

Is One Simple Poverty Scorecard Enough for India?

Mark Schreiner

October 19, 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

Is one simple poverty scorecard enough for India? Using out-of-sample bootstrap tests, this paper compares accuracy for an All-India scorecard versus urban- and rural-specific scorecards. Even though India is large and diverse, there is no evidence that segmentation provides large increases in accuracy.

Acknowledgements

This paper was funded by Grameen Foundation USA and managed by Nigel Biggar. Data was provided by India's National Sample Survey Organisation. I am grateful for comments from Nigel Biggar, Dean Caire, Shiva Makki, Geetha Nagarajan, Frances Sinha, and Jeff Toohig.

Is One Simple Poverty Scorecard Enough for India?

1. Introduction

Some potential users of the simple All-India poverty scorecard in Schreiner (2006a) have expressed concern about applying a single scorecard in such a large and diverse country. In particular, poverty rates differ sharply between urban and rural areas, suggesting that the indicators of poverty may also differ along urban/rural lines.

Is one poverty scorecard sufficient? Or would segmented scorecards improve accuracy a lot? A similar test in Mexico (Schreiner, 2006b) found only small benefits to segmentation. The research here for India comes to the same broad conclusion.

2. Methods

Poverty scorecards were constructed and tested for three segments:

- All-India (both urban and rural)
- Urban only
- Rural only

About 25 percent of Indians are urban and 75 percent are rural (Figure 1). In the survey, *rural* and *urban* appear to follow official census definitions.

Each scorecard was constructed from scratch using a one-half random sample of the relevant segment of the 41,013 households in Schedule 1.0 of the 59th Round (2003) of India's Socio-Economic Survey (NSSO, 2005). The average household represented

about 24,000 people. Sixteen households who each represented more than 500,000 people were omitted because their inclusion led to the breakdown of some bootstrap estimates (see Singh, 1998).

For each scorecard, the association between poverty scores (sum of scorecard points) and poverty likelihoods (probability of being poor) was derived by bootstrapping one-fourth of the remaining relevant segment from the 59th Round.

A household is “poor” if its expenditure is less than \$1/day/person.¹ The poverty likelihood associated with a given score in a given segment is the share of people—averaged across bootstrap samples—who have that score and who were poor.

Accuracy was tested by applying the scorecards to bootstrap samples of the final one-fourth of the 59th Round and looking at:

- Differences between predicted and true poverty likelihoods
- Differences between estimated and true overall poverty rates
- Rank-ordering of households by poverty scores

These measures are compared for the All-India scorecard applied to urban and rural and then for the urban and rural scorecards applied to their respective segments.

¹ Adjusted for purchasing-power parity and regional prices (see Sillers, 2006; Deaton, 2003; and the Appendix).

3. Poverty lines

No country has more poor people than India (200–900 million, depending on the poverty line). There is an entire literature—but little agreement—about Indian poverty lines (Deaton and Kozel, 2005a). Some of them include:

- \$1/person/day (Rs14.91 at purchasing power parity in 2003)
- Official all-India lines (Rs11.51 rural, Rs16.79 urban)
- Official state-wise and rural/urban lines, adjusted for cost-of-living

The state-wise lines are better because they adjust for cost-of-living. The cost-of-living adjustments in Deaton (2003) are better than the official adjustments because they account for what households actually buy.

Decades ago, the official lines were based on caloric benchmarks. Over time, however, the line fell behind the cost of food, and of course “poverty” is more than just lack of calories. Taking a “basic needs” approach, Abraham (2005) estimates a poverty rate of 90 percent. At the same time, Deaton (2004) and Aiyar (2003) argue that mechanization has decreased caloric requirements and that per-rupee food quality—and nutrient density—has improved.

Officially, about one in four Indians are poor. But less than half usually eat three meals per day, so common sense suggests that the official lines are too low.

The scorecards here use the \$1/day poverty line, adjusted for cost-of-living using Deaton (2003). This international benchmark is usually higher than the official lines in rural areas but lower in urban areas. It is higher than Deaton’s (2003) lines in all areas.

Figures 2 and 3 show 2003 poverty lines by state and rural/urban segment. The overall poverty rate was 23.6 percent by the official line, 15.1 percent by the Deaton-adjusted line, and 37.4 percent by the \$1PPP line.

4. Scorecards

The three scorecards here are simple, easy-to-use, and objective. They aim to help development programs to target services, track changes in poverty over time, and report clients' poverty rates.

Indicators were selected to be:

- Inexpensive to collect, easy to answer quickly, and simple to verify
- Strongly correlated with poverty
- Liable to change values 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). Scores can be computed quickly by hand in the field.

A score corresponds to a “poverty likelihood”, that is, the probability of being poor. For a group of people, the share who are poor is the average poverty likelihood. For a group over time, progress is the change in average poverty likelihood.

Schreiner (2006a) describes scorecard construction. Preference was given to verifiable indicators whose values are liable to change over time. Scorecard construction seeks not only to achieve technical accuracy but also to facilitate proper use (Schreiner, 2005). When scoring projects fail, the issue is usually not inaccuracy but rather the

refusal of users to accept scoring (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”.

To allow field workers to compute scores by hand in real time, the one-page scorecards feature:

- Only 10 indicators
- Only categorical indicators (“number of fans owned”, not “value of all assets”)
- User-friendly weights (non-negative integers, no arithmetic beyond simple addition)

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

The scorecards in Figures 4, 5, and 6 can be photocopied for immediate use. They can also serve as templates for data-entry screens for database software that records indicators, scores, and poverty likelihoods.

The rural scorecard (Figure 6) resembles the all-India scorecard because three-fourths of all Indians (and almost 9 in 10 poor Indians) are rural.

5. Scores and poverty likelihoods

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

- From a one-fourth hold-out sample from the relevant segment in the 59th Round, 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 bootstrapped samples

These poverty likelihoods are objective, that is, based on data. The process just described would produce objective poverty likelihoods *even if the scorecards were constructed without data.*² Of course, the scorecards here were constructed with data. The point is being made here only because some parties have misunderstood the significance of the fact that some choices in scorecard construction—as in any statistical analysis—were informed by the analyst’s judgment and knowledge of microfinance. That this judgment is explicitly acknowledged in no way impunes the objectivity of the poverty likelihoods; their objectivity depends on using data to associate scores with poverty likelihoods, not on whether only data was used to construct scorecards.

² In fact, scorecards of objective, proven accuracy are often constructed based *only* on qualitative judgment (Fuller, 2006; Caire, 2004; Schreiner *et al.*, 2004).

Poverty likelihoods by score for the All-India, urban, and rural scorecards are in Figures 7, 8, and 9. For example, scores of 5–9 in the All-India scorecard in Figure 7 correspond to a poverty likelihood of 96.3 percent. That is, on average across the 10,000 bootstrap samples, 96.3 percent of all Indians with scores of 5–9 were poor.

For the urban scorecard applied to urban areas, scores of 5–9 were associated with a poverty likelihood of 75.0 percent. For the rural scorecard applied to rural areas, scores of 5–9 correspond to a poverty likelihood of 88.3 percent.

The precision of poverty likelihoods is best expressed as confidence intervals. Figures 10, 11, and 12 show poverty likelihoods for the three scorecards along with 90-, 95-, and 99-percent confidence intervals.

For example, the average poverty rate across bootstrap samples for the All-India scorecard for people with scores of 30–34 (the poverty likelihood) was 52.8 percent. In 90 percent of the 10,000 samples, the rate was between 49.7–56.0 percent. In 95 percent of samples, the share was between 49.1–56.7, and in 99 percent of samples, the share was between 48.0–57.8.

Figure 13 shows average confidence intervals for the poverty likelihood, weighting by the number of people falling in each score range. In general, the 90-percent interval is about ± 3 percentage points, the 95-percent interval is about ± 4 percentage points, and the 99-percent interval is about ± 5 percentage points. The all-India scorecard has the narrowest intervals, followed by rural, then urban.

Narrower confidence intervals mean greater precision. The width of the intervals depends—among other things—on the size of the hold-out sample, the number of people with a given score, the accuracy of the scorecard, the extent of overfitting, and the presence of households with extreme weights in the bootstrap samples.

In general, the poverty likelihoods consistently decrease as scores increase, which makes sense. The main exception is in the urban scorecard (Figure 11) for scores of 5–9, 10–14, and 15–19. Why would scores go from 75.0 percent, down to 32.0 percent, then back up to 55.2 percent?

In the poverty-likelihood hold-out sample, about half of urban households who scores of 10–14 were poor. But the non-poor households represented about 30 percent more people, pulling the estimated poverty likelihood down. Unlike the case of the 16 households omitted from the overall survey who represented more than 500,000 people, the disproportionate weights here did not result from a small group of households, so it was not possible to omit a couple of outliers and get a more reasonable result.³

³ The dip in poverty likelihood persists with different random construction and bootstrap samples, so it is not simply due to an unfortunate draw.

6. Accuracy tests

Accuracy was tested by applying the scorecards to 10,000 bootstrap replicates of the final one-fourth of the data and looking at:

- Differences between predicted and true poverty likelihoods
- Differences between estimated and true overall poverty rates
- Rank-ordering of households by poverty scores

To measure gains from segmentation, the accuracy of the All-India scorecard applied to urban and rural areas was compared with the accuracy of the urban and rural scorecards applied to their respective segments.

6.1 Differences between estimated and true poverty likelihoods

Figures 14, 15, and 16 depict the differences (in 10,000 bootstrap samples) between estimated and true poverty likelihoods from the All-India, urban, and rural scorecards applied to their respective segments. The information in these three figures is summarized in Figure 17, which presents the average (population-weighted) absolute differences and 90-, 95-, and 99-percent confidence intervals.

The estimated poverty likelihoods of the All-India scorecard applied to all India are more accurate (smaller mean absolute difference, compared to true poverty likelihoods) than the urban or rural scorecards applied to their respective segments (3.6 percentage points versus 8.0 and 6.8). The confidence intervals for all three scorecards are not wide, between ± 4 or 5 percentage points.

Compared with the segment-specific scorecards, the estimated poverty likelihoods of the All-India scorecard applied to the urban and rural segments (Figures 18 and 19) are about as accurate (8.3 percentage-point mean absolute difference versus 8.0 for urban, Figure 17) or more accurate (4.0 versus 6.8 for rural). The All-India scorecard also has narrower confidence intervals. In terms of targeting accuracy, there no downside to the All-India scorecard vis-à-vis the segment-specific scorecards.

Are these estimates of poverty likelihoods accurate enough to justify using the scorecard for targeting? This is a judgment call. The judgment of the author is that the errors due to scorecard inaccuracy are probably not large relative to other sources of error (such as mistakes or fraud in data collection) and relative to the level of targeting accuracy available from other low-cost tools. Comparisons with other poverty scorecards are not possible because this sort of test has not been reported elsewhere.

6.2 Distribution of the differences between estimated and true overall poverty rates

One further aspect of accuracy is the “closeness” of the estimated overall poverty rate (average poverty likelihood) to the true poverty rate (share of surveyed people who are poor). Given a scorecard and a segment to which the scorecard is applied, the accuracy of the estimated overall poverty rate is tested here by:

- Drawing a bootstrap sample from the final one-fourth of the 59th Round
- Computing scores and converting them to poverty likelihoods
- Computing the difference between the average poverty likelihood (the estimated overall poverty rate) and the true overall poverty rate

- Repeating the previous three steps 10,000 times
- Checking the distribution of differences between true and estimated poverty rates

According to Figure 20, the All-India scorecard applied to all India gives a precise⁴ (standard error 0.0039) and somewhat biased⁵ estimate of overall poverty rates (on average, the estimate differs from the true value by 2.00 percentage points).

The bias is likely due to:

- Sampling variation in the selection of the one-half construction sample and in the two one-fourth hold-out samples
- Transforming and rounding the scorecard weights to be non-negative integers
- Overfitting

Of course, the bias is easily eliminated; apply the All-India scorecard to all India, and subtract 2.00 percentage points from the average poverty likelihood. The result will then be, with 90-percent certainty, within ± 0.7 percentage points of the true value.

The bias is similar (2.05 and 1.99 percentage points) for the All-India scorecard applied to urban and rural India. Likewise, the 90-percent confidence intervals are still less than a percentage point (Figure 20).

The urban scorecard applied to urban India produces a slightly less biased estimate of the overall poverty rate (-1.47 percentage points, Figure 20). Still, this is

⁴ A statistical estimator is *precise* if, in repeated samples (as in the bootstrap), the distribution of estimates is narrowly bunched so that all samples give very similar estimates, regardless of whether these estimates are close to the true value.

⁵ A statistical estimator is *unbiased* if, in repeated samples (as in the bootstrap), its average equals the true value being estimated, even though the distribution of estimates around the true value may be wide or narrow.

only half a percentage point closer to the true value than is the All-India scorecard. In rural areas, the All-India scorecard is *less* biased than the rural-specific scorecard.

In terms of estimating overall poverty rates, there is no evidence of large gains to from the use of segment-specific scorecards. In any case, there is not much room for improvement, considering that the biases are known and that the 90-percent confidence intervals are smaller than 1 percentage point.

6.3 Rank-ordering of households by poverty scores

ROC curves are standard tools for comparing, at a glance, how well scorecards rank-order cases, that is, give poor households lower poverty scores than non-poor households (Baulch, 2003; Wodon, 1997). They plot the share of poor and non-poor households against the share of all households ranked by score.

Figures 21 and 22 are ROC curves for:

- All-India scorecard applied to urban, versus urban scorecard applied to urban
- All-India scorecard applied to rural, versus rural scorecard applied to rural

What do the ROC curves mean? Suppose a program sets a policy cut-off so as to target the lowest-scoring x percent of potential participants. For a given scorecard, the ROC curves then show the share of the poor and non-poor who would be targeted. Greater ability to rank-order—with less leakage and less undercoverage—is shown by curves that are closer the northwest and southeast corners of the graph.

In Figures 21 and 22, the three 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 next two lines inward are from actual scorecards. The thinner solid line is the segmented urban (or rural) scorecard applied to its segment, while the dashed line is the All-India scorecard applied to urban (or rural) areas.

In general, the curves are quite close to each other, showing that the scorecards are about equal in rank-ordering. Because of the dip in poverty likelihood for scores of 10–14 discussed earlier, the All-India scorecard in Figure 21 (slightly) beats the urban-specific scorecard in terms of rank-ordering the urban households most-likely to be poor, although the urban-specific scorecard is (slightly) better for higher scores. For rural households, the rural-specific scorecard in Figure 22 is always better than the All-India scorecard, although the differences are again slight.

For example, targeting the 20 percent of urban cases with the lowest scores—regardless of scorecard—would target 35–37 percent of all the poor and 7–9 percent of all the non-poor. In the rural case, targeting the 20 percent with the lowest scores would target about 64–69 percent of all the poor and 9 percent of all the non-poor.⁶

Figures 21 and 22 report two other common measures of ability to rank-order: the Kolmogorov-Smirnov statistic (Hoadley and Oliver, 1998) and the ratio of the area

⁶ With a random targeting—the diagonal dotted line in the figures—the lowest-scoring 20 percent of potential participants would include 20 percent of all the poor and 20 percent of all the non-poor.

inside a ROC curves for a specific scorecard to the area inside the trapezoid of a hypothetical perfect scorecard. The KS statistic measures the maximum distance between the curves for the poor and non-poor, and it does not vary much by scorecard (44.9 versus 45.1 for urban, and 61.2 versus 64.0 for rural). Likewise, the area measure does not vary much (61.1 versus 59.9 for urban, and 78.9 versus 80.4 for rural).

This again suggests that for targeting, the All-India scorecard is sufficient, with perhaps some slight gains to be had from the rural-specific scorecard in rural segments.

7. Conclusion

Is one poverty scorecard sufficient? Or would segmented scorecards greatly improve accuracy?

A test in Mexico (Schreiner, 2006a) found only small benefits to segmentation, and a similar, more complete test here for India reaches the same broad conclusion. At least in these two cases, a single, simple country-wide scorecard is about as good as two segmented scorecards, one for urban areas and one for rural.

Accuracy was tested with out-of-sample bootstrap tests, looking at:

- Differences between predicted and true poverty likelihoods
- Differences between estimated and true overall poverty rates
- Rank-ordering of households by poverty scores

In general, the All-India scorecard was as accurate for targeting purposes as the rural- and urban-specific scorecards (the rural scorecard was slightly better at rank-

ordering). In terms of the overall poverty rate, the All-India and rural scorecards were in a dead heat, while the urban scorecard was slightly better than All-India.

In any case, the differences are small, smaller than several other potential sources of error, and certainly smaller than the level of precision that most applications will likely require. In sum, there is little evidence of large returns to incurring the cost and complexity of constructing and using segmented scorecards.

In the case of India, of course, segmented scorecards now exist, and so it will probably make sense to use them, not so much because they offer greater accuracy but rather because they are more likely to be accepted by potential users.

The results here reinforce the idea that simple poverty scorecards offer development programs a quick, easy, inexpensive, and accurate way to identify the poor and thus improve their targeting, track progress out of poverty through time, and report to external stakeholders the share of their participants who are poor.

A final question remains: How can a country-wide scorecard predict about as well as segment-specific scorecards?

Several factors probably are relevant:

- Poverty is relatively easy to predict
- The statistical phenomenon known as the “flat max”
- Because the All-India sample is larger, the bootstrap tests are less likely to be affected by a few households

Poverty is easy to predict because most households are similar; whether urban or rural, they all want an faster, cleaner way to cook and a sturdier, dryer house. Furthermore, poverty is a state that results from historical accumulation (of opportunity, effort, and luck). It turns out, therefore, that an excellent indicator of expenditure-based poverty is the ownership of a few specific assets that most households tend to accumulate as their expenditure increases.

In the rural and urban scorecards, five of the ten indicators are the same. What's more, any of the scorecards only requires about 5 indicators to achieve 95 percent of their predictive power. This fact—known in predictive modelling as the “flat max”⁷—suggests that the All-India scorecard captures enough variation in its indicators to cover both rural and urban areas.

Finally, bootstrap estimates converge to true population values more slowly (or not at all) when there are cases with extreme sampling weights and/or when sample sizes are smaller. Thus, part of the nondescriptedness of the performance of the urban- and rural-specific scorecards may be due to their smaller sample sizes' reducing the power of their bootstrap tests.

⁷ On the “flat max”, see Lovie and Lovie (1986), Dawes (1979), and Wainer (1976).

References

- Abraham, Ronald Joseph. (2005) "Tripping over the Poverty Line", Sept. 15, Center for Policy Alternatives, <http://www.cpasind.com/articles/ot-05-09-15-tripping-over-poverty-line.html>.
- Aiyar, Swaminathan S. Anklesaria. (2003) "It's not just Calories, Stupid", *Economic Times*, Dec. 10, http://www.swaminomics.org/et_articles/et20031210_not_calories_stupid.htm.
- 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>.
- Caire, Dean. (2004) "Building Credit Scorecards for Small Business Lending in Developing Markets", Bannock Consulting, http://www.microfinance.com/English/Papers/Scoring_SMEs_Hybrid.pdf.
- Dawes, Robyn M. (1979) "The Robust Beauty of Improper Linear Models in Decision Making", *American Psychologist*, Vol. 34, No. 7, pp. 571–582.
- Deaton, Angus. (2004) "Measuring Poverty", Research Program in Development Studies, Princeton University, http://www.wws.princeton.edu/rpds/downloads/deaton_povertymeasured.pdf.
- (2003) "Prices and Poverty in India, 1987–2000", *Economic and Political Weekly*, Jan. 25, pp. 362–368, http://www.wws.princeton.edu/rpds/downloads/deaton_prices_poverty_in_dia.pdf.
- ; and Valerie Kozel. (2005) *The Great Indian Poverty Debate*, Delhi: MacMillan, ISBN 1403-92644-1.
- Deaton, Angus; and Alessandro Tarozzi. (2000) "Prices and Poverty in India", Research Program in Development Studies, Princeton University, http://www.wws.princeton.edu/rpds/downloads/deaton_tarozzi_prices_poverty.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.
- 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.
- Lovie, A.D.; and P. Lovie. (1986) “The Flat Maximum Effect and Linear Scoring Models for Prediction”, *Journal of Forecasting*, Vol. 5, pp. 159–168.
- National Sample Survey Organisation. (2005) “Household Consumer Expenditure and Employment-Unemployment Situation in India: NSS 59th Round, January–December 2003”, Report No. 490, Ministry of Statistics and Programme Implementation, Government of India,
http://mospi.nic.in/rept%20_%20pubn/ftest.asp?rept_id=495&type=NSS0.
- Schreiner, Mark. (2006a) “A Simple Poverty Scorecard for India”, report to Grameen Foundation U.S.A.,
http://www.microfinance.com/English/Papers/Scoring_Poverty_India.pdf.
- (2006b) “Indices de Pobreza para México”, report to Grameen Foundation U.S.A.,
http://www.microfinance.com/Castellano/Documentos/Scoring_Poverty_Mexico.pdf.
- (2005) “IRIS questions on poverty scorecards”, memo for Grameen Foundation U.S.A., mark@microfinance.com.
- (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.
- ; 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.

Wainer, Howard. (1976) “Estimating Coefficients in Linear Models: It Don’t Make No Nevermind”, *Psychological Bulletin*, Vol. 83, pp. 213–217.

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

Appendix: \$1PPP Poverty Line Adjusted for Cost-of-Living Using Deaton (2003)

Poverty lines were selected to:

- Follow the World Bank's \$1/person/day (purchasing-power parity) line, as the poverty rates implied by India's official poverty line are unrealistically low (27.0 percent in urban India, and 23.5 percent in rural), and those of Deaton (2003) are even lower (9.5 percent urban, 21.6 percent rural)
- Account for rural/urban and state-wise cost-of-living using Deaton (2003)
- Match the average of the rural and urban \$1PPP lines to the all-India \$1PPP line
- Match the ratio of rural to urban \$1PPP lines to that same ratio for Deaton's lines

Basic inputs to the calculation are:

- \$1PPP/person/day for all-India in 2003 is Rs453.65/person/month
- In 2003, 74.82 percent of the population was rural, and 25.18 percent was urban
- Deaton's (2003) all-urban poverty line for 2000 is Rs354.11/person/month for urban and Rs309.32 for rural

The population-weighted average of rural and urban \$1PPP lines should match the all-India \$1PPP line:

$$\text{Rs}453.90 = (0.7482 \times \text{Rural } \$1\text{PPP line}) + (0.2518 \times \text{Urban } \$1\text{PPP line}).$$

Furthermore, the ratio of the two lines should match the ratio of Deaton's lines:

$$(\text{Rural } \$1\text{PPP line} \div \text{Urban } \$1\text{PPP line}) = \text{Rs}309.32 \div \text{Rs}354.11.$$

Solving the algebra gives:

- Rural \$1PPP line of Rs437.93/person/month
- Urban \$1PPP line of Rs501.35/person/month

To account for cost-of-living, Deaton's (2003) state-wise lines for 2000 are then adjusted by the ratio of the rural or urban \$1PPP line to Deaton's corresponding line. For both rural and urban areas, the adjustment factor is 1.41578.

In sum, the 2003 \$1PPP state poverty line for an urban (or rural) area is 1.41578 multiplied by the state's 2000 urban (or rural) line from Deaton (2003).

Figure 1: Households surveyed, people represented, and overall poverty rates by segment in Round 59

	<u>All</u>		
	Households	People	% poor
Constructing scorecards	20,345	488,499,660	37.7
Associating scores with likelihoods	10,387	250,311,643	37.6
Testing accuracy	10,265	242,156,618	36.7
Total:	40,997	980,967,921	37.4

	<u>Urban</u>		
	Households	People	% poor
Constructing scorecards	7,395	124,420,394	20.0
Associating scores with likelihoods	3,786	62,250,520	18.4
Testing accuracy	3,688	63,491,789	19.5
Total:	14,869	250,162,703	19.4

	<u>Rural</u>		
	Households	People	% poor
Constructing scorecards	12,950	364,079,266	43.7
Associating scores with likelihoods	6,601	188,061,122	43.9
Testing accuracy	6,577	178,664,829	42.8
Total:	26,128	730,805,217	43.5

Source: Calculations based on Schedule 1.0 of the 59th (2003) Rounds of India Socio-Economic Survey by the National Sample Survey Organization.

16 households were omitted because their very high weights (greater than 500,000 persons each) caused them to dominate some of the bootstrap tests. See Singh (1998).

**Figure 2: Rural poverty lines (Rs/person/month),
2003**

State	Official state-wise	Deaton-adjusted	\$1/day adjusted
Andhra Pradesh	280.93	330.80	438.35
Assam	390.43	363.19	481.28
Bihar	355.86	317.18	420.30
Gujarat	340.76	360.39	477.57
Haryana	387.63	332.03	439.98
Himachal Pradesh	392.59	386.06	511.58
Karnataka	330.77	344.67	456.73
Kerala	400.43	399.52	529.42
Madhya Pradesh	332.64	308.65	409.00
Maharashtra	340.43	341.73	452.84
Orissa	346.08	320.88	425.22
Punjab	387.49	338.14	448.08
Rajasthan	367.57	346.08	458.60
Tamil Nadu	328.69	359.54	476.44
Uttar Pradesh	359.93	299.68	397.11
West Bengal	374.13	327.83	434.42

Source: Calculations based on Schedule 1.0 of the 59th (2003) Rounds of India Socio-Economic Survey by the National Sample Survey Organization.

The first two columns are derived from Table 5 in Deaton (2003), scaled by 1.068403 for rural price increases from 2000 to 2003. The last column is the 2000 Deaton-adjusted line scaled by 1.41578 (see Appendix). As in Deaton and Tarozzi (2000), the poverty line for Jammu and Kashmir is taken as that of Himachal Pradesh; Chandigarh is taken as Punjab; Uttaranchal as Uttar Pradesh; rural Dehli as Haryana; Sikkim, Arunachal Pradesh, Nagaland, Manipur, Mizoram, Tripura, and Meghalaya to Assam; Jharkhand as Bihar; Chhattisgarh as Madhya Pradesh; Daman and Diu, Dadra and Nagar Haveli, and Goa as Maharashtra; Lakshadweep as Kerala; and Pondicherry and the Andaman and Nicobar Islands as Tamil Nadu.

Figure 3: Urban poverty lines (Rs/person/month), 2003

State	Official state-wise	Deaton-adjusted	\$1/day adjusted
Andhra Pradesh	514.45	387.76	488.10
Assam	386.89	426.26	536.57
Bihar	427.14	361.75	455.37
Gujarat	533.58	415.43	522.93
Haryana	472.61	403.08	507.39
Himachal Pradesh	472.61	424.75	534.67
Karnataka	575.23	413.02	519.90
Kerala	536.55	434.40	546.82
Madhya Pradesh	541.72	361.36	454.88
Maharashtra	607.02	433.42	545.58
Orissa	532.13	351.29	442.20
Punjab	436.56	394.25	496.27
Rajasthan	524.03	397.19	499.98
Tamil Nadu	534.92	411.74	518.29
Uttar Pradesh	468.21	360.38	453.64
West Bengal	460.26	386.35	486.33
Delhi	568.49	450.37	566.92

Source: Calculations based on Schedule 1.0 of the 59th (2003) Rounds of India Socio-Economic Survey by the National Sample Survey Organization.

The first two columns are derived from Table 5 in Deaton (2003), scaled by 1.1247165 for urban price increases from 2000 to 2003. The last column is the 2000 Deaton-adjusted line scaled by 1.41578 (see Appendix).

Figure 4: All-India poverty scorecard

Indicator		Values					Points
1.	How many children aged 0 to 17 are in the household?	75 0	4 8	3 11	2 17	1 22	Zero 31
2.	What is the household's primary energy source for cooking?	Firewood and chips, charcoal, or none			Any other fuel		
		0			8		
3.	Does the household own a television?				No 0	Yes 4	
4.	How many hectares of land does the household own?	Urban, any amount		Rural, 0 to 0.4	Rural, 0.41 to 2	Rural, >2	
		0		4	7	10	
5.	What is the principal occupation of the household?	Agricultural labourers	Operators and labourers, bricklayers, construction workers	Cultivators, farmers, fishers, hunters, loggers, unknown	Sales workers, service workers, transport-equipment operators	Professional, technical, clerical, administrative, managerial, executive, teachers	
		0	6	8	11	13	
6.	How many almira/h/dressing tables does the household own?	None			One	Two or more	
		0			2	9	
7.	Is the residence all pucca (burnt bricks, stone, cement, concrete, jackboard/cement-plastered reeds, timber, tiles, galvanised tin or asbestos cement sheets)?				No 0	Yes 5	
8.	Does the household own a pressure cooker or pressure pan?				No 0	Yes 5	
9.	Does the household own a sewing machine?				No 0	Yes 6	
10.	How many electric fans does the household own?	None		One or two	Three or more		
		0		5	10		
Source: Calculations based on Schedule 1.0 of the 59th Round (2003) of India's Socio-Economic Survey (NSSO, 2005).							Total:

Figure 5: Urban India poverty scorecard

Indicator		Values					Points	
1.	How many children aged 0 to 17 are in the household?	≥5	4	3	2	1	Zero	
		0	7	9	17	20	28	
2.	What is the household's primary energy source for cooking?				Firewood and chips, or charcoal	Not firewood and chips, charcoal, nor LPG	LPG	
					0	6	9	
3.	How many electric fans does the household own?				None	One or two	3 or more	
					0	5	10	
4.	Does the household hire its residence?					No	Yes	
						0	9	
5.	Does the household own a foam or rubber cushion (dunlopillo type)?					No	Yes	
						0	10	
6.	What is the principal occupation of the household?				Others or unknown	Executive, managerial, administrative, and teachers	Professional and technical workers	
					0	3	7	
7.	Does the household own a scooter or motorcycle?					No	Yes	
						0	9	
8.	Does the household own a refrigerator?					No	Yes	
						0	14	
9.	Does the household own a television?					No	Yes	
						0	2	
10.	Does the household own a tape recorder or CD player?					No	Yes	
						0	3	
							Total:	

Source: Calculations based on Schedule 1.0 of the 59th Round (2003) of India's Socio-Economic Survey (NSSO, 2005).

Figure 6: Rural India poverty scorecard

Indicator		Values					Points	
1.	How many children aged 0 to 17 are in the household?	≥5	4	3	2	1	Zero	
		0	7	10	16	21	29	
2.	How many electric fans does the household own?				None	One or two	Three or more	
					0	4	10	
3.	What is the principal occupation of the household?		Agricultural labourers		Cultivators, farmers, fishers, hunters, loggers, sales workers, service workers, operators and labourers, bricklayers, construction workers, transport-equipment operators, and unknown		Professional, technical, clerical, administrative, managerial, executive, teachers	
			0		9		13	
4.	Does the household own a pressure cooker or pressure pan?					No	Yes	
						0	7	
5.	How many almirah/dressing tables does the household own?				None	One	Two or more	
					0	3	9	
6.	Is the residence all pucca (burnt bricks, stone, cement, concrete, jackboard/cement-plastered reeds, timber, tiles, galvanised tin or asbestos cement sheets)?					No	Yes	
						0	6	
7.	Does the household own a sewing machine?					No	Yes	
						0	9	
8.	What is the household's primary energy source for cooking?					Not LPG	LPG	
						0	9	
9.	Does the household own a television?					No	Yes	
						0	5	
10.	How many hectares of land does the household own?				1 or less	1.01 to 2	More than 2	
					0	2	5	
							Total:	

Source: Calculations based on Schedule 1.0 of the 59th Round (2003) of India's Socio-Economic Survey (NSSO, 2005).

Figure 7: Scores and poverty likelihoods, All-India

Score	Poverty likelihood for people with score in range (%)	% of people <=score who are poor	% of people >score who are non-poor
0-4	88.5	88.5	61.8
5-9	96.3	91.8	62.2
10-14	81.7	85.7	63.4
15-19	82.0	83.5	66.8
20-24	73.6	78.7	71.9
25-29	60.6	72.5	77.1
30-34	52.8	67.9	82.1
35-39	39.0	62.2	86.5
40-44	36.3	58.3	92.0
45-49	22.1	54.0	95.6
50-54	5.2	49.5	95.9
55-59	6.6	46.1	96.8
60-64	6.8	44.1	98.1
65-69	3.6	42.1	99.0
70-74	2.8	40.8	100.0
75-79	0.0	39.8	100.0
80-84	0.0	39.1	100.0
85-89	0.0	38.8	100.0
90-94	0.0	38.7	100.0
95-100	0.0	38.7	0.0

Surveyed cases weighted to represent all India

Source: Calculations based on Schedule 1.0 of the 59th Round (2003) of India's Socio-Economic Survey (NSSO, 2005).

Figure 8: Scores and poverty likelihoods, urban

Score	Poverty likelihood for people with score in range (%)	% of people <=score who are poor	% of people >score who are non-poor
0-4	98.6	98.6	83.0
5-9	75.0	80.7	85.3
10-14	32.0	56.4	86.2
15-19	55.2	56.0	88.9
20-24	50.2	54.1	92.7
25-29	26.0	45.3	95.6
30-34	12.0	38.1	96.9
35-39	8.4	33.1	97.8
40-44	1.1	28.6	97.5
45-49	11.4	26.6	99.6
50-54	1.5	24.6	99.9
55-59	0.4	22.7	100.0
60-64	0.0	21.1	100.0
65-69	0.0	20.0	100.0
70-74	0.0	19.2	100.0
75-79	0.0	18.6	100.0
80-84	0.0	18.3	100.0
85-89	0.0	18.1	100.0
90-94	0.0	18.0	100.0
95-100	0.0	18.0	100.0

Surveyed cases weighted to represent urban India

Source: Calculations based on Schedule 1.0 of the 59th Round (2003) of India's Socio-Economic Survey (NSSO, 2005).

Figure 9: Scores and poverty likelihoods, rural

Score	Poverty likelihood for people with score in range (%)	% of people <=score who are poor	% of people >score who are non-poor
0-4	85.5	85.5	54.9
5-9	88.3	87.6	57.8
10-14	74.3	81.9	60.2
15-19	74.2	78.2	66.7
20-24	57.4	72.4	70.9
25-29	60.3	69.3	79.4
30-34	32.9	62.8	83.2
35-39	23.4	58.1	85.3
40-44	32.4	55.9	90.8
45-49	25.7	54.1	96.0
50-54	6.6	51.8	96.8
55-59	3.5	50.1	96.9
60-64	2.2	48.8	96.6
65-69	5.2	47.9	97.4
70-74	6.4	47.1	100.0
75-79	0.0	46.6	100.0
80-84	0.0	46.2	100.0
85-89	0.0	46.0	100.0
90-94	0.0	45.9	100.0
95-100	0.0	45.9	100.0

Surveyed cases weighted to represent rural India

Source: Calculations based on Schedule 1.0 of the 59th Round (2003) of India's Socio-Economic Survey (NSSO, 2005).

Figure 10: Confidence intervals for estimated poverty likelihoods, All-India scorecard applied to all of India

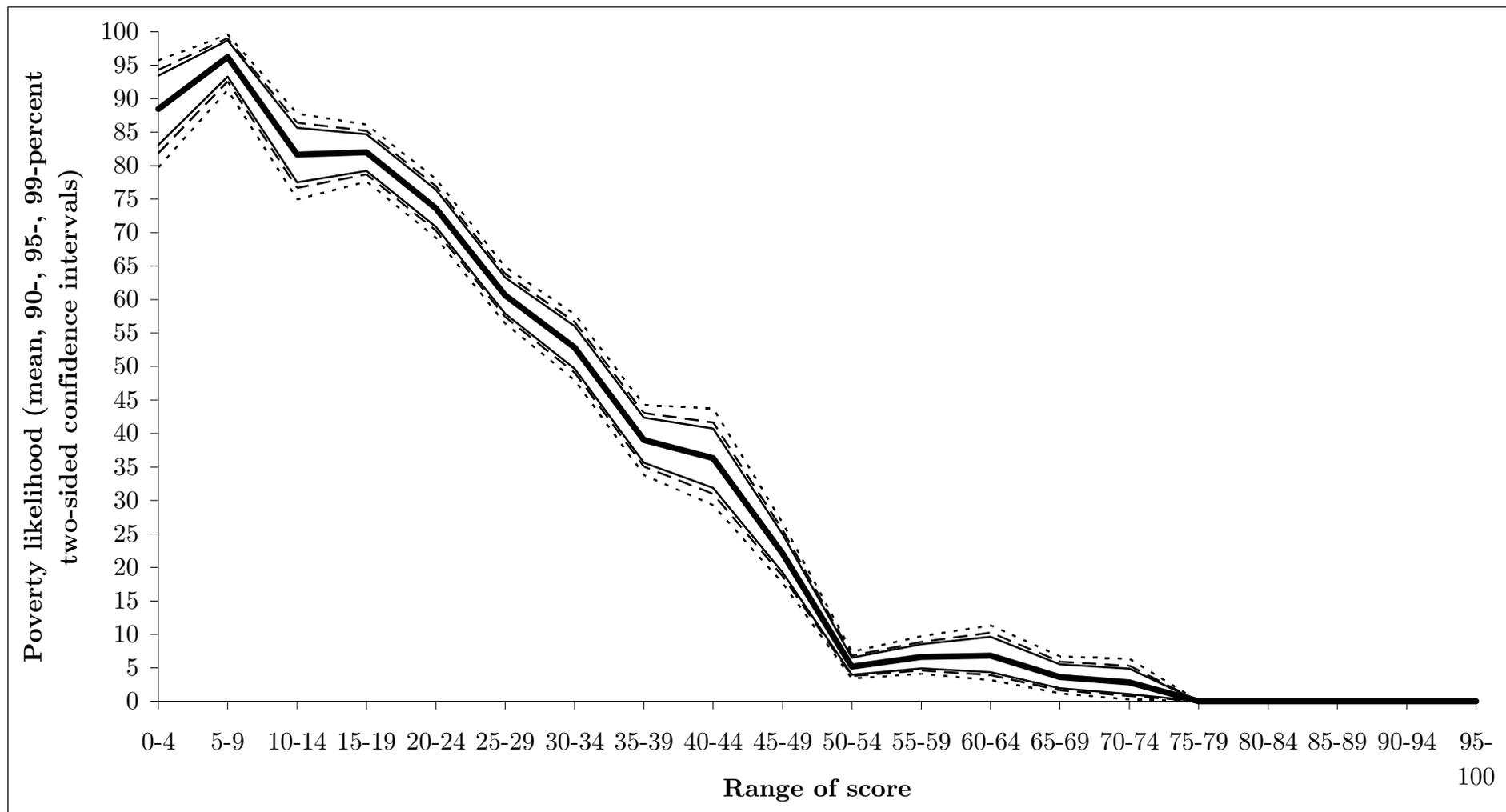


Figure 11: Confidence intervals for estimated poverty likelihoods, urban scorecard applied to urban areas

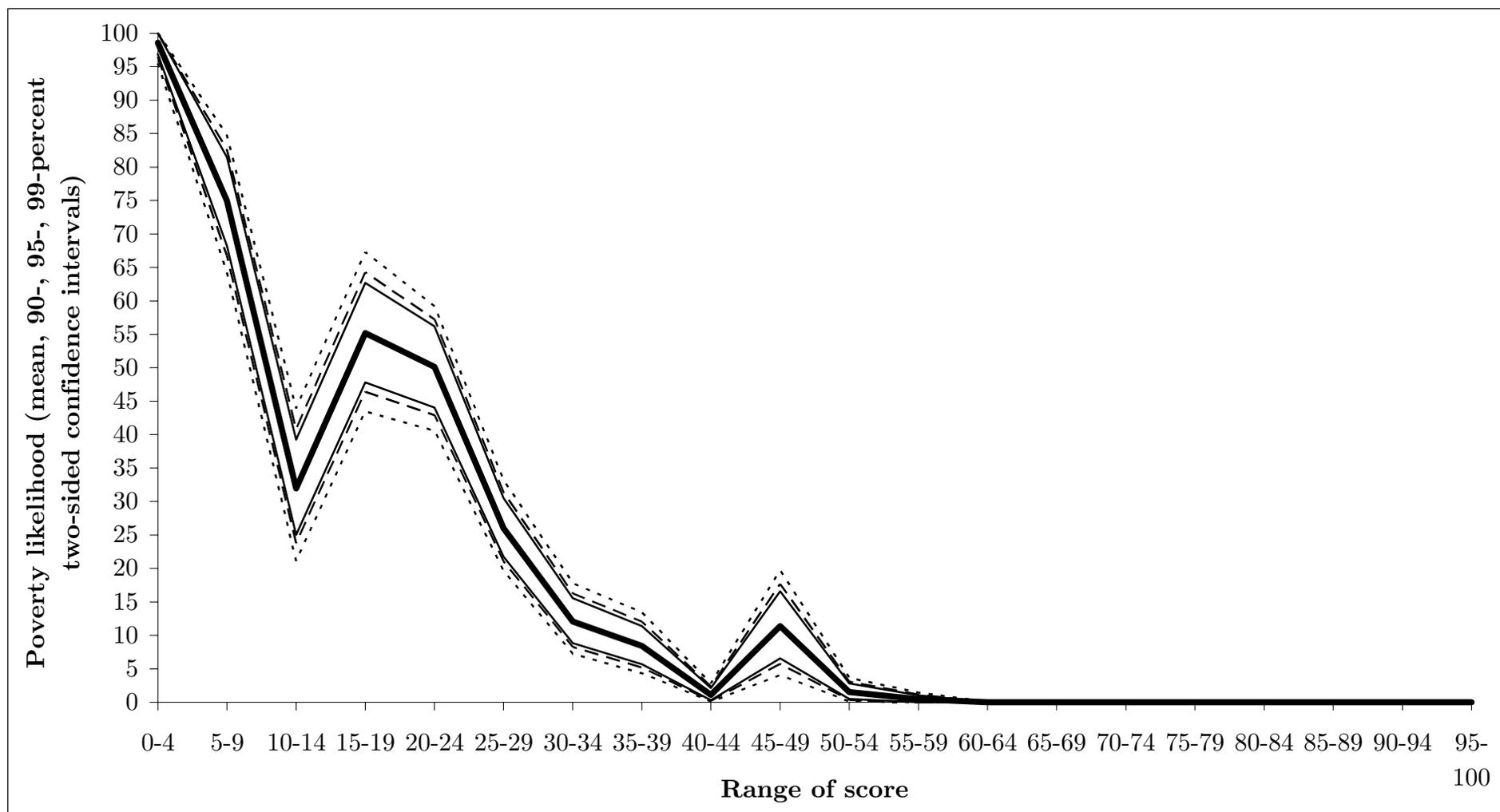


Figure 12: Confidence intervals for estimated poverty likelihoods, rural scorecard applied to rural areas

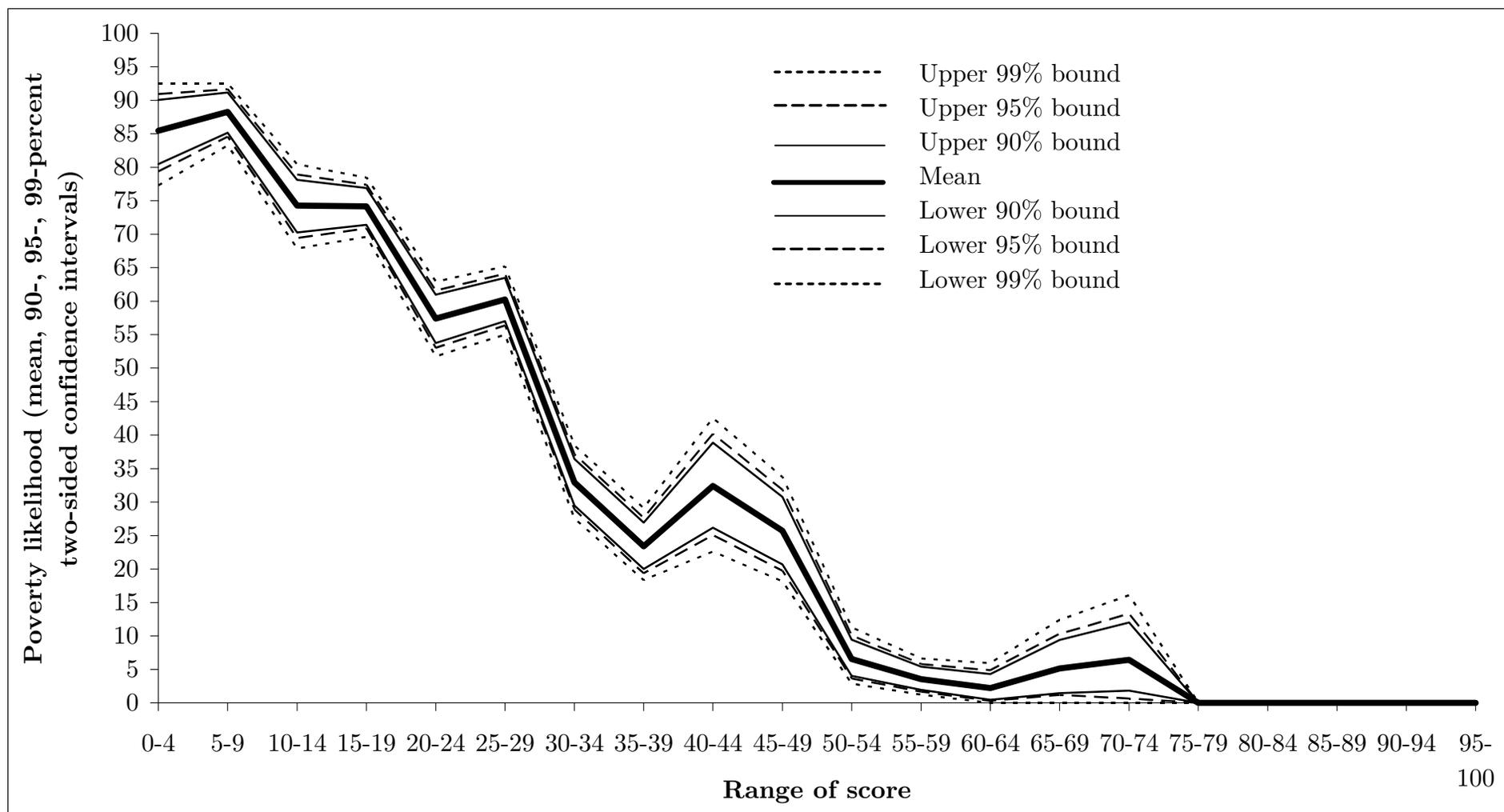


Figure 13: Population-weighted average confidence intervals for estimated poverty likelihoods associated with scores

Scorecard segment	All-India	Urban	Rural
	applied to	applied to	applied to
Application segment	All India	Urban	Rural
90-percent	2.7	3.1	3.5
95-percent	3.2	3.6	4.1
99-percent	4.2	4.8	5.5

Source: 59th Round (2003) of India's Socio-Economic Survey (NSSO, 2005).

Figure 14: Differences between predicted and actual poverty likelihoods, All-India scorecard applied to all India

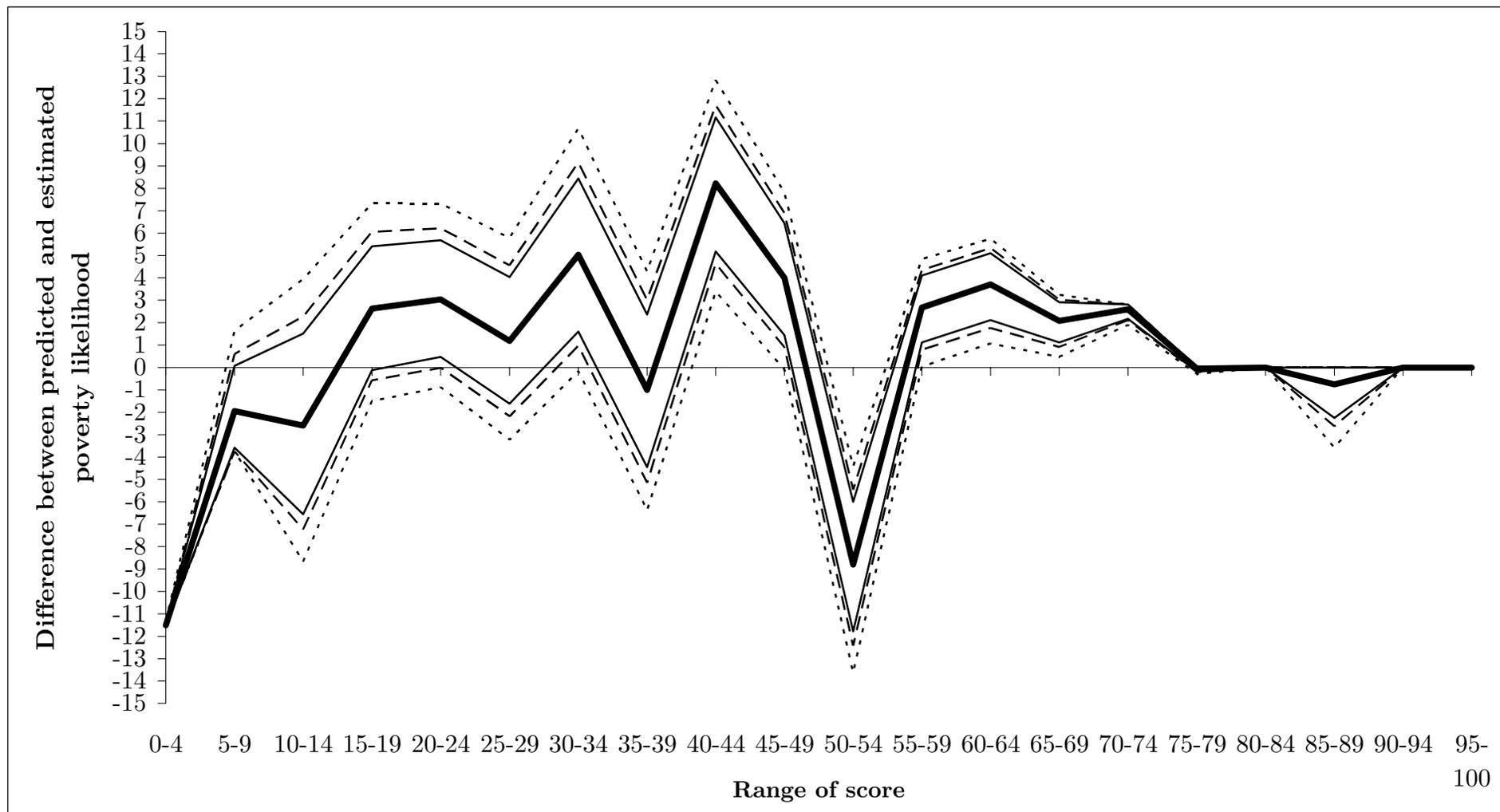


Figure 15: Differences between predicted and actual poverty likelihoods, urban scorecard applied to urban areas

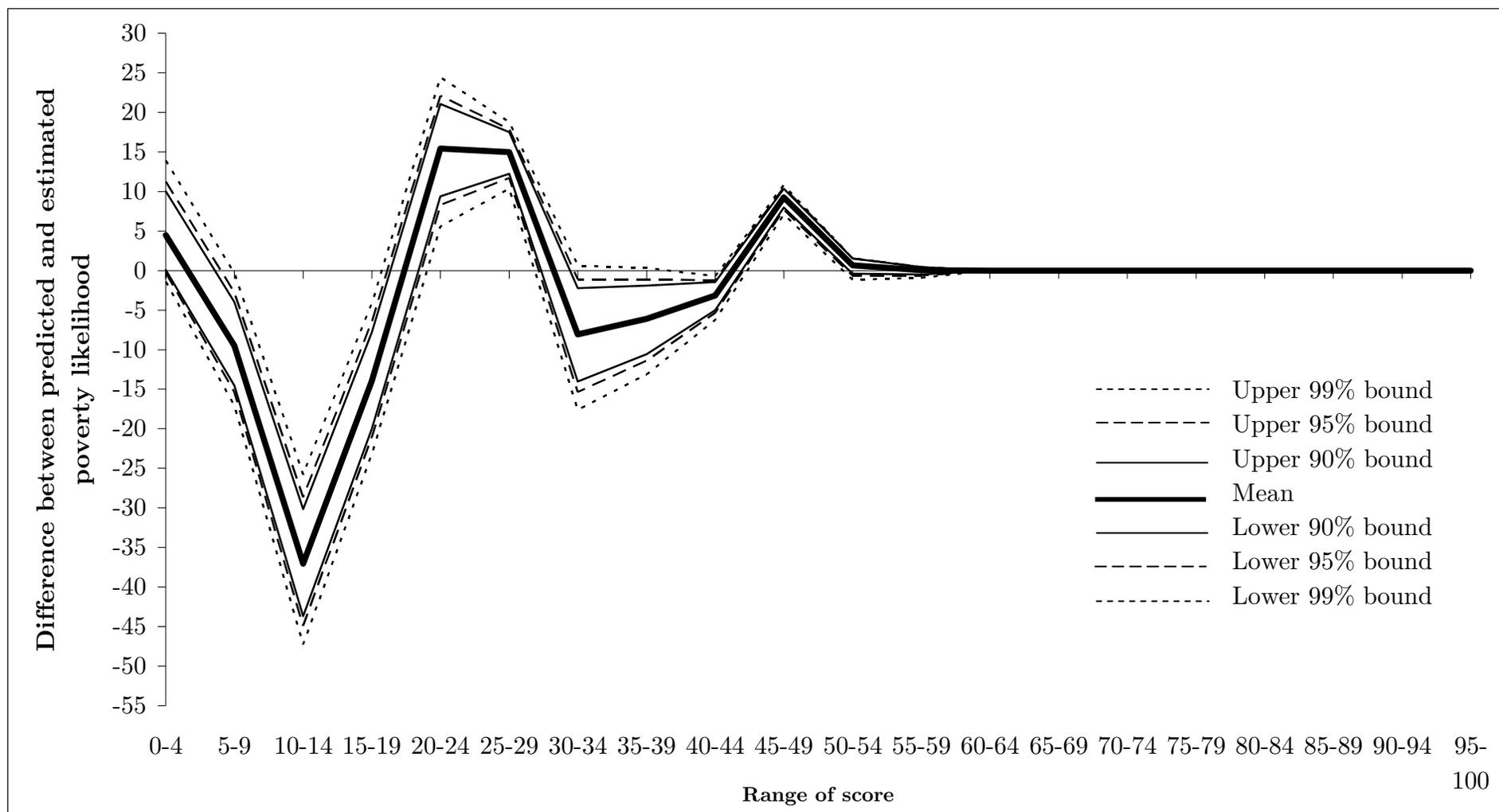


Figure 16: Differences between predicted and actual poverty likelihoods, rural scorecard applied to rural areas

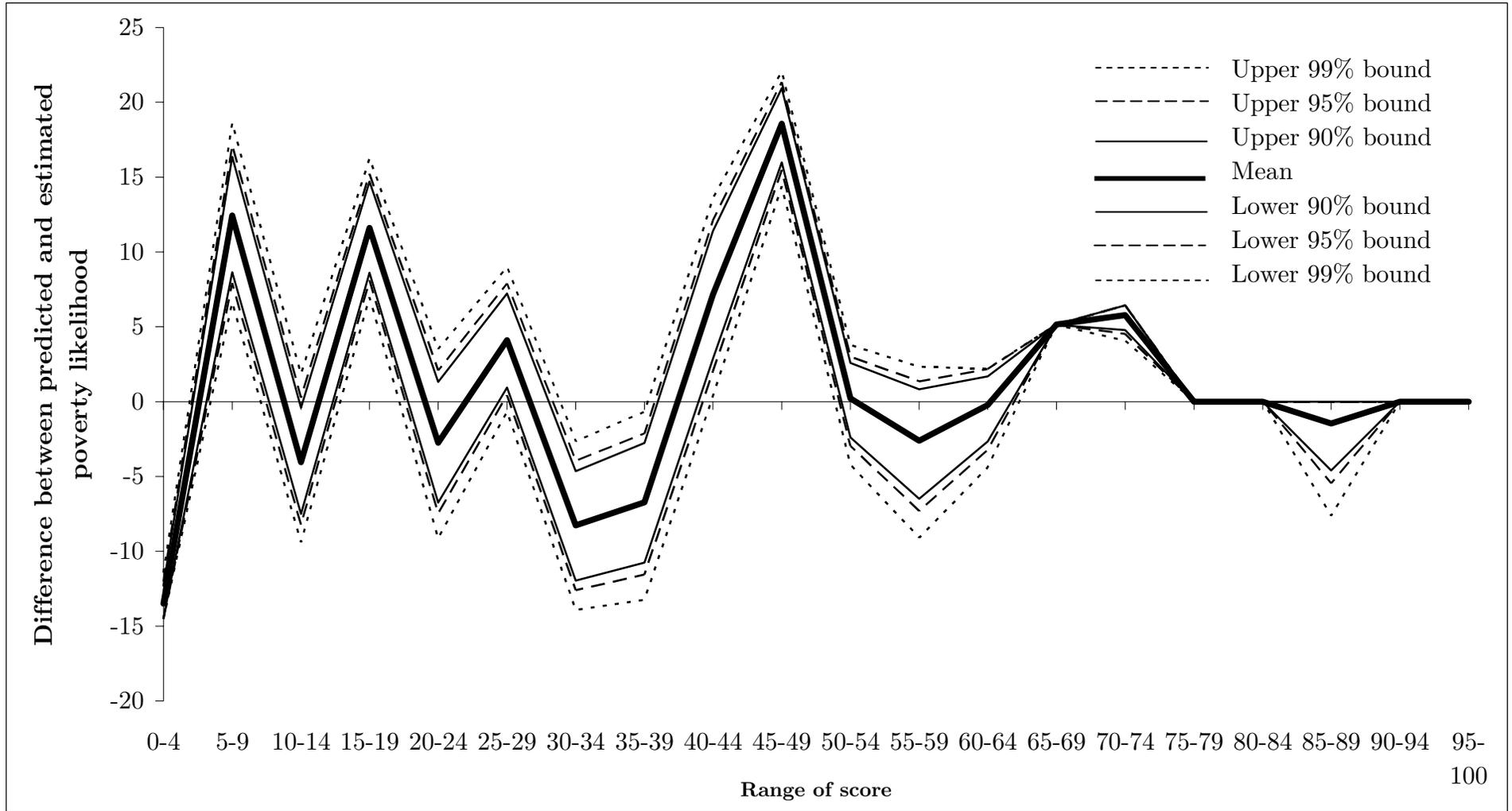


Figure 17: Population-weighted average absolute differences (and confidence intervals) between predicted and actual poverty likelihoods

Scorecard segment	All-India applied to	Urban applied to	Rural applied to	All-India applied to	All-India applied to
Application segment	All India	Urban	Rural	Urban	Rural
Absolute differences (percentage points)					
Mean	3.6	8.0	6.8	8.3	4.0
Confidence intervals around mean absolute difference (+/- percentage points)					
90-percent	2.6	2.9	3.1	3.4	3.2
95-percent	3.1	3.4	3.7	4.0	3.8
99-percent	4.1	4.5	4.9	5.3	5.0

Source: 59th Round (2003) of India's Socio-Economic Survey (NSSO, 2005).

Figure 18: Differences between predicted and actual poverty likelihoods, All-India scorecard applied to urban areas

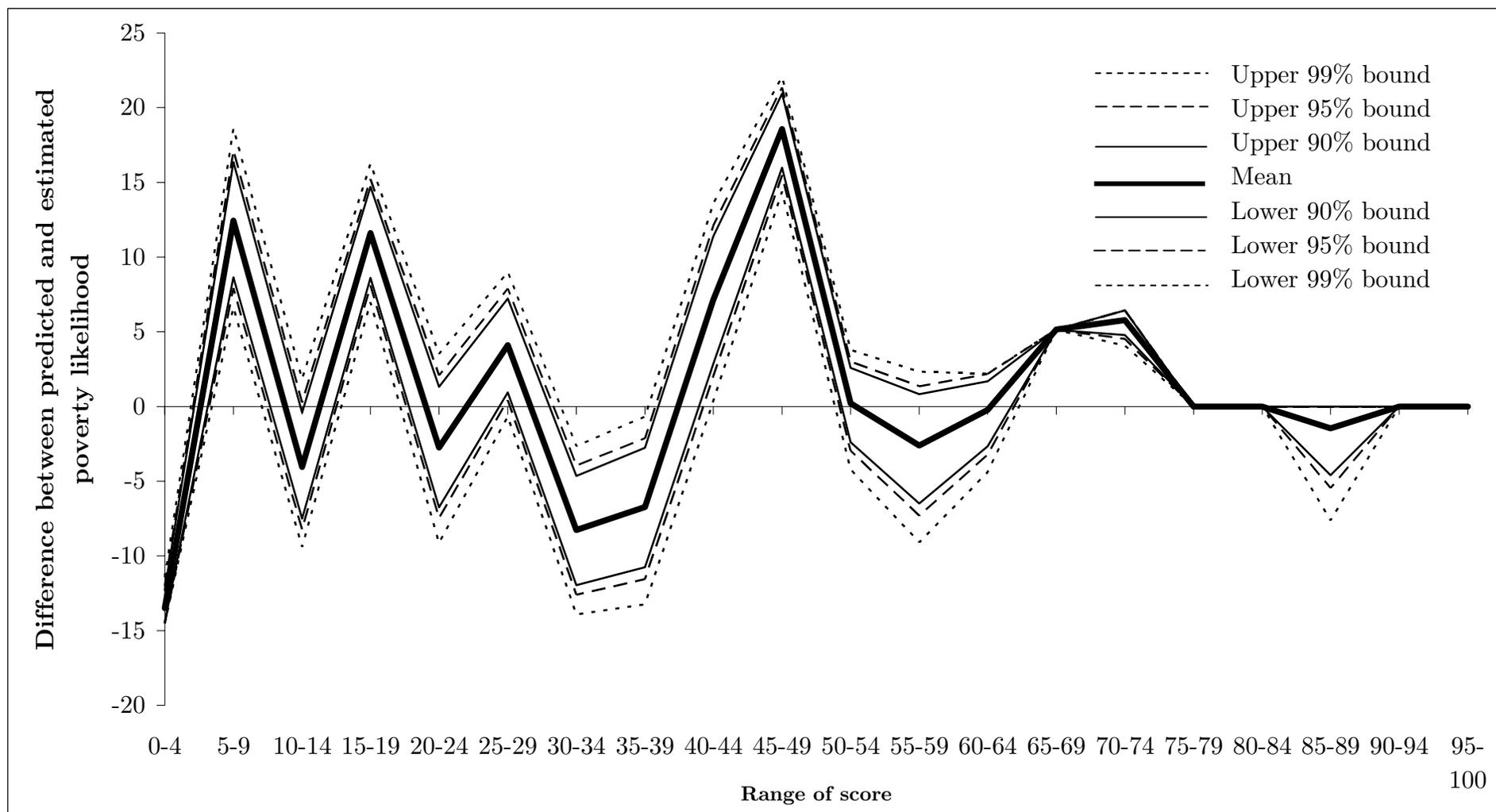


Figure 19: Differences between predicted and actual poverty likelihoods, All-India scorecard applied to rural areas

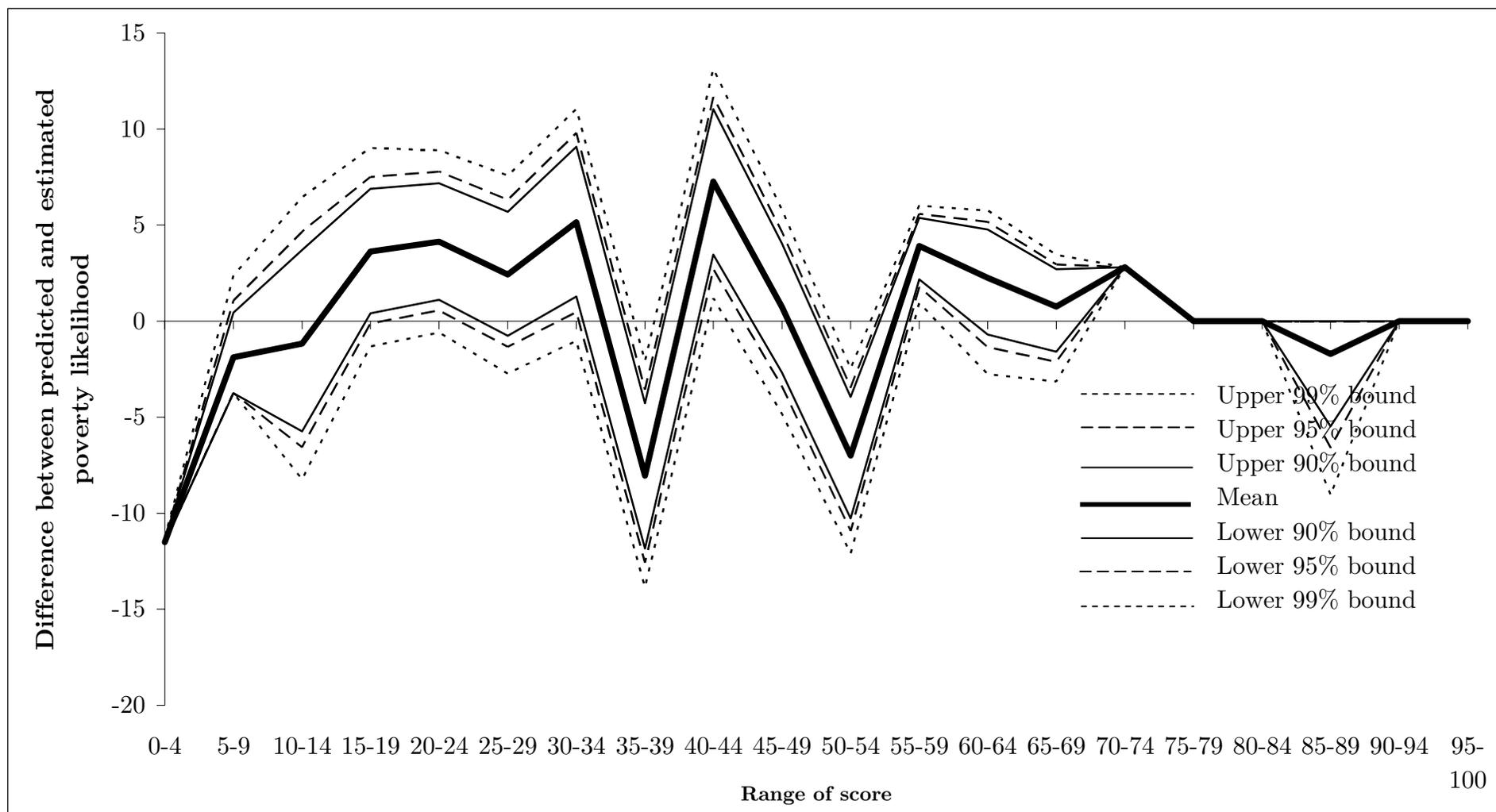


Figure 20: Distribution of differences between estimated and true overall poverty rates

Scorecard segment	All-India applied to	Urban applied to	Rural applied to	All-India applied to	All-India applied to
Application segment	All India	Urban	Rural	Urban	Rural
<u>Distribution of differences (percentage points)</u>					
Mean	2.00	-1.47	3.14	2.05	1.99
Standard deviation	0.39	0.53	0.52	0.51	0.52
<u>Confidence intervals (+/- percentage points)</u>					
90-percent	0.7	0.9	0.9	0.8	0.9
95-percent	0.8	1.0	1.0	1.0	1.0
99-percent	1.0	1.4	1.3	1.3	1.4
Source: 59th Round (2003) of India's Socio-Economic Survey (NSSO, 2005).					

Figure 21: Ability to rank-order households by poverty status, All-India scorecard applied to urban, versus urban scorecard applied to urban

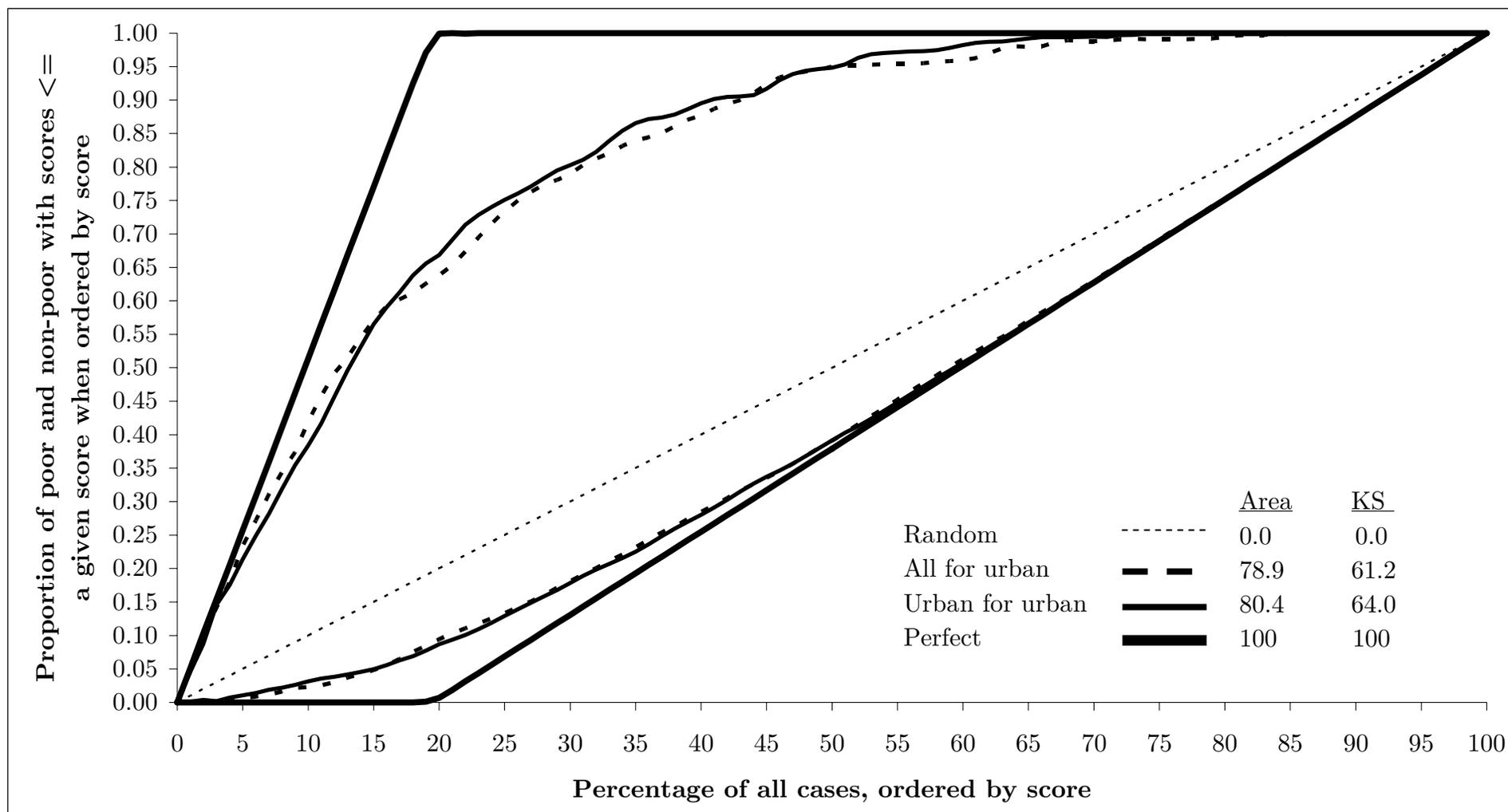


Figure 22: Ability to rank-order households by poverty status, All-India scorecard applied to rural, versus rural scorecard applied to rural

