Poverty Assessment Tool Accuracy Submission
USAID/IRIS Tool for Serbia
Submitted: June 27, 2008
Updated: February 15, 2013 (text clarification; added decimal values to coefficients)

The following report is divided into six sections. Section 1 provides a general overview of the tool development process. Section 2 describes the data set used to create Poverty Assessment Tool for Serbia. Section 3 details the set of statistical procedures used for selecting indicators and for estimating household expenditure or, for some models, the probability that a household is very poor. Section 4 reports on the in-sample accuracy of each prediction model considered. Sections 5 and 6 explain how regression coefficients are used in poverty prediction and how these predictions are used to classify households into the "very poor" and "not very poor" categories.

Annex 1 to this report provides accuracy results for an additional poverty line beyond that required by the Congressional legislation. Annex 2 supplies a careful consideration of out-of-sample accuracy for Serbia Poverty Assessment Tool.

1. Overall approach to the tool development

The approach used to develop the poverty assessment tool for Serbia is built on the lessons learned and methods refined during the original USAID/IRIS project, Developing Poverty Assessment Tools (September 2003 to October 2006). In the initial phase of the project, the IRIS Center analyzed data from existing national household surveys and from surveys it conducted itself in countries where survey data were not available. The aim was to identify household indicators most closely associated with a household being "very poor" in terms of per-capita expenditures or income. IRIS used statistical methods to identify the 15 indicators that most closely track the per-capita expenditures or income of each household, as revealed by the household survey data. In addition, IRIS compared the performance of 8 different statistical methods in quantifying the statistical links between these 15 indicators and household expenditures/income; the accuracy of each method was assessed using criteria developed especially for this project. In this manner, IRIS identified the best-performing set of indicators (with associated weights) and statistical method for identifying the poverty status of households in each country. Statistical testing for accuracy was carried out for twelve countries in total.

In addition, the indicators that appeared among the "best 15" in at least one of the twelve countries were included in the next part of the project: testing for practicality. USAID selected seventeen microenterprise organizations to conduct field tests of practicality. Each question was rated as to whether the respondent found it to be sensitive, difficult, or that it was perceived that she falsified her answer. The lessons learned from the practicality testing were used to remove impractical indicators from consideration for the final poverty assessment tools.

The end result of this development process was a country-specific poverty assessment tool that estimates—rather than directly measures—household per capita consumption

expenditure (or income) or the probability that a household is very poor based on a short set of practical indicators. Each country tool is incorporated into a data entry template that allows a microenterprise practitioner to easily enter and store the responses of its sampled clients to indicator questions and will also estimate the percentage of that practitioner's client households who are very poor.

In October 2006, USAID contracted the IRIS Center to build on the statistical methods and practicality information generated during the original project to develop poverty assessment tools for use by microenterprise practitioners in additional countries such as Serbia. As part of this new phase, IRIS will explore the use of existing household survey data sets beyond specifically the LSMS to develop new poverty assessment tools, refine the tool development and testing methodology, and strive to make the tools even simpler and easier to implement. IRIS will also provide a Help Desk to assist practitioners with the implementation of approved USAID poverty assessment tools.

2. Data source

For Serbia, existing data from the 2007 LSMS were used to construct the poverty assessment tool. The full sample of 5,557 households is nationally representative. The sample used for tool construction is comprised of a randomly selected 2,779 households (50 percent of the full sample). The remainder, another randomly selected 2,778 households, is reserved for out-of-sample accuracy testing, which investigates the robustness of in-sample poverty estimation.

3. Process used to select included indicators

Suitable household surveys, such as the LSMS, typically include variables related to education, housing characteristics, consumer durables, agricultural assets, illness and disability, and employment. For Serbia, more than 90 indicators from all categories were considered.

The MAXR procedure in SAS was used to select the best poverty indicators (for variables found to be practical) from the pool of potential indicators in an automated manner. MAXR is commonly used to narrow a large pool of possible indicators into a more limited, yet statistically powerful set of indicators. The MAXR technique seeks to maximize explained variance (i.e., R^2) by adding one variable at a time (per step) to the regression model, and then considering all combinations among pairs of regressors to move from one step to the next. Thus, the MAXR technique allows us to identify the best model containing 15 variables (not including control variables for household size, age of the household head, and location).

The MAXR procedure yielded the best 15 variables for the OLS model (also used for the Quantile model) and another set of best 15 variables for the Linear Probability model (also used for the Probit model). The final set of indicators and their weights, therefore depended on selecting one of these four statistical models—OLS, Quantile, Linear

Probability, or Probit—as the best model.[1] This selection of the best model was based on the Balance Poverty Accuracy Criterion (BPAC) and the Poverty Incidence Error (PIE), along with practicality considerations.[2]

4. Estimation methods used to identify final indicators and their weights/coefficients

As explained more fully in Section 6, the line used to construct the poverty tool for Serbia is the "median poverty line" – the level of monthly expenditure that divides the poorest half of those living below the national poverty line from the less-poor half of the officially poor. Table 1 summarizes the accuracy results achieved by each of the eight estimation methods in predicting household poverty relative to this poverty line. For Serbia, the most accurate method, on the basis of BPAC, is the 1-step Quantile regression.

Table 1: In-sample Accuracy Results for Prediction at the Legislative Poverty Line

SERBIA Median line***** Share of "very poor": 3.1%	**Total Accuracy**	**Poverty Accuracy**	**Under-coverage**	**Leakage**	**PIE**	**BPAC**
Single-step methods						
OLS	97.24	17.87	82.13	8.08	-2.26	-56.18
Quantile regression (estimation point: 26)	**96.63**	**45.33**	**54.67**	**55.63**	**0.03**	**44.37**
Linear Probability	97.10	6.18	93.82	1.06	-2.83	-86.59
Probit	97.64	30.47	69.53	7.51	-1.89	-31.55
Two-step methods						
OLS – 11 percentile cutoff	97.66	35.95	64.05	12.37	-1.58	-15.73
Quantile (estimation points: 26, 3) 11 percentile cutoff	96.77	44.35	55.65	49.95	-0.17	38.64
LP – 13 percentile cutoff	97.90	38.65	61.35	7.24	-1.66	-15.46
Probit – 11 percentile cutoff	97.71	38.01	61.99	12.86	-1.50	-11.11
* Median poverty line is 7,551 dinars per adult equivalent per month in June 2007 prices. This poverty line is based on the official national poverty line of 8,883 dinars.						

For Serbia, the functionality of predicting the poverty rate at another poverty line—in this case, the national poverty line—has been added. When running the analysis routine with the Epi Info template, the user is presented the option to predict the extreme poverty rate (using the median line), the poverty rate (national line), or both. The methodology and the accuracy results for this prediction are discussed in Annex 1.

[1] The set of indicators and their weights also depended on the selection of a 1-step or 2-step statistical model.

[2] For a detailed discussion of these accuracy criteria, see "Note on Assessment and Improvement of Tool Accuracy" at www.povertytools.org

5. How coefficients and weights are used to estimate poverty status or household expenditures

For the Quantile regression method, the estimated regression coefficients indicate the weight placed on each of the included indicators in estimating the household expenditures of each household in the sample. These estimated coefficients are shown in Table 3. In constructing the Poverty Assessment Tool for each country, these weights are inserted into the "back-end" analysis program of the EPI template used to calculate the incidence of extreme poverty among each implementing organization's clients. While a skilled EPI user would be able to locate the model's weights in the backend, they would not be seen by the client or the interviewer during the normal course of interviewing, entering the data, or calculating the extreme poverty rate.

6. Decision rule used for classifying households as very poor and not very-poor

The legislation governing the development of USAID tools defines the "very poor" as either the bottom (poorest) 50 percent of those living below the poverty line established by the national government or those living on the local equivalent of less than $1/day (in Purchasing Power Parity, or PPP, terms). The applicable poverty line for USAID tool development is the one that yields the higher household poverty rate for a given country.

In Serbia the applicable threshold is the median poverty line, the household per capita expenditure value of the 50^{th} percentile below the national poverty line of 8,883 dinars per adult equivalent per month, at the level of prices prevailing in June 2007 when the household survey data were collected. At 7,551 per month, the median poverty line identifies 3.1% of households as "very poor". Alternatively, the $1/day poverty line of 1534 dinars yields a household poverty rate of 0.0%. Hence the decision rule for Serbia's USAID poverty assessment tool in classifying the "very poor" (and the "not very-poor") is whether that predicted per capita monthly expenditures of a household fall below (or above) the median poverty line.

Note that the national poverty line reported above (8,883 dinar) is expressed in the prevailing prices at the time of data collection June 2007. The poverty line and other monetary variables used in the accuracy results (and thus, in the poverty calculation), are inflation-adjusted to the most current time period possible (May 2008). This yields a (national) poverty line of 10,246 dinars per adult equivalent per month and an extreme (median national) poverty line of 8,709 dinars per adult equivalent per month.

Because the selected tool is based on a Quantile model, each household whose estimated per capita consumption expenditures according to the tool fall below the median poverty line is identified as "very poor," and each household whose estimated per capita consumption expenditures exceeds the median poverty line is identified as "not very-poor."

An additional requirement for using the median poverty line is that the national poverty line on which it depends is actively used by the local government. This is indeed the case in Serbia, where the government uses the national poverty line for poverty monitoring.[3]

Table 2 below compares the poverty status of the sample households as identified by the selected model, versus their true poverty status as revealed by the data from the benchmark household survey (in-sample test). The upper-left and lower-right cells show the number of households correctly identified as "very poor" or "not very-poor," respectively. Meanwhile, the upper-right and lower-left cells indicate the twin errors possible in poverty assessment: misclassifying very poor households as not very-poor; and the opposite, misclassifying not very-poor households as very poor.

Table 2: Poverty Status of Sample Households, as Estimated by Model and Revealed by the Benchmark Survey

	Number of households identified as very poor by the tool	**Number of households identified as not very-poor by the tool**
Number of "true" very poor households (as determined by benchmark survey)	39 (1.4%)	46 (1.7%)
Number of "true" not very-poor households (as determined by benchmark survey)	47 (1.7%)	2647 (95.2%)

[3] http://webrzs.statserb.sr.gov.yu/axd/en/index.php

Table 3: Regression Estimates using 1-step Quantile Method for Prediction at the Median Poverty Line

SERBIA 1 STEP MAXR/QUANT: variables from MAXR/OLS 100 percentile model
Regression results, estimation point of 26 percentile

.26 Quantile regression — Number of obs = 2779
Raw sum of deviations 933.7058 (about 9.677865)
Min sum of deviations 633.1381 — Pseudo R2 = 0.3219

Variable	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
Intercept	9.586478	0.081	118.360	0.000	9.428	9.745
Household size	-0.1493846	0.014	-10.610	0.000	-0.177	-0.122
Household size squared	0.0057735	0.002	3.260	0.001	0.002	0.009
Household head age	-0.0021213	0.003	-0.750	0.452	-0.008	0.003
Household head age squared	-0.0000327	0.000	-1.320	0.187	0.000	0.000
Household lives in Belgrade	0.0750103	0.020	3.800	0.000	0.036	0.114
Household lives in West Serbia	0.0302452	0.019	1.560	0.119	-0.008	0.068
Household lives in Central Serbia	0.0258658	0.018	1.430	0.152	-0.010	0.061
Household lives in East Serbia	-0.0432627	0.021	-2.070	0.039	-0.084	-0.002
Household lives in South-East Serbia	-0.0661076	0.019	-3.420	0.001	-0.104	-0.028
Household lives in rural area	-0.0027688	0.015	-0.190	0.849	-0.031	0.026
Share of household members who have never attended school	-0.3242225	0.069	-4.730	0.000	-0.459	-0.190
Household head has completed a post-secondary, non-university program	0.1658878	0.025	6.600	0.000	0.117	0.215
Household head has completed a university degree	0.1987236	0.024	8.300	0.000	0.152	0.246
Number of rooms in the dwelling	0.0476407	0.005	9.040	0.000	0.037	0.058
Household has sewerage	0.1974638	0.024	8.300	0.000	0.151	0.244
Household has central/autonomous heating	0.0741951	0.017	4.410	0.000	0.041	0.107
Household owns one or more landline telephone	0.137356	0.018	7.540	0.000	0.102	0.173
Household uses solid fuel, such as firewood or coal, for one of their sources of heating	-0.0360967	0.017	-2.080	0.037	-0.070	-0.002
Household owns one or more air conditioners	0.0944093	0.022	4.280	0.000	0.051	0.138
Household owns one or more dishwashers	0.1127693	0.026	4.280	0.000	0.061	0.164
Household owns one or more vacuums	0.0777192	0.021	3.760	0.000	0.037	0.118
Household owns one or more CD/DVD players	0.1308288	0.015	8.830	0.000	0.102	0.160
Household owns one or more radios	0.0535486	0.012	4.310	0.000	0.029	0.078
Household owns one or more PC/laptop computers	0.1730987	0.016	10.550	0.000	0.141	0.205
Household owns one or more cars	0.2331569	0.014	16.840	0.000	0.206	0.260

Annex 1: Poverty Prediction at the National Poverty Line

Strictly construed, the legislation behind the USAID poverty assessment tools concerns "very poor" and "not very-poor" beneficiaries. Nevertheless, the intended outcome of the legislation is to provide USAID and its implementing partners with poverty measurement tools that they will find useful.

After discussions among USAID, IRIS, and other members of the microenterprise community, a consensus emerged that the tools would benefit from predictive capacity beyond legislatively-defined extreme poverty. To that end, on agreement with USAID, IRIS has used the best indicators and regression type for predicting the "very poor" to also identify the "poor." For $1/day PPP models, this will be the $2/day PPP; for median poverty models, the "poor" threshold will be the national poverty line. Following this logic, then, the "poor" ("not poor") in Serbia are defined as those whose predicted incomes fall below (above) the national poverty line.

Table 4 summarizes the predictive accuracy results for the national poverty line using the Quantile model specification from the median poverty line. The indicators are the same as those in the model for the median line, but the percentile of estimation and the coefficients of the model were allowed to change (compare Tables 3 and 6). This methodology allows the content and length of the questionnaire to remain the same, but permits greater accuracy in predicting at the national poverty line.

Table 4: Accuracy Results Obtained for Prediction at the National Poverty Line

Serbia National Line Share of Poor: 6.2%	**Total Accuracy**	**Poverty Accuracy**	**Under-coverage**	**Leakage**	**PIE**	**BPAC**
Single-step methods						
Quantile regression (estimation point: 30)	93.23	45.12	54.88	53.91	-0.06	44.15

Table 5 below compares the poverty status of the sample households as identified by the selected model, versus their true poverty status as revealed by the data from the benchmark household survey (in-sample test). The upper-left and lower-right cells show the number of households correctly identified as "poor" or "not poor," respectively. Meanwhile, the upper-right and lower-left cells indicate the twin errors possible in poverty assessment: misclassifying poor households as not poor; and the opposite, misclassifying not poor households as poor.

Table 5: Poverty Status of Sample Households, as Estimated by Model and Revealed by the Benchmark Survey, at National Poverty Line

	Number of households identified as poor by the tool	Number of households identified as not poor by the tool
Number of "true" poor households (as determined by benchmark survey)	78 (2.8%)	93 (3.3%)
Number of "true" not poor households (as determined by benchmark survey)	95 (3.4%)	2513 (90.5%)

Table 6: Regression Estimates using 1-step Quantile Method for Prediction at the National Poverty Line

Regression results, estimation point of 30 percentile

.3 Quantile regression — Number of obs = 2779
Raw sum of deviations 1003.837 (about 9.736639)
Min sum of deviations 682.8626 — Pseudo R2 = 0.3197

Variable	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
Intercept	9.680543	0.119	81.520	0.000	9.448	9.913
Household size	-0.17198	0.021	-8.030	0.000	-0.214	-0.130
Household size squared	0.009378	0.003	3.440	0.001	0.004	0.015
Household head age	-0.00385	0.004	-0.920	0.356	-0.012	0.004
Household head age squared	-1.9E-05	0.000	-0.530	0.598	0.000	0.000
Household lives in Belgrade	0.085875	0.029	2.980	0.003	0.030	0.142
Household lives in West Serbia	0.026143	0.028	0.920	0.356	-0.029	0.082
Household lives in Central Serbia	0.041191	0.026	1.560	0.119	-0.011	0.093
Household lives in East Serbia	-0.02662	0.031	-0.870	0.383	-0.087	0.033
Household lives in South-East Serbia	-0.05871	0.029	-2.060	0.039	-0.115	-0.003
Household lives in rural area	0.01336	0.021	0.630	0.529	-0.028	0.055
Share of household members who have never attended school	-0.33768	0.098	-3.450	0.001	-0.530	-0.146
Household head has completed a post-secondary, non-university program	0.149948	0.038	3.910	0.000	0.075	0.225
Household head has completed a university degree	0.22699	0.035	6.450	0.000	0.158	0.296
Number of rooms in the dwelling	0.048185	0.007	6.470	0.000	0.034	0.063
Household has sewerage	0.203508	0.035	5.810	0.000	0.135	0.272
Household has central/autonomous heating	0.067787	0.024	2.790	0.005	0.020	0.115
Household owns one or more landline telephone	0.151453	0.027	5.710	0.000	0.100	0.203
Household uses solid fuel, such as firewood or coal, for one of their sources of heating	-0.0428	0.025	-1.690	0.092	-0.093	0.007
Household owns one or more air conditioners	0.096214	0.032	3.010	0.003	0.034	0.159
Household owns one or more dishwashers	0.117177	0.039	3.010	0.003	0.041	0.194
Household owns one or more vacuums	0.069889	0.030	2.310	0.021	0.011	0.129
Household owns one or more CD/DVD players	0.130514	0.021	6.090	0.000	0.089	0.173
Household owns one or more radios	0.053437	0.018	2.930	0.003	0.018	0.089
Household owns one or more PC/laptop computers	0.182178	0.024	7.640	0.000	0.135	0.229
Household owns one or more cars	0.227992	0.020	11.270	0.000	0.188	0.268

Annex 2: Out-of-Sample Accuracy Tests

In statistics, prediction accuracy can be measured in two fundamental ways: with in-sample methods and with out-of-sample methods. In the in-sample method, a single data set is used. This single data set supplies the basis for both model calibration and for the measurement of model accuracy. In the out-of-sample method, at least two data sets are utilized. The first data set is used to calibrate the predictive model. The second data set tests the accuracy of these calibrations in predicting values for previously unobserved cases.

The previous sections of this report provide accuracy results of the first type only. The following section presents accuracy findings of the second type, as both a supplement to certification requirements and as an exploration of the robustness of the best model outside of the 'laboratory' setting.

As noted in section 2, the data set used to construct the Serbia tool was divided randomly into two data sets of equal size (2,779 and 2,778 households). A naïve method for testing out-of-sample accuracy—or for overfitting—is to simply apply the model calibrated on the first data set to the observations contained in the holdout data set. These results are show in Table 7. The best model (1-step Quantile) does not perform very well in terms of BPAC, losing 16 percentage points, but does perform well in predicting the overall extreme poverty rate, with PIE diminishing only marginally.

Table 7: Comparison of In-Sample and Out-of-Sample Accuracy Results

	Total Accuracy	Poverty Accuracy	Under-coverage	Leakage	PIE	BPAC
In-Sample Prediction						
	96.63	45.33	54.67	55.63	0.03	44.37
Out-of-Sample Prediction						
	96.14	34.05	65.95	59.94	-0.18	28.04

Another, more rigorous method for testing the out-of-sample accuracy performance of the tool is to provide confidence intervals for the accuracy measures, derived from 1,000 bootstrapped samples from the holdout sample.[4] Each bootstrapped sample is constructed by drawing observations, with replacement, from the holdout sample. The calibrated model is then applied to each sample to yield poverty predictions; across 1,000 samples, this method provides the sampling distributions for the model's accuracy measures.

Table 8 presents the out-of-sample, bootstrapped confidence intervals for the 1-step Quantile model. The lower bound is defined by the 2.5th percentile of the sample

[4] This method of out-of-sample testing is used by Mark Schreiner for the PPI scorecards as detailed on www.microfinance.com

distribution for each accuracy measure; the upper bound is defined by the 97.5th percentile (note that this method does not assume normality).

As with the naïve result in Table 7, the performance of this model is mixed when judged using an out-of-sample approach. The 95% confidence interval is quite wide for BPAC. On the other hand, the confidence interval is extremely tight for PIE.

Table 8: Bootstrapped Confidence Intervals Computed Empirically from Sampling Distribution Without Normality Assumption

Accuracy Measure	95% Confidence Interval	
	LB	UB
Total Accuracy	95.12	96.55
Poverty Accuracy	26.56	46.42
Undercoverage	53.58	73.44
Leakage	50.73	94.01
PIE	-0.57	0.88
BPAC	4.33	40.43

One potential explanation for the out-of-sample BPAC performance for Serbia is the almost negligible rate of extreme poverty in that country; only 3.1% of the households can be classified as very poor. Predicting any phenomenon that occurs so infrequently in a population is difficult; errors of misclassification—undercoverage and leakage—can be quite pronounced. It is perhaps also worth noting that the other three regression models (OLS, Linear Probability, and Probit) perform very poorly, even in-sample, in Serbia.

However, the primary purpose of the PAT is to assess the overall extreme poverty rate across a group of households. The out-of-sample results for PIE in Table 8 indicate that the extreme poverty rate predicted the PAT is less than one percentage point from the true value in the population. By this measure, the predictive model is quite accurate.