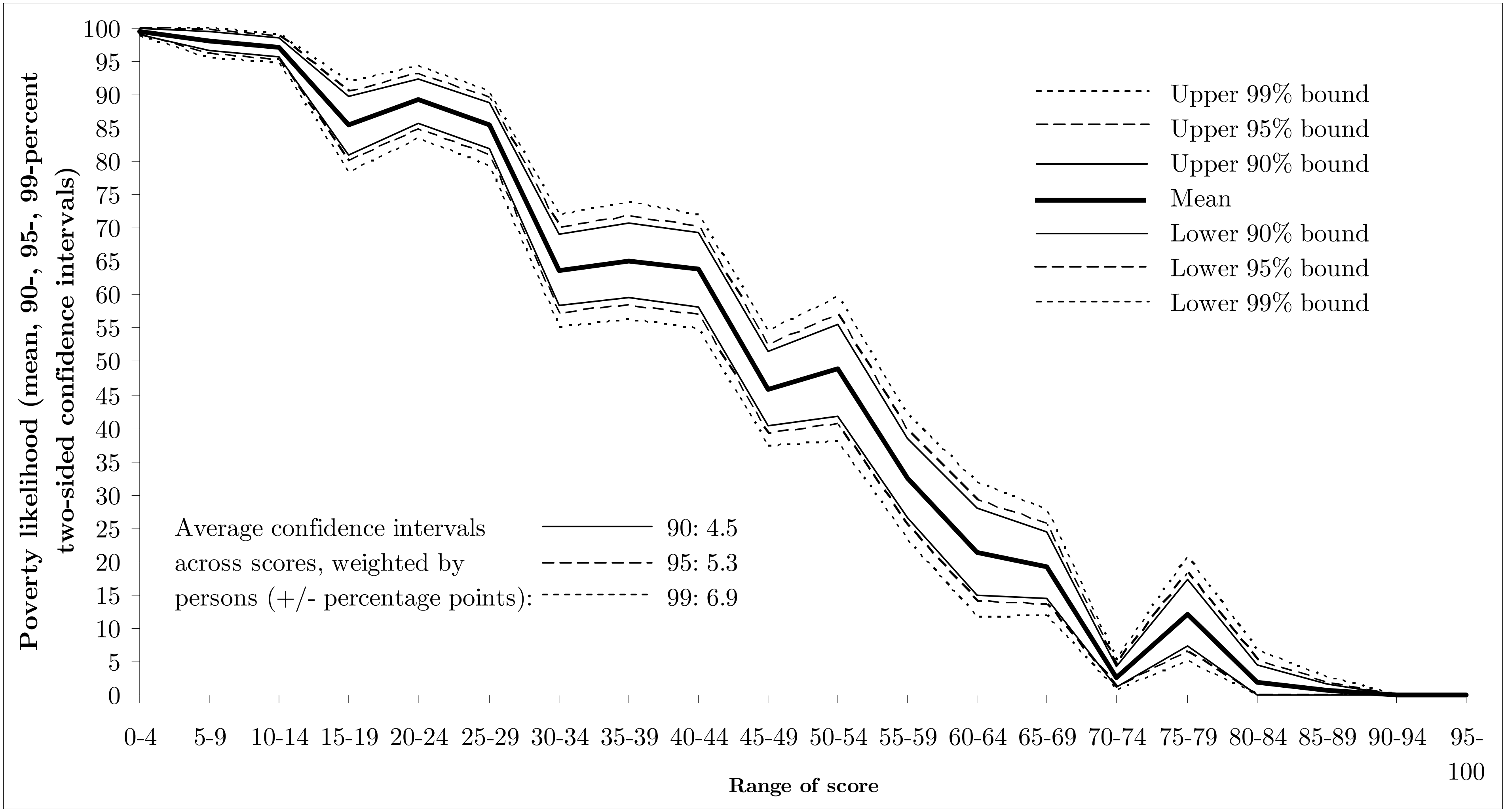

Figure 8: Differences between estimated and true poverty likelihoods

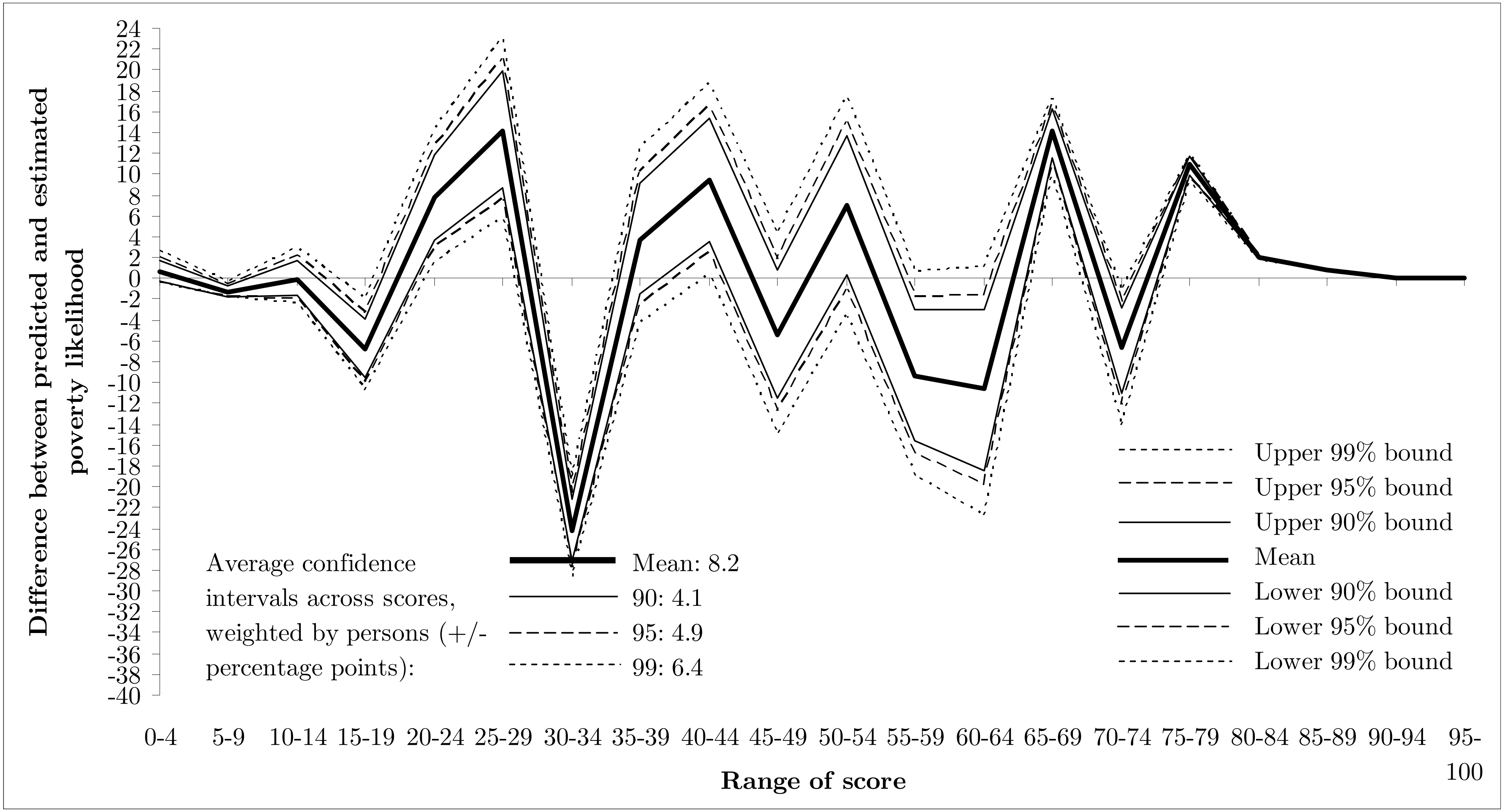

Figure 9: ROC curve of ability to rank-order households by poverty status

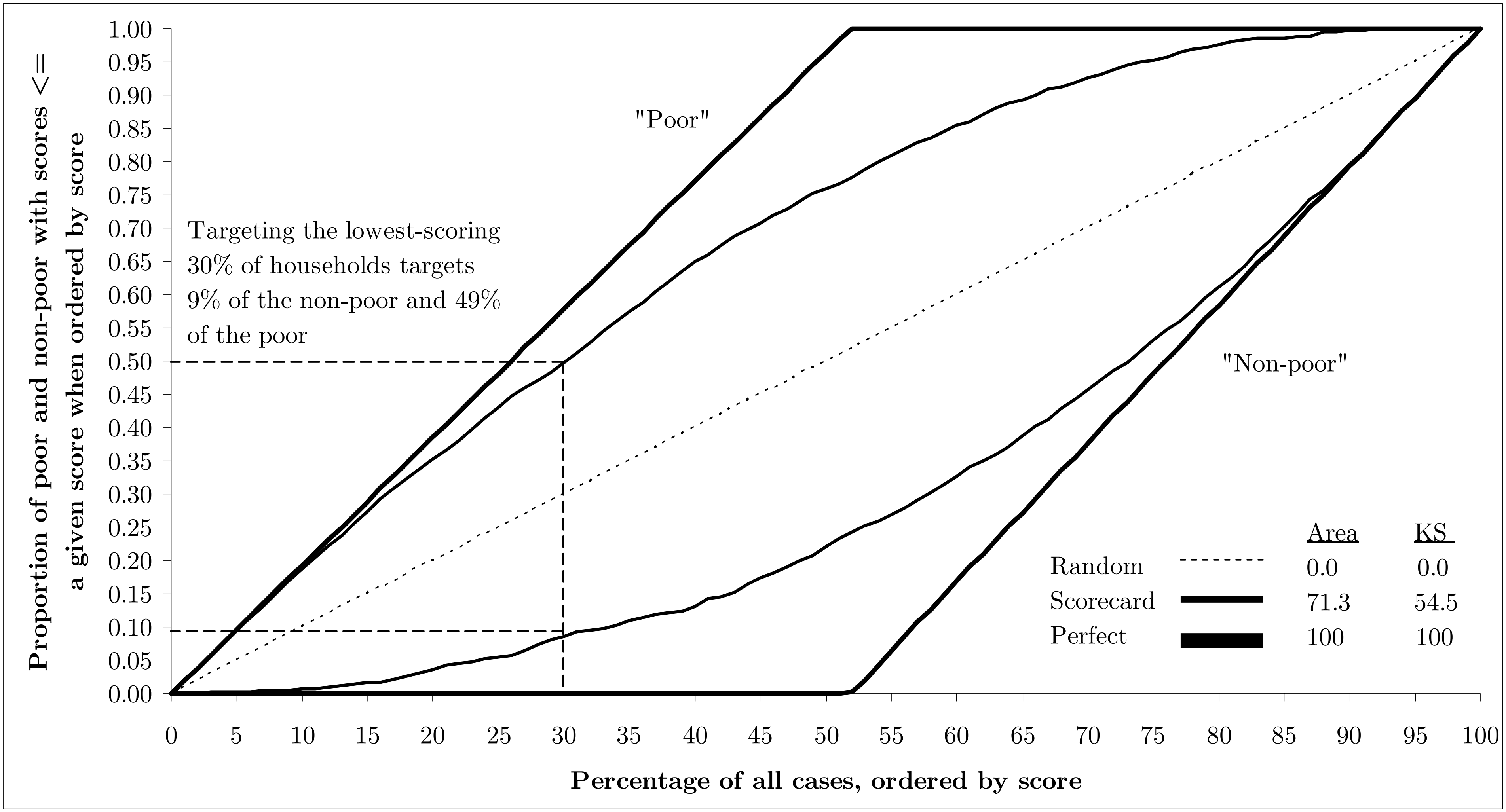

Figure 10: General classification matrix

		Targeting segment	
		Targeted	**Non-targeted**
True poverty status	**Poor**	**A.** Truly poor correctly targeted	**B.** Truly poor mistakenly non-targeted
	Non-poor	**C.** Truly non-poor mistakenly targeted	**D.** Truly non-poor correctly non-targeted

Figure 11: People by targeting classification and score

Score	A. Truly poor correctly targeted	B. Truly poor mistakenly non-targeted	C. Truly non-poor mistakenly targeted	D. Truly non-poor correctly non-targeted
0-4	3.1	52.4	0.0	44.5
5-9	5.8	49.6	0.1	44.4
10-14	10.6	44.9	0.2	44.3
15-19	15.9	39.6	1.1	43.4
20-24	22.5	32.9	1.9	42.6
25-29	29.1	26.4	3.0	41.5
30-34	33.8	21.7	5.7	38.8
35-39	39.4	16.1	8.7	35.8
40-44	44.0	11.5	11.3	33.2
45-49	47.2	8.3	15.0	29.5
50-54	50.1	5.4	18.1	26.4
55-59	52.6	2.9	23.2	21.3
60-64	53.6	1.9	26.9	17.6
65-69	54.6	0.8	31.3	13.2
70-74	54.7	0.8	34.6	9.9
75-79	55.4	0.0	39.8	4.7
80-84	55.5	0.0	40.9	3.6
85-89	55.5	0.0	44.1	0.4
90-94	55.5	0.0	44.5	0.0
95-100	55.5	0.0	44.5	0.0

Figures normalized to sum to 100.

Source: Based on the 2002 ENIGH.

Figure 12: General net-benefit matrix

		Targeting segment	
		Targeted	Non-targeted
True poverty status	Poor	α	β
	Non-poor	γ	δ

Figure 13: "Total Accuracy" net-benefit matrix

True poverty status	Targeting segment: Targeted	Targeting segment: Non-targeted
Poor	1	0
Non-poor	0	1

Figure 14: Total net benefit for some common net-benefit matrices

	Total Accuracy (A + B)	Poverty Accuracy 100*A / (A+B)	Non-poverty Accuracy 100*D / (C+D)	Undercoverage 100*B / (A+B)	Leakage 100*C / (A+C)
	1 0	1 0	0 0	0 -1	0 0
Score	0 1	0 0	0 1	0 0	-1 0
0-4	47.6	5.7	100.0	94.3	0.4
5-9	50.3	10.5	99.9	89.5	1.1
10-14	54.9	19.1	99.5	80.9	1.9
15-19	59.3	28.6	97.5	71.4	6.5
20-24	65.1	40.6	95.7	59.4	7.8
25-29	70.6	52.5	93.2	47.5	9.4
30-34	72.6	60.9	87.2	39.1	14.4
35-39	75.2	70.9	80.5	29.1	18.0
40-44	77.2	79.2	74.7	20.8	20.4
45-49	76.6	85.0	66.2	15.0	24.2
50-54	76.5	90.3	59.3	9.7	26.5
55-59	73.9	94.7	47.9	5.3	30.6
60-64	71.2	96.6	39.5	3.4	33.4
65-69	67.8	98.5	29.6	1.5	36.4
70-74	64.6	98.6	22.2	1.4	38.7
75-79	60.1	99.9	10.5	0.1	41.8
80-84	59.1	100.0	8.2	0.0	42.4
85-89	55.9	100.0	0.9	0.0	44.3
90-94	55.5	100.0	0.0	0.0	44.5
95-100	55.5	100.0	0.0	0.0	44.5

All figures in percentage units.

Figure 15: "Poverty Accuracy" net-benefit matrix

True poverty status	Targeting segment: Targeted	Targeting segment: Non-targeted
Poor	1	0
Non-poor	0	0

Figure 16: Net-benefit matrix for BPAC

True poverty status		Targeting segment: Targeted	Targeting segment: Non-targeted
	Poor	1	1
	Non-poor	-1	0

Figure 17: Poverty likelihoods for the very poor, poor, and non-poor by score

Score	Poverty likelihood in score range			Share of cases <= score		
	Very Poor	Poor	Non-poor	Very Poor	Poor	Non-poor
0-4	78.0	21.6	0.4	78.0	21.6	0.4
5-9	76.4	21.8	1.8	77.2	21.7	1.1
10-14	61.4	35.8	2.9	70.0	28.1	1.9
15-19	45.8	39.7	14.5	61.2	32.3	6.5
20-24	45.8	43.4	10.8	56.5	35.7	7.8
25-29	54.1	31.5	14.5	55.9	34.7	9.4
30-34	28.6	35.1	36.3	50.8	34.8	14.4
35-39	29.3	35.8	34.9	47.0	34.9	18.0
40-44	22.5	41.3	36.2	43.8	35.8	20.4
45-49	16.0	30.0	54.1	40.7	35.1	24.2
50-54	12.1	36.8	51.1	38.2	35.3	26.5
55-59	9.4	23.2	67.5	35.3	34.1	30.6
60-64	10.9	10.5	78.6	33.9	32.7	33.4
65-69	7.0	12.3	80.7	32.2	31.4	36.4
70-74	0.8	1.9	97.4	31.0	30.3	38.7
75-79	0.3	11.8	87.9	29.1	29.1	41.8
80-84	1.4	0.6	98.0	28.8	28.8	42.4
85-89	0.0	0.8	99.2	27.8	27.9	44.3
90-94	0.0	0.0	100.0	27.7	27.8	44.5
95-100	#N/A	#N/A	#N/A	27.7	27.8	44.5

Figure 18: Confidence intervals for estimated poverty likelihoods for being very poor associated with scores

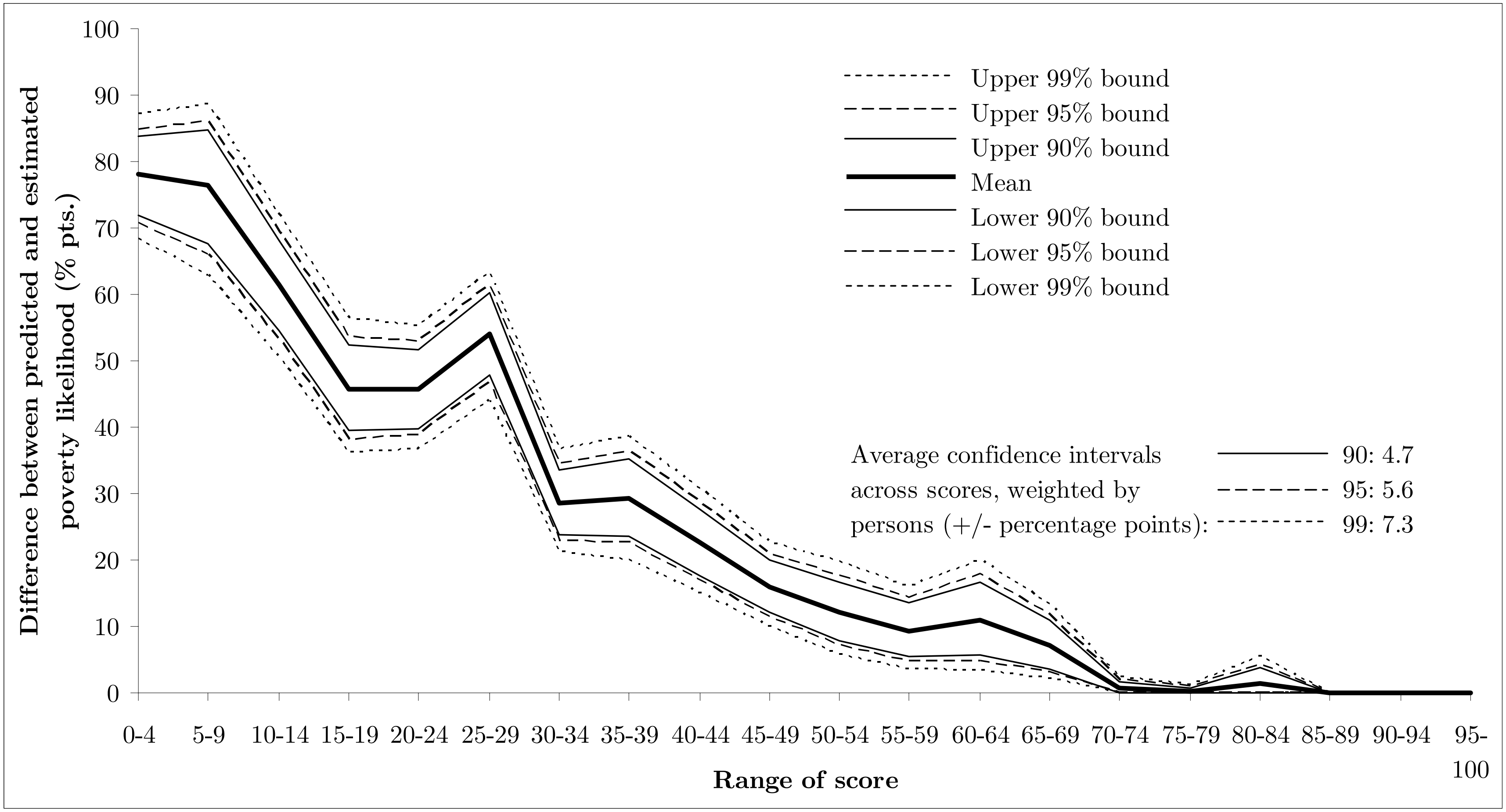

Figure 19: Differences between estimated and true poverty likelihoods for the very poor

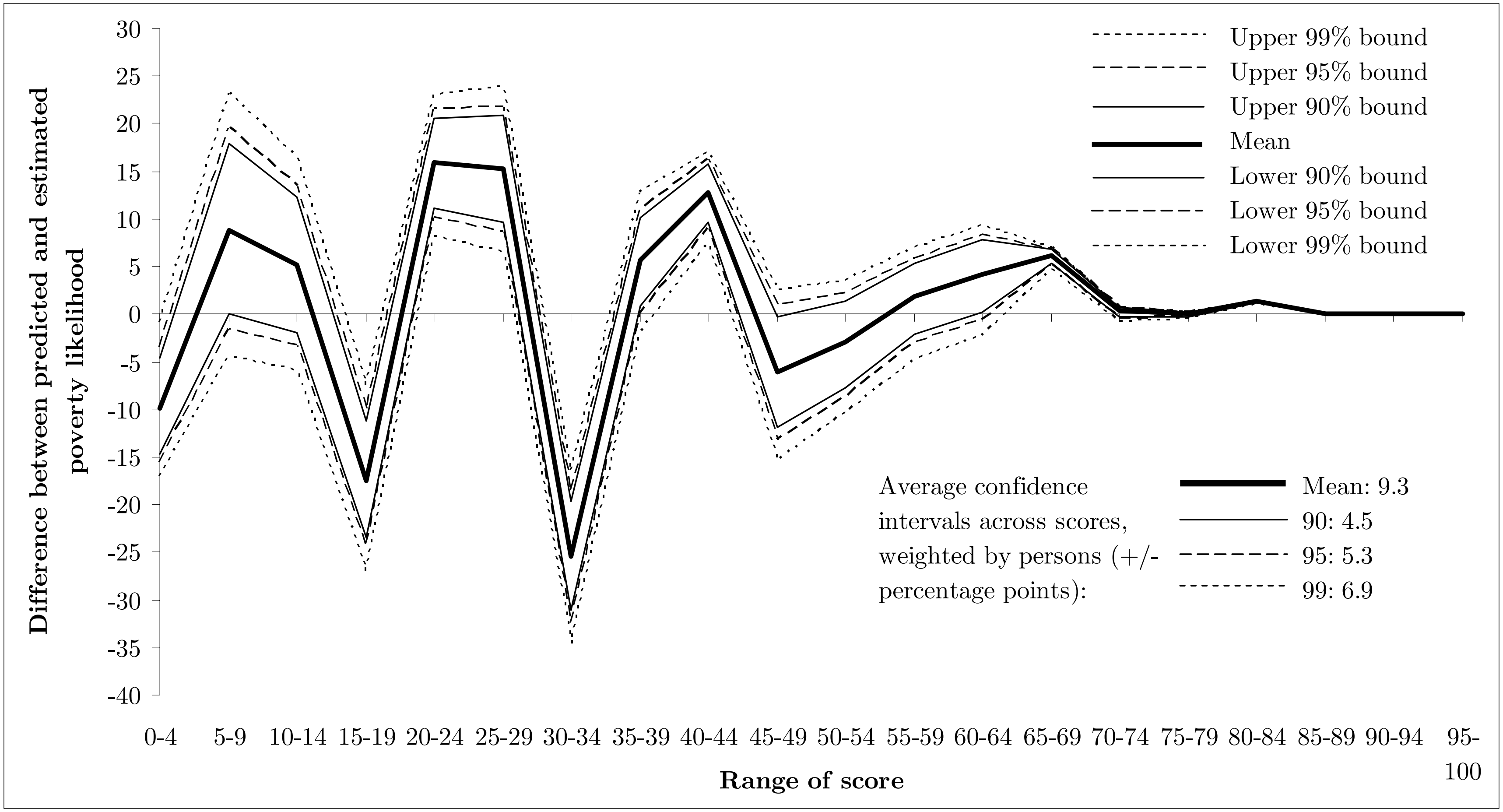

Figure 20: ROC curve of ability to rank-order households by very poor versus not very poor poverty status

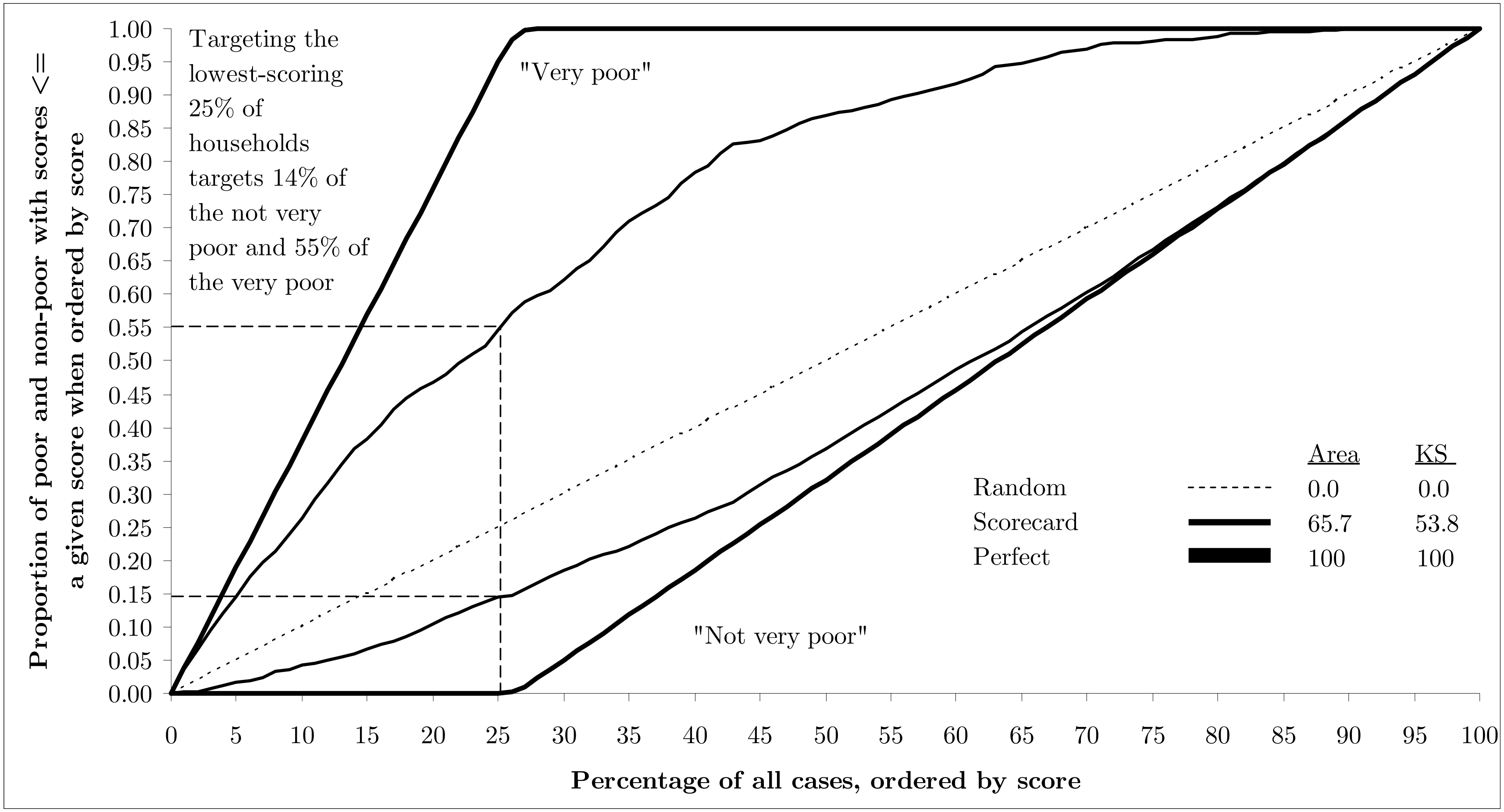

Figure 21: Classification matrix, three segments

		Targeting segment		
		Very Poor	**Poor**	**Non-poor**
True poverty status	**Very Poor**	**A.** Truly very poor correctly targeted as very poor	**B.** Truly very poor incorrectly targeted as poor	**C.** Truly very poor incorrectly targeted as non-poor
	Poor	**D.** Truly poor incorrectly targeted as very poor	**E.** Truly poor correctly targeted as poor	**F.** Truly poor incorrectly targeted as non-poor
	Non-poor	**G.** Truly non-poor incorrectly targeted as very poor	**H.** Truly non-poor incorrectly targeted as poor	**I.** Truly non-poor correctly targeted as poor

Figure 22: Net-benefit matrix, three segments

True poverty status	Targeting segment		
	Very Poor	**Poor**	**Non-poor**
Very Poor	α	β	γ
Poor	δ	ε	ζ
Non-poor	η	θ	ι

Figure 23: Classification results, very poor/poor cut-offs from 0 to 44 and poor/non-poor cut-offs from 5 to 49

Upper bound, very poor segment	Upper bound, poor segment																										
	5-9			**10-14**			**15-19**			**20-24**			**25-29**			**30-34**			**35-39**			**40-44**			**45-49**		
0-4	59	51	555	59	122	483	59	190	416	59	272	333	59	372	233	59	423	183	59	483	123	59	522	84	59	548	57
	16	14	636	16	56	594	16	115	535	16	193	457	16	251	399	16	313	337	16	386	264	16	458	192	16	508	142
	0	1	1,066	0	5	1,063	0	26	1,041	0	45	1,022	0	72	995	0	136	931	0	207	860	0	270	797	0	360	707
5-9				109	72	483	109	140	416	109	222	333	109	322	233	109	372	183	109	432	123	109	471	84	109	498	57
				31	42	594	31	101	535	31	179	457	31	237	399	31	299	337	31	372	264	31	443	192	31	493	142
				2	3	1,063	2	25	1,041	2	44	1,022	2	71	995	2	135	931	2	206	860	2	269	797	2	359	707
10-14							181	68	416	181	150	333	181	250	233	181	300	183	181	360	123	181	399	84	181	426	57
							73	59	535	73	137	457	73	195	399	73	257	337	73	330	264	73	401	192	73	451	142
							5	21	1,041	5	41	1,022	5	68	995	5	131	931	5	203	860	5	265	797	5	356	707
15-19										249	82	333	249	182	233	249	233	183	249	293	123	249	332	84	249	358	57
										131	78	457	131	136	399	131	198	337	131	271	264	131	343	192	131	393	142
										26	19	1,022	26	46	995	26	110	931	26	181	860	26	244	797	26	334	707
20-24													331	100	233	331	150	183	331	210	123	331	249	84	331	276	57
													209	58	399	209	120	337	209	193	264	209	265	192	209	315	142
													46	27	995	46	91	931	46	162	860	46	225	797	46	315	707
25-29																431	50	183	431	110	123	431	149	84	431	176	57
																268	62	337	268	135	264	268	206	192	268	256	142
																73	64	931	73	135	860	73	198	797	73	288	707
30-34																			482	60	123	482	99	84	482	126	57
																			329	73	264	329	145	192	329	195	142
																			136	71	860	136	134	797	136	224	707
35-39																						542	39	84	542	66	57
																						403	71	192	403	122	142
																						208	63	797	208	153	707
40-44																									581	27	57
																									474	50	142
																									270	90	707
45-49																											

Figures in units of 10,000 people.

Figure 23 (cont.): Classification results, very poor/poor cut-offs from 0 to 49 and poor/non-poor cut-offs from 50 to 100

		Upper bound, poor segment																													
		50-54			55-59			60-64			65-69			70-74			75-79			80-84			85-89			90-94			95-100		
Upper bound, very poor segment	0-4	59	566	40	59	583	23	59	595	11	59	604	1	59	605	1	59	605	0	59	606	0	59	606	0	59	606	0	59	606	0
		16	561	89	16	603	47	16	615	35	16	631	19	16	633	17	16	649	1	16	649	1	16	650	0	16	650	0	16	650	0
		0	434	633	0	556	511	0	645	422	0	751	316	0	830	237	0	955	113	0	980	87	0	1,058	10	0	1,067	0	0	1,067	0
	5-9	109	515	40	109	532	23	109	545	11	109	554	1	109	554	1	109	555	0	109	555	0	109	555	0	109	555	0	109	555	0
		31	547	89	31	588	47	31	600	35	31	617	19	31	618	17	31	635	1	31	635	1	31	636	0	31	636	0	31	636	0
		2	433	633	2	555	511	2	644	422	2	750	316	2	829	237	2	954	113	2	979	87	2	1,056	10	2	1,066	0	2	1,066	0
	10-14	181	443	40	181	460	23	181	473	11	181	482	1	181	483	1	181	483	0	181	483	0	181	483	0	181	483	0	181	483	0
		73	505	89	73	547	47	73	559	35	73	575	19	73	576	17	73	593	1	73	593	1	73	594	0	73	594	0	73	594	0
		5	429	633	5	551	511	5	641	422	5	747	316	5	825	237	5	950	113	5	975	87	5	1,053	10	5	1,063	0	5	1,063	0
	15-19	249	376	40	249	393	23	249	405	11	249	414	1	249	415	1	249	415	0	249	416	0	249	416	0	249	416	0	249	416	0
		131	446	89	131	488	47	131	500	35	131	516	19	131	518	17	131	534	1	131	534	1	131	535	0	131	535	0	131	535	0
		26	408	633	26	530	511	26	619	422	26	725	316	26	804	237	26	929	113	26	954	87	26	1,032	10	26	1,041	0	26	1,041	0
	20-24	331	293	40	331	310	23	331	323	11	331	332	1	331	333	1	331	333	0	331	333	0	331	333	0	331	333	0	331	333	0
		209	368	89	209	410	47	209	422	35	209	438	19	209	439	17	209	456	1	209	456	1	209	457	0	209	457	0	209	457	0
		46	389	633	46	511	511	46	600	422	46	706	316	46	784	237	46	909	113	46	934	87	46	1,012	10	46	1,022	0	46	1,022	0
	25-29	431	193	40	431	210	23	431	223	11	431	232	1	431	233	1	431	233	0	431	233	0	431	233	0	431	233	0	431	233	0
		268	310	89	268	352	47	268	364	35	268	380	19	268	381	17	268	398	1	268	398	1	268	399	0	268	399	0	268	399	0
		73	362	633	73	484	511	73	573	422	73	679	316	73	758	237	73	883	113	73	908	87	73	985	10	73	995	0	73	995	0
	30-34	482	143	40	482	160	23	482	172	11	482	182	1	482	182	1	482	183	0	482	183	0	482	183	0	482	183	0	482	183	0
		329	248	89	329	290	47	329	302	35	329	318	19	329	320	17	329	336	1	329	336	1	329	337	0	329	337	0	329	337	0
		136	298	633	136	420	511	136	509	422	136	615	316	136	694	237	136	819	113	136	844	87	136	922	10	136	931	0	136	931	0
	35-39	542	83	40	542	100	23	542	112	11	542	122	1	542	122	1	542	123	0	542	123	0	542	123	0	542	123	0	542	123	0
		403	175	89	403	217	47	403	229	35	403	245	19	403	246	17	403	263	1	403	263	1	403	264	0	403	264	0	403	264	0
		208	227	633	208	349	511	208	438	422	208	544	316	208	622	237	208	747	113	208	772	87	208	850	10	208	860	0	208	860	0
	40-44	581	44	40	581	61	23	581	73	11	581	83	1	581	83	1	581	84	0	581	84	0	581	84	0	581	84	0	581	84	0
		474	103	89	474	145	47	474	157	35	474	173	19	474	175	17	474	192	1	474	192	1	474	192	0	474	192	0	474	192	0
		270	164	633	270	286	511	270	375	422	270	481	316	270	560	237	270	685	113	270	710	87	270	788	10	270	797	0	270	797	0
	45-49	607	17	40	607	34	23	607	47	11	607	56	1	607	57	1	607	57	0	607	57	0	607	57	0	607	57	0	607	57	0
		524	53	89	524	95	47	524	107	35	524	123	19	524	125	17	524	141	1	524	142	1	524	142	0	524	142	0	524	142	0
		360	74	633	360	196	511	360	285	422	360	391	316	360	470	237	360	595	113	360	620	87	360	697	10	360	707	0	360	707	0

Figures in units of 10,000 people.

Figure 23 (cont.): Classification results, very poor/poor cut-offs from 50 to 94 and poor/non-poor cut-offs from 55 to 100

Upper bound, very poor segment		55-59			60-64			65-69			70-74			75-79			80-84			85-89			90-94			95-100		
50-54		625	17	23	625	29	11	625	39	1	625	39	1	625	40	0	625	40	0	625	40	0	625	40	0	625	40	0
		577	42	47	577	54	35	577	70	19	577	71	17	577	88	1	577	88	1	577	89	0	577	89	0	577	89	0
		434	122	511	434	211	422	434	317	316	434	396	237	434	521	113	434	546	87	434	624	10	434	633	0	434	633	0
55-59					642	12	11	642	22	1	642	22	1	642	23	0	642	23	0	642	23	0	642	23	0	642	23	0
					619	12	35	619	28	19	619	30	17	619	46	1	619	47	1	619	47	0	619	47	0	619	47	0
					556	89	422	556	195	316	556	274	237	556	399	113	556	424	87	556	502	10	556	511	0	556	511	0
60-64								654	9	1	654	10	1	654	10	0	654	11	0	654	11	0	654	11	0	654	11	0
								631	16	19	631	18	17	631	34	1	631	35	1	631	35	0	631	35	0	631	35	0
								646	106	316	646	185	237	646	309	113	646	334	87	646	412	10	646	422	0	646	422	0
65-69											663	1	1	663	1	0	663	1	0	663	1	0	663	1	0	663	1	0
											647	1	17	647	18	1	647	18	1	647	19	0	647	19	0	647	19	0
											752	78	237	752	203	113	752	228	87	752	306	10	752	316	0	752	316	0
70-74														664	0	0	664	1	0	664	1	0	664	1	0	664	1	0
														649	17	1	649	17	1	649	17	0	649	17	0	649	17	0
														830	125	113	830	150	87	830	228	10	830	237	0	830	237	0
75-79																	664	0	0	664	0	0	664	0	0	664	0	0
																	666	0	1	666	1	0	666	1	0	666	1	0
																	955	25	87	955	103	10	955	113	0	955	113	0
80-84																				665	0	0	665	0	0	665	0	0
																				666	1	0	666	1	0	666	1	0
																				980	78	10	980	87	0	980	87	0
85-89																							665	0	0	665	0	0
																							666	0	0	666	0	0
																							1,058	10	0	1,058	10	0
90-94																										665	0	0
																										666	0	0
																										1,068	0	0

Figures in units of 10,000 people.

Figure 24: Classification results, very poor 0–39, poor 40–54, and non-poor 55–100

		People with score in range					
Segment	**Score**	**Very Poor**		**Poor**		**Non-poor**	
	0-4		59		16		0
	5-9		51		14		1
	10-14		72		42		3
Very poor	15-19	**542**	68	**403**	59	**208**	21
0-24	20-24	**81%**	82	**60%**	78	**19%**	19
	25-29		100		58		27
	30-34		50		62		64
	35-39		60		73		71
Poor	40-44	**83**	39	**175**	71	**227**	63
25-34	45-49	**12%**	27	**26%**	50	**21%**	90
	50-54		17		53		74
	55-59		17		42		122
	60-64		12		12		89
	65-69		9		16		106
Non-poor	70-74	**40**	1	**89**	1	**633**	78
35-100	75-79	**6%**	0	**13%**	17	**59%**	125
	80-84		0		0		25
	85-89		0		1		78
	90-94		0		0		10
	95-100		0		0		0
	Total:		665		666		1,068

Counts of people are in units of 10,000.

Figure 25: An example net-benefit matrix reflecting common values

True poverty status	Targeting segment: Very Poor	Targeting segment: Poor	Targeting segment: Non-poor
Very Poor	+3	-2	-6
Poor	-1	+2	-2
Non-poor	-2	-1	+1

Note: This is an example. Each program should define its own net-benefit matrix.

Figure 26: Computation of total net benefit for a cut-off pair of 35–39 and 50–54

Cell			Persons	Net benefit/person	Net benefit
A.	Truly very poor	as very poor	542	+3	+1,626
B.	Truly very poor	as poor	83	–2	–166
C.	Truly very poor	as non-poor	40	–6	–240
D.	Truly poor	as very poor	403	–1	–403
E.	Truly poor	as poor	175	+2	+350
F.	Truly poor	as non-poor	89	–2	–178
G.	Truly non-poor	as very poor	208	–2	–416
H.	Truly non-poor	as poor	227	–1	–227
I.	Truly non-poor	as non-poor	633	+1	+633
				Total net benefit:	+979

Note: Persons are counted in units of 10,000.