Poverty Assessment Tool Accuracy Submission
USAID/IRIS Tool for Peru
Submitted: September 15, 2011

The following report is divided into five sections. Section 1 describes the data used to create the Poverty Assessment Tool for Peru. Section 2 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 3 reports on the in-sample accuracy of each prediction model considered. Sections 4 and 5 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 reviews the out-of-sample accuracy for the Peru Poverty Assessment Tool.

1. Data source

For Peru, existing data from the Encuesta Nacional de Hogares (2009) were used to construct the poverty assessment tool. The full sample of 21,753 households is nationally representative. The sample used for tool construction comprises a randomly selected 16,315 households (75 percent of the full sample). The remainder, another randomly selected 5,438 households, is reserved for out-of-sample accuracy testing, which will investigate the robustness of in-sample poverty estimation. This PAT is intended to replace an earlier tool for Peru that was based on original data collected by the IRIS Center in 2004.

2. 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, and employment. The data collected in the household roster was used to construct household indicators such as *source of lighting*, *type of cooking fuel used,* and *type of toilet facility used*. For Peru, more than 64 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 the 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 Balanced Poverty Accuracy Criterion (BPAC) and the Poverty Incidence Error (PIE), along with practicality considerations.[2]

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

As explained more fully in Section 5, the line used to construct the poverty tool for Peru is the median line. Table 1 summarizes the accuracy results achieved by each of the eight estimation methods in predicting household poverty relative to this poverty line. For Peru, on the basis of BPAC, the 2-step Quantile regression model is slightly more accurate than the 1-step model. However, the 1-Step Quantile regression includes fewer variables. Following precedent from previous decisions made in consultation 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

Peru Median line* Share of "very poor": 15.9%	**Total Accuracy**	**Poverty Accuracy**	**Under-coverage**	**Leakage**	**PIE**	**BPAC**
Single-step methods						
OLS	88.03	55.08	44.92	29.86	-2.41	40.03
Quantile regression (estimation point: 42 percentile)	**87.68**	**62.11**	**37.89**	**39.05**	**0.18**	**60.95**
Linear Probability	88.43	48.91	51.09	21.21	-4.78	19.02
Probit	88.78	55.27	44.73	24.38	-3.10	35.91
Two-step methods						
OLS –28 percentile cutoff	88.95	58.25	41.75	27.28	-2.32	43.77
Quantile (estimation points: 42, 12) 28 percentile cutoff	88.52	67.23	32.77	38.88	0.98	61.12
LP – 22 percentile cutoff	89.15	59.33	40.67	27.14	-2.17	45.79
Probit – 22 percentile cutoff	89.02	59.23	40.77	27.82	-2.07	46.27
*The Median poverty line is 2056.75 Peruvian Nuevos Soles (PEN) per year per capita in 2009 prices.						

[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.

For Peru, the functionality of predicting the poverty rate at other poverty lines—in this case, the food line, median line, halfway between median and national line, national line, and 50% above national line —has been added. This functionality is based on statistical models for prediction at the median and national lines. The methodology and the accuracy results for this prediction are discussed in Annex 1.

4. 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 CSPro template used to calculate the incidence of extreme poverty among each implementing organization's clients.

5. 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 (1) those living on the local equivalent of less than the international poverty line ($1.25/day in 2005 PPP terms)[3] or (2) the poorest half of those living below the poverty line established by the national government. The applicable poverty line for developing USAID tools is the one that yields the higher household poverty rate in a given country.

In Peru the applicable threshold is the median line, at the level of prices prevailing when the household survey data were collected (January – December 2009). 31.8% of Peruvian households in the 2009 sample lived below the national poverty line of 3084 PEN per person per year. Based on this national poverty line, the median poverty line is 2056.75 Peruvian Nuevos Soles (PEN) per year per capita in 2009 prices. At these values, the median line identifies 15.9% of households as "very poor."

According to the second definition of extreme poverty in the legislation, the international poverty line of 853.42[4] PEN per adult equivalent per year identifies 1.1% of households as "very poor."[5]

[3] The 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. The World Bank adopted the $1.25/day line in 2008 to incorporate improved estimates of PPP based on data from 2005. The applicable 2005 PPP rate for Peru is 1.6534499.

[4] The calculation for the $1.25/day poverty line is 1.25*1.6534499*365.25*(113.05/100) where the final term is the CPI adjustment from 2005 prices to 2009 prices.

[5] The poverty rate (using population weights) is 5.9% on PovcalNet for 2009.

Hence the decision rule for Peru's USAID poverty assessment tool in classifying the "very poor" (and the "not very-poor") is whether that predicted per capita daily 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 is less than or equal to 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."

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)	1,622 (9.9%)	990 (6.1%)
Number of "true" not very-poor households (as determined by benchmark survey)	1,020 (6.3%)	12,683 (77.7%)

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

.42 Quantile regression — Number of obs = 16,315
Min sum of deviations 5284.281 — Pseudo R2 = 0.4576

Variable	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
Household size	-0.2170	0.0086	-25.26	0.000	-0.2338	-0.2002
Household size squared	0.0104	0.0008	13.80	0.000	0.0089	0.0119
Household head age	0.0002	0.0022	0.10	0.922	-0.0040	0.0044
Household head age squared	-0.0000	0.0000	-0.60	0.549	0.0000	0.0000
Household lives in rural location	-0.1751	0.0145	-12.06	0.000	-0.2036	-0.1467
Household lives in Costa Norte	-0.1118	0.0187	-5.98	0.000	-0.1485	-0.0751
Household lives in Costa Centro	-0.0628	0.0212	-2.96	0.003	-0.1043	-0.0213
Household lives in Costa Sur	-0.1885	0.0252	-7.49	0.000	-0.2378	-0.1392
Household lives in Sierra Norte	-0.1805	0.0217	-8.34	0.000	-0.2230	-0.1381
Household lives in Sierra Centro	-0.1438	0.0159	-9.05	0.000	-0.1750	-0.1127
Household lives in Sierra Sur	-0.2305	0.0169	-13.61	0.000	-0.2637	-0.1973
Household lives in Lima Metropolitan area	0.0522	0.0188	2.78	0.005	0.0156	0.0891
Dependency ratio	-0.0811	0.0084	-9.65	0.000	-0.0976	-0.0646
Household head lives with a partner	-0.0712	0.0127	-5.59	0.000	-0.0961	-0.0462
Household head has no education	-0.1761	0.0200	-8.81	0.000	-0.2152	-0.1369
Household head has completed technical school	0.1204	0.0214	5.62	0.000	0.0784	0.1624
Household head has completed university	0.2950	0.0234	12.59	0.000	0.2491	0.3409
Share of household members with no education	-0.0811	0.0084	-9.65	0.000	-0.0976	-0.0646
Number of rooms in household’s dwelling	0.0495	0.0059	8.38	0.000	0.0379	0.0611
Roof is predominantly made of reinforced concrete	0.1540	0.0148	10.43	0.000	0.1251	0.1830
Roof is predominantly made of straw, palm leaves, or etc.	-0.1889	0.0200	-9.46	0.000	-0.2280	-0.1497
Main cooking fuel is firewood	-0.0927	0.0137	-6.79	0.000	-0.1194	-0.0659
Household owns one or more gas stoves	0.1200	0.0143	8.40	0.000	0.0920	0.1480
Number of radios owned by household	0.0430	0.0083	5.18	0.000	0.0267	0.0592
Number of color TVs owned by household	0.1535	0.0090	16.97	0.000	0.1358	0.1712
Number of refrigerators or freezers owned by household	0.1666	0.0137	12.15	0.000	0.1397	0.1934
Number of cars, vans or pick-up trucks owned by household	0.1801	0.0182	9.87	0.000	0.1443	0.2158
Intercept	8.7703	0.0561	156.44	0.000	8.6605	8.8802

Annex 1: Poverty Prediction at the National Poverty Line and Discussion of Additional Poverty Lines

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 $1.25/day PPP models, this will be the $2.50/day PPP; for median poverty models, the "poor" threshold will be the national poverty line. Following this logic, then, the "poor" ("not poor") in Peru are defined as those whose predicted expenditures 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.

Based on the statistical models underlying prediction at these two lines, IRIS has also introduced the functionality of prediction at five lines to increase the usefulness of the tool to partner organizations. For Peru, these five lines are the food line, median line, halfway between median and national line, national line, and the 50% above national line. Poverty rates at the five lines are predicted using the best model for the median line, but poverty rates at the last two lines are predicted using only the best model for the national line. As discussed in this document, accuracy has been tested at the median and national lines. Given this, the predictions made at the other lines are intended for indicative use by implementing partners.

The tabulation of poverty prevalence has also been expanded to provide a fuller summary of the incidence of poverty among the implementing organization's clients. Poverty status at the five poverty lines is cross tabulated with regional location, household head's gender, household size, and housing conditions. Again, the additional information provided is for indicative purposes rather than statistical inference.

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

Peru National line* Share of "poor": 31.8%	**Total Accuracy**	**Poverty Accuracy**	**Under-coverage**	**Leakage**	**PIE**	**BPAC**
Single-step method						
Quantile regression (estimation point: 46)	85.13	77.19	22.81	23.64	0.27	76.36
*The National poverty line is 3084 Peruvian Nuevos Soles per capita per year in 2009 prices.						

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 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)	4,033 (24.7%)	1,191 (7.3%)
Number of "true" not poor households (as determined by benchmark survey)	1,235 (7.6%)	9,856 (60.4%)

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

.46 Quantile regression — Number of obs = 16,315
Min sum of deviations 5370.903 — Pseudo R2 = 0.4574

Variable	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
Household size	-0.2278	0.0085	-26.76	0.000	-0.2444	-0.2111
Household size squared	0.0114	0.0007	15.41	0.000	0.0099	0.0128
Household head age	-0.0011	0.0021	-0.50	0.615	-0.0053	0.0031
Household head age squared	0.0000	0.0000	0.04	0.969	0.0000	0.0000
Household lives in rural location	-0.1696	0.0143	-11.82	0.000	-0.1977	-0.1415
Household lives in Costa Norte	-0.1103	0.0184	-6.00	0.000	-0.1464	-0.0743
Household lives in Costa Centro	-0.0556	0.0210	-2.65	0.008	-0.0968	-0.0145
Household lives in Costa Sur	-0.1889	0.0248	-7.60	0.000	-0.2376	-0.1402
Household lives in Sierra Norte	-0.1852	0.0214	-8.66	0.000	-0.2271	-0.1433
Household lives in Sierra Centro	-0.1417	0.0157	-9.04	0.000	-0.1724	-0.1110
Household lives in Sierra Sur	-0.2384	0.0167	-14.26	0.000	-0.2712	-0.2056
Household lives in Lima Metropolitan area	0.0464	0.0186	2.50	0.012	0.0100	0.0829
Dependency ratio	-0.0815	0.0084	-9.73	0.000	-0.0979	-0.0651
Household head lives with a partner	-0.0741	0.0125	-5.90	0.000	-0.0987	-0.0495
Household head has no education	-0.1684	0.0197	-8.52	0.000	-0.2071	-0.1296
Household head has completed technical school	0.1271	0.0213	5.98	0.000	0.0855	0.1688
Household head has completed university	0.3062	0.0233	13.15	0.000	0.2605	0.3518
Share of household members with no education	-0.4695	0.0398	-11.79	0.000	-0.5476	-0.3914
Number of rooms in household’s dwelling	0.0533	0.0058	9.17	0.000	0.0419	0.0647
Roof is predominantly made of reinforced concrete	0.1633	0.0146	11.19	0.000	0.1347	0.1919
Roof is predominantly made of straw, palm leaves, or etc.	-0.1878	0.0196	-9.58	0.000	-0.2263	-0.1494
Main cooking fuel is firewood	-0.0896	0.0134	-6.67	0.000	-0.1159	-0.0632
Household owns one or more gas stoves	0.1206	0.0142	8.49	0.000	0.0928	0.1485
Number of radios owned by household	0.0456	0.0082	5.55	0.000	0.0295	0.0618
Number of color TVs owned by household	0.1533	0.0090	17.10	0.000	0.1357	0.1709
Number of refrigerators or freezers owned by household	0.1625	0.0136	11.95	0.000	0.1358	0.1891
Number of cars, vans or pick-up trucks owned by household	0.1836	0.0182	10.11	0.000	0.1480	0.2193
Intercept	8.8498	0.0556	159.21	0.000	8.741	8.9588

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 1, the data set used to construct the Peru tool was divided randomly into two data sets 16,315 households (75 percent of the sample) and 5,438 households (25 percent sample). 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.

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

	Total Accuracy	Poverty Accuracy	Under-coverage	Leakage	PIE	BPAC
In-Sample Prediction						
	87.68	62.11	37.89	39.05	0.18	60.95
Out-of-Sample Prediction						
	87.66	60.86	39.14	40.18	0.16	52.82

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.[6] 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 performance of this model is very good. The confidence interval around the sample mean BPAC is moderately narrow at +/- 6.38 percentage points. For

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

PIE, which measures the difference between the predicted poverty rate and the actual poverty rate, the confidence interval is +/- 1.0 percentage points.

Table 8: Bootstrapped Confidence Intervals on Assumption of Normality

Variable	Mean	Std. Dev.	Confidence interval	
			LB	UB
Total Accuracy	86.94	0.48	85.99	87.88
Poverty Accuracy	68.72	1.66	63.46	71.98
Undercoverage	31.28	1.66	28.02	34.54
Leakage	52.68	3.24	46.33	59.03
PIE	3.32	0.52	2.30	4.35
BPAC	47.32	3.24	40.97	53.67

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.

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

Accuracy Measure	95% Confidence Interval	
	LB	UB
Total Accuracy	85.65	88.81
Poverty Accuracy	63.66	73.76
Undercoverage	26.24	36.34
Leakage	43.99	65.06
PIE	1.82	5.00
BPAC	34.94	56.01

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 and Table 9 indicate that the extreme poverty rate estimate produced by the Peru PAT appears to be somewhat biased toward overestimating the actual extreme poverty rate, with a moderately narrow confidence interval of 2.30 to 4.35. By this measure, the predictive model behind the Peru PAT is accurate.