

Poverty Assessment Tool Accuracy Submission
USAID/IRIS Tool for Kosovo
Submitted: September 30, 2008

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 Kosovo. 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 the Kosovo Poverty Assessment Tool.

1. Overall approach to the tool development

The approach used to develop the poverty assessment tool for Kosovo 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 Kosovo. 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 Kosovo, existing data from the 2000 LSMS were used to construct the poverty assessment tool. The full sample of 2,880 households is nationally representative. The sample used for tool construction is comprised of a randomly selected 1,440 households (50 percent of the full sample). The remainder, another randomly selected 1,440 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 Kosovo, more than 85 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 Kosovo 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 Kosovo, the most accurate method, on the basis of BPAC, is the 2-step Quantile regression. However, the 1-step Quantile regression is slightly less accurate (roughly 3 percentage points lower in BPAC and 0.4 percentage point in PIE) and requires only 15 indicators. Following the precedent established in discussions with USAID, the 1-step Quantile was selected as the best model, taking into consideration both accuracy and practicality.

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

KOSOVO Median line* Share of “very poor”: 24.6%	Total Accuracy	Poverty Accuracy	Under-coverage	Leakage	PIE	BPAC
Single-step methods						
OLS	80.11	43.54	56.46	22.76	-8.46	9.84
Quantile regression (estimation point: 43)	79.24	59.42	40.58	42.57	0.49	57.43
Linear Probability	80.59	42.86	57.14	20.17	-9.28	5.89
Probit	80.78	47.20	52.79	23.77	-7.29	18.17
Two-step methods						
OLS – 43 percentile cutoff	79.84	51.21	48.79	31.53	-4.33	33.94
Quantile (estimation points: 43, 20) 43 percentile cutoff	80.36	60.45	39.55	39.11	-0.11	60.01
LP – 43 percentile cutoff	81.59	54.95	45.05	28.29	-4.21	38.19
Probit – 43 percentile cutoff	81.61	54.69	45.31	27.94	-4.36	37.33
* Median poverty line is 75.47 DM per adult equivalent per month in 2000 prices. This poverty line is based on the official national poverty line of 106.54 DM per adult equivalent per month. ³						

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

³ The poverty line reported above is expressed in the prevailing prices at the time of data collection in November 2000.

For Kosovo, 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.

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).⁴ The applicable poverty line for USAID tool development is the one that yields the higher household poverty rate for a given country.

In Kosovo the applicable threshold is the median poverty line, the household per capita expenditure value of the 50th percentile below the national poverty line of 106.54 DM per adult equivalent per month, at the level of prices prevailing in November 2000 when the household survey data were collected. This median poverty line identifies 24.6% of households in the sample as “very poor”.

On the other hand, there is currently no PPP exchange rate for Kosovo. While it is technically possible to construct a PPP line of Kosovo using Serbia’s PPP exchange rate, the resulting value would be an extremely rough proxy for the true value.⁵ In addition, past experience indicates that “\$1/day” lines are in general much lower than national poverty lines in countries in Eastern Europe and Eurasia.

⁴ The congressional legislation specifies the international poverty line as the “equivalent of \$1 per day (as calculated using the purchasing power parity (PPP) exchange rate method).” USAID and IRIS interpret this to mean the international poverty line used by the World Bank to track global progress toward the Millennium Development Goal of cutting the prevalence of extreme poverty in half by 2015. This poverty line has recently been recalculated by the Bank to accompany new, improved estimates of PPP.

⁵ This approximation would depend on the exchange rate between Serbia and Kosovo and also Serbian CPI data (which has a much higher inflation rate in overlapping years than Kosovo).

Hence, in consultation with USAID, the decision rule for Kosovo’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.

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 Kosovo, where the government uses the national poverty line for poverty monitoring.⁶

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)	214 (14.9%)	146 (10.1%)
Number of “true” not very-poor households (as determined by benchmark survey)	153 (10.6%)	927 (64.4%)

⁶ See <http://www.ks.gov.net/ESK/esk/pdf/shqip/sociale/Vlerësimi%20i%20varfërisë%20në%20Kosovë,%20Vëllimi%20I.pdf>

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

Regression results, estimation point of 43 percentile
(sum of wgt is 1.5141e+05)

.43 Quantile regression
Raw sum of deviations 602.7759 (about 4.5731387)
Min sum of deviations 466.176

Number of obs = 1440

Pseudo R2 = 0.2266

Variable	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
Intercept	4.130	0.252	16.390	0.000	3.635	4.624
Household size	0.007	0.007	0.920	0.357	-0.008	0.021
Household size squared	0.000	0.000	-1.230	0.217	0.000	0.000
Household head age	-0.105	0.015	-7.130	0.000	-0.134	-0.076
Household head age squared	0.003	0.001	4.880	0.000	0.002	0.005
Household lives in Pec	-0.040	0.041	-0.970	0.330	-0.122	0.041
Household lives in Pristina	0.146	0.048	3.040	0.002	0.052	0.240
Household lives in Prizren	0.059	0.050	1.190	0.232	-0.038	0.157
Household lives in Kosovska Mitrovica	-0.063	0.049	-1.280	0.201	-0.159	0.033
Household lives in rural area	0.133	0.047	2.850	0.004	0.041	0.224
Share of household members that can write	0.288	0.086	3.340	0.001	0.119	0.457
Household head is literate	0.418	0.180	2.330	0.020	0.065	0.771
Household head can write	-0.374	0.178	-2.100	0.036	-0.724	-0.024
Household has raised one or more milk cows in the last twelve months	0.116	0.043	2.710	0.007	0.032	0.201
Household owns one or more ploughs	-0.091	0.052	-1.730	0.085	-0.194	0.012
Household owns one or more water heaters	0.109	0.057	1.890	0.058	-0.004	0.221
Household owns one or more washing machines	0.152	0.045	3.380	0.001	0.064	0.241
Household owns one or more cameras or video cameras	0.133	0.053	2.500	0.012	0.029	0.238
Household owns one or more satellite dishes	0.144	0.037	3.930	0.000	0.072	0.216
Household owns one or more pieces of jewelry	0.078	0.036	2.170	0.030	0.008	0.149
Household owns one or more cars or trucks	0.381	0.038	10.080	0.000	0.306	0.455
Main source of hot water is wood stove	0.111	0.054	2.060	0.040	0.005	0.217

Main source of hot water is other than centralized supply, electric water heater, coal stove, wood stove, gas water heater	0.019	0.068	0.280	0.781	-0.114	0.152
Household has access to electricity	0.080	0.162	0.490	0.622	-0.238	0.398
Household disposes of garbage in street containers	-0.074	0.045	-1.650	0.099	-0.163	0.014

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. However, the underlying goal 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 median poverty models, the “poor” threshold will be the national poverty line. Following this logic, then, the “poor” (“not poor”) in Kosovo 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

Kosovo National Line Share of Poor: 49.3%	Total Accuracy	Poverty Accuracy	Under-coverage	Leakage	PIE	BPAC
Single-step methods						
Quantile regression (estimation point: 52)	72.96	73.33	26.67	26.81	0.07	73.19

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)	534 (37.1%)	194 (13.5%)
Number of “true” not poor households (as determined by benchmark survey)	195 (13.5%)	517 (35.9%)

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

Regression results, estimation point of 52 percentile
(sum of wgt is 1.5141e+05)

.52 Quantile regression
Raw sum of deviations 611.4075 (about 4.688765)
Min sum of deviations 475.6802

Number of obs = 1440

Pseudo R2 = 0.2220

Variable	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
Intercept	4.439	0.268	16.560	0.000	3.913	4.965
Household size	-0.006	0.008	-0.790	0.429	-0.022	0.009
Household size squared	0.000	0.000	0.460	0.642	0.000	0.000
Household head age	-0.106	0.016	-6.600	0.000	-0.137	-0.074
Household head age squared	0.003	0.001	4.530	0.000	0.002	0.005
Household lives in Pec	0.173	0.051	3.400	0.001	0.073	0.273
Household lives in Pristina	0.122	0.053	2.320	0.021	0.019	0.226
Household lives in Prizren	-0.036	0.052	-0.690	0.488	-0.138	0.066
Household lives in Kosovska Mitrovica	0.164	0.049	3.310	0.001	0.067	0.260
Household lives in rural area	-0.041	0.045	-0.910	0.364	-0.128	0.047
Share of household members that can write	0.303	0.091	3.330	0.001	0.125	0.482
Household head is literate	0.336	0.193	1.740	0.082	-0.042	0.714
Household head can write	-0.252	0.191	-1.320	0.188	-0.627	0.123
Household has raised one or more milk cows in the last twelve months	0.097	0.047	2.070	0.039	0.005	0.188
Household owns one or more ploughs	-0.056	0.056	-1.000	0.315	-0.166	0.054
Household owns one or more water heaters	0.124	0.060	2.070	0.038	0.007	0.242
Household owns one or more washing machines	0.115	0.048	2.390	0.017	0.020	0.209
Household owns one or more cameras or video cameras	0.139	0.057	2.410	0.016	0.026	0.251
Household owns one or more satellite dishes	0.138	0.039	3.520	0.000	0.061	0.215
Household owns one or more pieces of jewelry	0.082	0.038	2.140	0.032	0.007	0.158
Household owns one or more cars or trucks	0.374	0.041	9.220	0.000	0.295	0.454
Main source of hot water is wood stove	0.096	0.057	1.690	0.091	-0.015	0.207
Main source of hot water is other than centralized supply, electric water heater, coal stove, wood stove, gas water heater	0.064	0.072	0.890	0.373	-0.077	0.206
Household has access to electricity	0.165	0.172	0.960	0.338	-0.173	0.503

Household disposes of garbage in street containers	-0.096	0.049	-1.970	0.049	-0.192	-0.001
--	--------	-------	--------	-------	--------	--------

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 Kosovo tool was divided randomly into two data sets of equal size (1,440 and 1,440 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 shown in Table 7. The best model (1-step Quantile) performs moderately well in terms of BPAC, losing less than 10 percentage points. The model performs reasonably well in predicting the overall extreme poverty rate, with PIE worsening by less than 1.5 percentage points.

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

	Total Accuracy	Poverty Accuracy	Under-coverage	Leakage	PIE	BPAC
In-Sample Prediction						
	79.24	59.42	40.58	42.57	0.49	57.43
Out-of-Sample Prediction						
	80.12	55.28	44.72	37.11	-1.85	47.68

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.⁷ 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.

⁷ This method of out-of-sample testing is used by Mark Schreiner for the PPI scorecards as detailed on www.microfinance.com

Table 8 presents the out-of-sample, bootstrapped confidence intervals for the 1-step Quantile model. As with the naïve result in Table 7, the performance of this model is modest, with better results for PIE than for BPAC. The confidence interval around the sample mean of BPAC (47.81) is +/- 13.2 percentage points. For PIE, which measures the difference between the actual poverty rate in the sample and the predicted poverty rate, the confidence interval is fairly tight at +/- 2.6 percentage points. This result is especially relevant, as the primary purpose of the PAT is to assess the overall extreme poverty rate across a group of households.

Table 8: Bootstrapped Confidence Intervals Computed on Assumption of Normality

Accuracy Measure	Mean	Std. Err.	95% Confidence Interval	
			LB	UB
Total Accuracy	80.13	1.15	77.88	82.38
Poverty Accuracy	55.44	2.99	49.57	61.30
Undercoverage	44.56	2.99	38.70	50.43
Leakage	37.30	4.19	29.09	45.50
PIE	-1.80	1.30	-4.35	0.75
BPAC	47.81	6.76	34.57	61.05

The results presented in Table 8 assume a normal distribution for the accuracy measures from the bootstrapped samples. This ignores the possibility that these estimates may have a skewed distribution. Table 9 presents alternative 95% confidence intervals. The lower bound is defined by the 2.5th percentile of the sample distribution for each measure; the upper bound is defined by the 97.5th percentile. On the whole, the results are quite similar between Tables 8 and 9, which adds confidence to the conclusion that the 1-step Quantile model on which the Kosovo PAT is based performs moderately well out-of-sample.

Table 9: Bootstrapped Confidence Intervals Computed Empirically from Sampling Distribution Without Normality Assumption

Accuracy Measure	95% Confidence Interval	
	LB	UB
Total Accuracy	77.86	82.41
Poverty Accuracy	49.69	61.40
Undercoverage	38.60	50.31
Leakage	29.45	45.78
PIE	-4.51	0.58
BPAC	33.43	59.20