Developing Poverty Assessment Tools Project

Constructing an International Poverty Assessment Tool: A Methodological Note with Illustrations[1]

The IRIS Center

April 25, 2006

The Challenge

The IRIS/USAID project on developing poverty assessment tools has collected new data in four countries (Bangladesh, Kazakhstan, Peru, and Uganda) to determine how well a limited set of indicators is able to correctly predict a household's extreme poverty status. Alternative statistical methods have been tested to identify the 5, 10 or 15 best indicators and to determine how well these indicators predict extreme poverty. In addition, comparative analysis of eight existing LSMS data sets (Albania, Ghana, Guatemala, India (Bihar and Uttar Pradesh), Jamaica, Madagascar, Tajikistan, and Vietnam) has been undertaken to assess the robustness of the study results over a larger number of countries.

For each of these 12 countries, the best statistical model has been identified and reported elsewhere.[2] From the country-specific model results, a poverty assessment tool can be derived which is optimal for each country in question. The current challenge is to construct an international poverty assessment tool on the basis of the 12 best country-specific models.

Number of International Tools

This note proposes a methodology for constructing a single international tool or a set of international tools. Although in principle it is possible to construct a single international tool based on the results from all 12 countries, this is not very desirable in practice due to the large differences in extreme poverty rates and other characteristics across the countries. For example, extreme poverty rates vary from a low of 4.5% in Kazakhstan to a high of 47% in Tajikistan and 78% in the two states covered by the LSMS in India.

For that reason, this note will show results for separate international tools for two sets of countries classified according to their poverty incidence: 5 countries with low extreme poverty incidence (below 20% poverty rate) and 6 countries with high extreme poverty

---

[1] This note was written by Christiaan Grootaert and Anthony Leegwater. It has greatly benefited from comments by Thierry van Bastelaer and Manfred Zeller. All calculations for the international tool were undertaken by Anthony Leegwater.
[2] Full results can be found on the poverty tools website: http://www.povertytools.org/Project_Documents/Accuracy%20Results%20for%2012%20Countries.pdf

rates (above 20%).[3] The Indian LSMS data set was excluded due to the exceptionally high poverty rate.

It would also be possible to construct an additional international tool based on the results from the 4 test countries alone, in order to take advantage of the larger set of poverty predictors available for these countries.

* * * * *

Constructing an International Poverty Assessment Tool in Seven Steps

The international poverty assessment tool consists of three elements: a set of variables (poverty predictors), a set of weights associated with these variables, and a decision rule to determine whether a household is predicted as very poor or not very-poor. The 7 steps below describe how to undertake the three associated tasks of identifying the variables, calculating their weights, and setting the decision rule.

Prior to starting the construction of the international tool, it is necessary to select the estimation methodology which will be used in all phases of tool construction. It is especially important to decide whether a 1-step or a 2-step method will be used, because the structure of the decision rule is fundamentally different between these methods. For that reason, it is not possible to base the international tool on 1-step methods for some countries and 2-step methods for others.

It is to be noted that if the selected estimation method is a 2-step procedure, each of the seven steps below will have to be implemented separately for the first step and the second step of the method. The task of constructing an international tool is thus more complex with 2-step methods.

Step 0: Select for the international tool the estimation methodology that yielded the best model (highest BPAC) in the greatest number of countries.

The following table shows which estimation method yielded the best results for each of the 11 retained countries.[4]

[3] To avoid unwieldy adjectives in the following note, these two groups of countries are referred to as 'low-poverty' and 'high-poverty' in the text.
[4] A list defining acronyms and an explanation of the various poverty accuracy measures can be found at the end of this note.

	(1)	(2)	(3)	(4)	(5)	(6)
	Extreme poverty incidence	BPAC	PIE	Best model	BPAC reduction if no 2-step quantile	BPAC reduction if no quantile
Kazakhstan	5%	41	0.00	2-step quant.	2.7	32.4
Jamaica	8%	53	-0.13	2-step quant.	4.7	40.6
Albania	10%	52	0.20	1-step quant.	n/a	43.5
Ghana	13%	53	0.11	2-step quant.	4.2	55.4
Vietnam	15%	70	0.17	2-step quant.	2.2	18.3
Guatemala	23%	74	-0.06	2-step LPM	n/a	n/a
Peru	27%	72	0.13	1-step quant.	n/a	10.2
Bangladesh	31%	73	0.50	2-step quant.	4.8	4.8
Uganda	32%	67	0.13	2-step quant.	5.9	7.1
Madagascar	35%	78	0.08	2-step LPM	n/a	n/a
Tajikistan	47%	72	0.00	2-step quant.	0.0	0.0

The table shows that the best performing model is the two-step quantile model. Unfortunately, this model is by far the most complex to implement. From the perspective of constructing an international tool, it has the significant drawback of requiring uniformity of the estimation points through which the quantile model is fitted in both the first and second step, as well as uniformity of the sample cutoff point for the second step. Since earlier sensitivity analysis with the quantile model has revealed that its performance is sensitive to even small deviations from the optimal estimation point, we opted to remove the two-step quantile model from consideration. As column (5) of the table above shows, this leads to only small losses in performance. Thus, for the group of 5 low-poverty countries, the international tool will be based on the one-step quantile model. For the 6 high-poverty countries, the best performing model becomes the two-step linear probability model, after the two-step quantile model has been removed from consideration.

An alternative to Step 0 as formulated above is to select the methodology that yields the highest average BPAC over the countries included in the tool construction effort. While this ensures the highest average accuracy, it may lead to a sub-optimal estimation methodology in more countries. Consider the following example of BPAC values for estimation methods A and B in 3 countries:

	Method A	Method B
Country 1	60	59
Country 2	70	80
Country 3	80	79
Average	70	72.7

The first criterion proposed above would select method A because it has a higher BPAC in 2 out of 3 countries, but the alternative criterion would select method B because it has an average BPAC of 72.7 against an average BPAC of 70 for method A.

1. Finding the set of variables for the international tool

Step 1: Identify categories of variables that will be included in the tool

A practical way to implement this step is to retain a category if it includes at least one variable that was selected as poverty predictor in two or more countries, by the best methodology identified in Step 0.

In the case of the international tool for the low-poverty countries, 9 categories were thus retained: education, housing, clothing expenditures, agricultural assets, financial assets, other assets, consumer durables, illness, and a residual "other" category.

Step 2: Identify predictor variables within each retained category

In line with the previous step, a practical way to implement Step 2 is to include a variable if it was selected as poverty predictor in two or more countries, by the best methodology identified in Step 0.

In the case of the international tool for the low-poverty countries, the table below shows the variables that were selected in the 9 categories retained in Step 1.

Education	**Financial assets**
Members with complete primary	Amount of loans
Members with incomplete primary	Amount of savings
Members with no education	Received remittances as proportion of expenditures
Members with incomplete secondary	Amount of sent remittances
Household head with university	
Proportion of household in school	**Other assets**
	Total land owned
Housing	
House size: number of rooms	**Consumer durables**
House size: square area	Refrigerator
Charcoal cooking fuel	Car
Gas cooking fuel	Cellular phone
Earthen floors	Computer
Electric lighting	Phone (land-line)
Kerosene lighting	Stove
Other lighting	Video-cassette recorder (VCR)

Flush toilet	Motorcycle
Latrine toilet	Musical instrument
Other toilet facilities	Radio
Clay walls	Truck
Water source: tap	Television
	Value of dishes
Clothing expenditures per capita[5]	Value of jewelry owned
Agricultural assets	**Illness**
Resale value of animals	Proportion of household with illness
Number of cows owned	
Owner of farm	**Other**
Generator	Female household head
Number of sheep and goats owned	Household head is craftsman
Water pump	Household head is salaried worker

While this approach is simple and practical, it does run the risk of including variables that are either irrelevant for certain countries or, even if relevant, poor predictors. This is the case because variables in the data set differ greatly across countries and because the predictive ability of a given variable can also differ greatly across countries. For that reason, it may be preferable to delete Step 2 and design the international tool only at the level of variable categories rather than individual variables. The selection of individual variables within the common categories would then be up to participating countries. However, it could be argued that this is not truly an international tool, because variables would not be the same across countries. Also, it is impossible to test what the loss in accuracy is from using an international tool defined only at the category level relative to country-specific tools (which are defined at the variable level). Thus, for the purposes of this note, we will attempt to construct the international tool at the variable level.

However, in order to accommodate this difficulty, we will introduce an alternative method for variable selection. It consists of standardizing the selected variables across countries. For example, instead of using "thatched roof" as a specific variable in the international tool, we construct an index of roof quality for all countries and use this index variable in the international tool. The construction of the index for roof quality was as follows (the same technique was used for the other housing characteristics).[6] First, the minimum number of roof categories across the countries is determined. In the group of five countries with low extreme poverty, the minimum number of roofing categories was four. Thus, the index variable for roofing had values from 0 to 3; the categories in each country were ranked along this scale according to mean household expenditures. If a country had more than four roofing categories, they were combined according to mean expenditures.

[5] Clothing expenditures and other monetary variables (value of dishes, etc.) are in logarithmic form.
[6] The other index variables for housing characteristics were: wall, cooking fuel, floor, lighting, toilet facilities, and drinking water.

Indices were also created for consumer durables in four categories: entertainment, communication, transportation, and other. The value of an index for each household was the unweighted sum of the durables it owned in that category. For example, a family with a television and radio but no other entertainment devices would have a value of 2 for the entertainment durables index.

Finally, educational attainment was also captured by indices. The education of the household head was measured on a scale from 0 to 5 (no education, primary incomplete, primary complete, secondary incomplete, secondary complete, university). The index for member educational attainment was based on the same scale; the number of household members attaining a particular educational level was multiplied by the relevant number, and then the multiplicative terms were summed across all education levels. For example, if a household had three members who did not complete primary school and two members who completed secondary school, the index value for that household would be: 3*1 + 2*4 or 11.

It is likely that the application of Steps 1 and 2 will lead to a data set that contains more than 15 variables. There remains thus a need to narrow the variable set to the required 15 (or 10 or 5) best variables. This is the purpose of Steps 3 and 4.

Step 3: Re-estimate the "best 15" model for each country starting with the set of variables selected in Step 2. Use the estimation methodology that was identified as best for that country.

The purpose of this step is to reduce the number of country cases where a selected variable is unavailable or irrelevant.

It needs to be pointed out that the estimation methodology for Step 3 must be of the same type (i.e. one-step or two-step) as the method selected in Step 0 for the international tool. For example, the international tool for the low-poverty countries is based on the one-step quantile model. If the best model for an individual country in this group is a two-step model, one needs to reject this model for Step 3 and use instead the best one-step model.

Step 4: Retain the 15 variables most frequently selected as poverty predictors in the regressions of Step 3. These are the "best 15" variables for the international tool.

It is possible that the regression results from Step 4 lead initially to retaining more than 15 variables. In that case, the variable set can be reduced to the desired number by adding considerations of variable availability and ease of data collection. For the five countries with low extreme poverty incidence, there were 21 variables that were found in the best 15 for two or more countries. To narrow down this list, we charted the availability of the variables that occurred only twice. We dropped the least available indicators (the five found in three countries or less), reducing the list to 16. We then dropped the one variable (remittances received as a proportion of total expenditures) of those available in four countries that was least practical to collect in the field.

It needs to be pointed out that Steps 3 and 4 do not address the problem posed by the control variables. Common to all countries are household size (and its square) and age of the household head (and its square). However, the location control variables differ from country to country and, by definition, can not be harmonized. Therefore, the location variables need to be made part of the inevitable country-specific part of the international tool, i.e. they are to be combined with the intercept (see below, in the section on the decision rule).

2. Determining the weights for the retained variables

Step 5: Re-estimate the "best" model for each country with the 15 variables from Step 4, using the earlier identified (see Step 0) common estimation methodology, and using data sets adjusted with sampling weights.

The purpose of this step is to provide comparable results across countries, both in terms of type of model and set of variables.

The use of a common estimation methodology obviously implies that for some countries suboptimal estimation methodologies will be used. This is a first reason why an international tool will likely have worse accuracy than country-specific tools. The use of a common estimation methodology is, however, necessary due to the incomparability of coefficients between estimation methodologies with continuous dependent variables (such as OLS and quantile regression with ln(household expenditure per capita) as dependent variable) and those with discrete dependent variables (such as LPM and probit with a dichotomous extreme poverty indicator as dependent variable). Furthermore, even within models with continuous dependent variables, coefficients from quantile models are not comparable with those from OLS models.

If the common estimation method is a two-step method, the cutoff point for the second-step estimation has to be standardized. This is likely to lead to additional losses in accuracy. Moreover, since the cutoff point can not fall below the extreme poverty line, the country with the highest extreme poverty rate will determine the location of the cutoff point. If this country has a very high extreme poverty rate, say, above 50%, this would in practice render the two-step method useless for the purpose of constructing an international tool. This is a major reason why the construction of an international tool should only be done for groups of countries with similar extreme poverty rates. In the case of our set of 6 countries with high extreme poverty rates, the inclusion of Tajikistan, with a extreme poverty rate of 47%, did pose some difficulties for the construction of the international tool based on the two-step LPM.

If the common method is a quantile model, the estimation point (the point on the distribution through which the quantile model is fitted) has to be standardized. A practical way to do this is to use the average of the estimation points that were used in the country-specific best models. This procedure will, however, also lead to losses in accuracy.

Step 6: Calculate the unweighted average of the coefficients from the country-specific models from Step 5 in order to determine the weights for the international tool. If a variable is not available in a given country, this country is excluded from the calculation of the weight for that variable.

If the common estimation method is an OLS or quantile model, for which the dependent variable is expressed in local currency, all coefficients need to be converted into PPP-terms before the average can be calculated. Given that the dependent variable is in logarithmic terms, this is best achieved by re-estimating the regression with the natural logarithm of household expenditure per capita expressed in PPP-dollars as the dependent variable. All explanatory variables that are monetary (e.g. value of assets, loan amount) also need to be expressed in PPP-terms. This will produce internationally comparable regression coefficients. However, for some countries the PPP-conversion factors have proven to be of questionable validity, and this can introduce unknown errors in the calculation process.

With regards to the common control variables (household size, age of the household head, and their squares), the coefficients can either be averaged, like all other variables, or they can be kept country-specific, like the location control variables. In the tables below, we show results for both options.

3. Decision rule

The ultimate purpose of the tool is to determine whether a household is predicted as very poor or not very-poor. To do this, a critical value needs to be calculated which is then compared to the extreme poverty line. If the critical value falls below the extreme poverty line, the household is deemed to be very poor, and not very-poor otherwise.[7]

Step 7: Calculate the critical value C = intercept + Σ (weight x variable), where all variables take the values specific to the household.

The decision rule can not be internationalized because the intercept is always country-specific. Thus the intercept in the formula above is country-specific, as estimated in the regressions in Step 5. It is necessary to include in this formula the coefficients of the location control variables, which, since they are dummy variables, directly modify the intercept. In contrast, the expression *Σ (weight x variable)* is common to all countries, and is the core of the international tool

The tables below show the international tool, constructed separately for the low-poverty countries and the high-poverty countries.

---

[7] The decision rule is slightly different for the linear probability and probit models. In a 2-step model, the critical value for the 1st step is the predicted probability that the household's per capita expenditures exceed a certain threshold. In the 2nd step, the critical value is the predicted probability that a household is very poor. If this probability exceeds a cutoff probability of .5 in the 2nd step, the household is predicted to be very poor.

Five Countries with Low Extreme Poverty Incidence: 1-step Quantile

Dependent variable: ln(household expenditure per capita in PPP-dollars)

Variables	Weight
Intercept	country-specific
Location controls	country-specific
Household head age	0.009
Household size	-0.234
Household head age squared	-8.60E-05
Household size squared	0.011
Clothing expenditures per capita	0.171
Number of rooms	0.039
Mud or clay walls	-0.123
Refrigerator	0.193
Car	0.385
Phone (land-line)	0.150
Stove	0.141
VCR	0.101
Members with complete primary	-0.269
Members with no education	-0.190
Members with incomplete primary education	-0.169
Household head with university education	0.128
Amount of remittances sent	0.018
Proportion of household with illness	0.058
Female household head	-0.005

Five Countries with High Extreme Poverty Incidence: 2-step Linear Probability

1st Step Dependent Variable: Household Expenditure Above Cutoff

1st Step Variables	**Weight**
Intercept	country-specific
Location controls	country-specific
Household head age	0.005
Household size	-0.128
Household head age squared	-4.15E-05
Household size squared	0.005
Clothing expenditures per capita	0.059
Members with no education	-0.189
Remittances sent as proportion of expenditures	0.659
Number of rooms	0.033
Car	0.103
Wood as cooking fuel	-0.058
Clay floors	-0.127
TV	0.053
Value of TV and VCR	0.046
Bicycle	0.054
Number of cows owned	0.012
Refrigerator	0.020
Female household head	-0.060
Radio	-0.027
Roof of leaves	-0.042

2nd Step Dependent Variable: Extreme Poverty Status

2nd Step Variables	**Weight**
Intercept	country-specific
Location controls	country-specific
Household head age	-0.010
Household size	0.110
Household head age squared	8.07E-05
Household size squared	-0.003
Bicycle	-0.071
Car	-0.117
Cellular phone	-0.113
Days lost to illness	-0.003
Refrigerator	-0.320

Clothing expenditures per capita	-0.061
Head with university education	-0.280
Candles as lighting	-0.115
Number of pigs owned	-0.050
Proportion of household with illness	-0.063
Remittances received as proportion of expenditures	-0.259
Number of rooms	-0.060
Flush toilets	-0.065
No toilet facilities	0.072
Standing water as water source	-0.009

* * * * *

Comparing the International Tool with Country-Specific Tools

As the process explained above indicates, constructing an international tool involves a number of compromises made for the sake of comparability, which in each case reduce the accuracy of the tool for a specific country. The compromises pertain to both variable selection and calculation of a common set of weights. It would thus be useful to apply the constructed international tool to the countries on which it is based to see how its accuracy compares to that of the original country-specific models. This comparison is presented in the tables below and shows the loss in accuracy from internationalizing the poverty assessment tool.

Five Countries with Low Extreme Poverty Incidence

ALBANIA (median) Poverty Rate[8]: 10.42%	Total Accur.	Pov. Accur.	Under-coverage	Leakage	PIE	BPAC
Single-step -- MAXR variable selection						
Quantile regression (estimation point: 34)	90.21	53.97	46.03	47.94	0.20	52.06
Single-step -- INTL model						
Quantile regression (estimation point: 34)	90.91	11.50	88.50	8.36	-7.52	-68.64
Quantile regression – country-specific age, size (estim. point: 34)[9]	91.20	30.74	69. 26	24.50	-4.20	-14.03
Single-step -- INTL model with indices[10]						
Quantile regression (estimation point: 34)	91.10	9.35	90.65	4.18	-8.12	-77.13

[8] "Poverty rate" is the percentage of the household sample that is very poor.
[9] Country-specific coefficients are attached to the control variables for the age of the household head and household size (and also to their respective, squared values).
[10] Index variables were constructed for housing characteristics, education levels of the household head and household members, and certain categories of consumer durables (entertainment, communication, transportation, and other).

GHANA (median) Poverty Rate: 13.42%	Total Accur.	Pov. Accur.	Under-coverage	Leakage	PIE	BPAC
Single-step -- MAXR variable selection						
Quantile regression (estimation point: 38)	86.74	50.00	50.00	48.82	-0.16	48.81
Single-step -- INTL model						
Quantile (estimation point: 34)	87.26	42.88	57.12	37.80	-2.59	23.56
Quantile regression – country-specific age, size (estim. point: 34)	83.55	66.44	33.56	88.98	7.44	11.02
Single-step -- INTL model with indices						
Quantile regression (estimation point: 34)	85.99	59.83	40.17	64.24	3.23	35.76

JAMAICA (median) Poverty Rate: 8.03%	Total Accur.	Pov. Accur.	Under-coverage	Leakage	PIE	BPAC
Single-step -- MAXR variable selection						
Quantile regression (estimation point: 34)	91.72	48.44	51.56	51.56	0.00	48.44
Single-step -- INTL model						
Quantile (estimation point: 34)	92.62	5.21	94.79	2.96	-6.94	-86.62
Quantile regression – country-specific age, size (estim. point: 34)	91.83	14.39	85.61	22.53	-4.77	-48.69
Single-step -- INTL model with indices						
Quantile regression (estimation point: 34)	92.74	6.22	93.78	2.29	-6.91	-85.26

KAZAKHSTAN (median) Poverty rate: 4.52%	Total Accur.	Pov. Accur.	Under-coverage	Leakage	PIE	BPAC
Single-step -- MAXR variable selection						
Quantile regression (estimation point: 23)	94.74	45.95	54.05	62.16	0.37	37.84
Single-step -- INTL model						
Quantile (estimation point: 34)	93.64	21.62	78.38	62.16	-0.73	5.41
Quantile regression – country-specific age, size (estim. point: 34)	94.25	8.11	91.89	35.14	-2.57	-48.65
Single-step -- INTL model with indices						
Quantile regression (estimation point: 34)	92.41	16.22	83.78	83.78	0	16.22

VIETNAM (median) Poverty Rate: 14.52%	Total Accur.	Pov. Accur.	Under-coverage	Leakage	PIE	BPAC
Single-step -- MAXR variable selection						
Quantile regression (estimation point: 42)	91.05	68.87	31.13	30.55	-0.08	68.29
Single-step -- INTL model						
Quantile (estimation points: 34)	34.46	99.74	0.26	415.98	65.46	-315.98
Quantile regression – country-specific age, size (estim. point: 34)	50.10	99.60	0.40	316.50	49.77	-216.50
Single-step -- INTL model with indices						
Quantile regression (estimation point: 34)	38.16	99.91	0.09	392.67	61.81	-292.67

Six Countries with High Extreme Poverty Incidence

BANGLADESH (ppp) Poverty rate: 31.41%	Total Accur.	Pov. Accur.	Under-coverage	Leakage	PIE	BPAC
Two-step method -- MAXR variable selection						
LP – 43 percentile cutoff	84.85	73.71	26.29	21.91	-1.37	69.32
Two-step method -- INTL model						
LP – 49 percentile cutoff	37.05	100	0	200.40	62.95	-100.40
LP – country-specific age, size – 49 percentile cutoff	70.96	84.06	15.94	76.49	19.02	23.51
Two-step method -- INTL model with indices						
LP – 50 percentile cutoff	45.43	100	0	173.71	54.57	-73.71

GUATEMALA (median) Poverty Rate: 22.96%	Total Accur.	Pov. Accur.	Under-coverage	Leakage	PIE	BPAC
Two-step method – MAXR variable selection						
LP – 30 percentile cutoff	88.16	74.10	25.90	25.66	-0.06	73.86
Two-step method -- INTL model						
LP – 49 percentile cutoff	78.89	8.89	91.11	0.86	-20.71	-81.35
LP – country-specific age, size – 49 percentile cutoff	85.79	57.31	42.69	19.22	-5.39	33.85
Two-step method -- INTL model with indices						
LP – 50 percentile cutoff	79.38	11.29	88.71	1.14	-20.10	-76.29

MADAGASCAR (ppp) Poverty Rate: 35.22%	Total Accur.	Pov. Accur.	Under-coverage	Leakage	PIE	BPAC
Two-step method -- MAXR variable selection						
LP – 37 percentile cutoff	84.85	78.61	21.39	21.62	0.08	78.38
Two-step method -- INTL model						
LP – 49 percentile cutoff	75.25	67.91	32.09	18.68	-6.54	54.50
LP – country-specific age, size – 49 percentile cutoff	71.85	51.90	48.10	9.64	-18.75	13.44
Two-step method -- INTL model with indices						
LP – 50 percentile cutoff	78.67	80.23	19.77	23.98	2.06	76.02

PERU (median) Poverty rate: 26.88%	Total Accur.	Pov. Accur.	Under-coverage	Leakage	PIE	BPAC
Two-step method -- MAXR variable selection						
LP – 59 percentile cutoff	86.87	69.30	30.69	18.13	-3.37	56.74
Two-step method -- INTL model						
LP – 49 percentile cutoff	79.50	29.77	70.23	6.05	-17.25	-34.42
LP – country-specific age, size – 49 percentile cutoff	81.13	43.72	56.28	13.95	-11.38	1.40
Two-step method -- INTL model with indices						
LP – 50 percentile cutoff	82.50	56.74	43.26	21.86	-5.75	35.35

TAJIKISTAN (median) Poverty Rate: 47.29%	Total Accur.	Pov. Accur.	Under-coverage	Leakage	PIE	BPAC
Two-step method -- MAXR variable selection						
LP – 52 percentile cutoff	74.19	73.49	26.51	28.07	0.74	71.93
Two-step method -- INTL model						
LP – 49 percentile cutoff	66.89	60.72	39.28	30.72	-4.05	52.17
LP – country-specific age, size – 49 percentile cutoff	64.67	33.37	66.63	8.07	-27.69	-25.18
Two-step method -- INTL model with indices						
LP – 50 percentile cutoff	61.31	22.29	77.71	4.10	-34.81	-51.33

UGANDA (ppp) Poverty rate: 32.36%	Total Accur.	Pov. Accur.	Under-coverage	Leakage	PIE	BPAC
Two-step method -- MAXR variable selection						
LP – 49 percentile cutoff	79.44	64.71	35.29	28.23	-2.28	57.64
Two-step method -- INTL model						
LP – 49 percentile cutoff	67.51	0	100	0.39	-32.23	-99.61
LP – country-specific age, size – 49 percentile cutoff	69.29	10.19	89.80	5.09	-27.41	-74.51
Two-step method -- INTL model with indices						
LP – 50 percentile cutoff	67.77	0.39	99.61	0	-32.23	-99.22

For each set of countries, the international tool has performance measures that are far below those obtained by the country-specific models. In the case of the low-poverty countries, BPAC declines from an average value of 51.1 points to -88.5. Negative values for BPAC reflect very low poverty accuracy and/or large imbalances between leakage and undercoverage, to an extent that a tool has to be rejected. This is confirmed by the values of PIE. The country-specific models have an average value of PIE that is not distinguishable from zero, i.e. they yield close-to-perfect poverty rate predictions. The international tool, in contrast, is off by an average of 16.8 percentage points from the true poverty incidence. For the high-poverty countries, BPAC drops on average from a very good 68.0 points to a dismal -34.9. The values of PIE indicate that the international tool incorrectly predicts the poverty rate by 22.6 percentage points on average. These numbers represent drops in accuracy so huge as to render the international tool practically useless. There is, as is to be expected, significant variation from country to country.

While for a few countries, the performance of the international tool is respectable, the practical difficulty is that it is not possible to determine *a priori* for which countries this would be the case.

The alternative approaches to international tool construction discussed earlier in this note do not lead to a tool that is much better. Specifically, keeping the weights of the control variables at their country-specific values leads generally to better results than the "pure" international tool, but tool performance is still unacceptably low. The same is true for the use of standardized variables for education, housing and consumer durables across countries.

The conclusion clearly is that an argument needs to be made in support of either calibration of an international tool to specific countries, or abandoning the international tool altogether in favor of only country-specific tools.

Calibration of the International Tool to Specific Countries

Calibration consists of adjusting the variables and/or the weights in the international tool to the situation of a specific country. It is possible to adjust the weights without changing the variables, but it is not possible to change the variables without re-calculating weights. Since the decision rule is always country-specific, calibrating both variables and weights is tantamount to estimating a new country-specific model.

Changing variables can be done on the basis of expert opinion, or, more rigorously, by re-estimating a prediction model on data from a recent household survey, such as LSMS, household expenditure survey, SDA Priority Survey (in sub-Saharan Africa), etc. This estimation requires access to the micro-data of the survey.

Adjusting the weights always requires re-estimating the prediction model on recent household survey data from the country. This implies that it will rarely be pragmatic to calibrate weights only. There is a fixed cost involved in obtaining access to the data from a household survey, and the difference in effort in calibrating weights only versus calibrating both variables and weights is minimal. The main difference is that in the former case a model is estimated with a predetermined set of variables, while in the latter case a variable search program (such as MAXR in SAS) must be included in the estimation routine. This represents only a marginal increase in cost or complexity.

Setting up the decision rule requires, in addition to determining the variables and their weights, estimating the intercept (and, preferably, also its location-specific shift factors). By definition, it is not possible to do this correctly without estimating an econometric model. However, there is an approximation possible based on the empirical observation that in most models with ln(expenditure) as dependent variable, the estimated intercept is very close (often within 1%) of the mean. Thus, if the statistical office of a country has published the results from a recent expenditure survey, the report will usually include the mean of the expenditure distribution. Frequently, the report will also include location-specific means, such as by region and by urban/rural. In such a case, access to the micro-data is not needed. However, if the international tool is based on a two-step estimation procedure, the intercept must be known for both steps. Since the first step is estimated on the full sample, this is not a problem, because the full-sample mean will typically be published in the survey report. The mean for the second step is more difficult to estimate, because it pertains to a subsample consisting of the x% poorest households. The calculation of this mean requires access to the micro-data or to a program, such as the World Bank's POVCAL, which can approximate the distribution from published aggregated data.

It is clear from this discussion that calibration is best done when access can be gained to the data from a recent household survey. If such access is not possible or if there is no recent survey in the country, calibration becomes a very subjective exercise with an unknown degree of error. It is also clear that, once data access is secured, the effort required to calibrate an existing international tool is virtually identical to that of estimating a new country-specific model.[11] In light of this, one can raise the question whether it makes sense to attempt to calibrate an international tool, since this will always be second-best to estimating a new country-specific model. This seems to be a case where the first-best can be obtained at little or no additional cost.

* * * * *

Conclusion

The motivation behind an international poverty assessment tool is compelling: a simple tool which, once developed, can be applied across a broad range of countries at low cost. However, in developing the international tool, much of the poverty accuracy displayed by a tool specialized to a particular country is lost.

In order to develop a viable international tool, one must select a limited number of poverty indicators, estimate weights to attach to these indicators, and determine a decision rule to categorize households as very poor or not very-poor. These tasks were accomplished in the steps outlined above to narrow the list of indicators to 15, calculate the average coefficients across countries, and determine the predicted household expenditures for each household.

Constructing the international tool from multiple countries required a number of compromises in variable selection and in the development of common weights for the sake of comparability. These compromises proved to have a heavy cost in poverty accuracy when compared to the country-specific models. Our preferred measure of

[11] Calibrating an existing international tool may be slightly cheaper than a country-specific tool. Data construction and checking for the 15 to 30 poverty indicators in the international tool will take less time than a similar effort on the full range of possible poverty indicators for a country-specific model. The estimation technique is also known before calibration, but the best technique would need to be selected from multiple alternatives for a country-specific model.

poverty accuracy, BPAC, fell from an average of 51.1 points to -88.5 for the group of five low-poverty countries and from an average of 68.0 points to -34.9 for the six high-poverty countries. Additional changes were tested on the model to circumvent some of these compromises, but yielded little overall improvement in performance. The additional step of calibrating an international tool to better fit a particular country can be accomplished reasonably with access to a recent household survey, but this step is likely to be only slightly cheaper than developing a new country-specific model with better accuracy.

Acronyms

BPAC: Balanced Poverty Accuracy Criterion

LPM: Linear Probability Model

LSMS: Living Standards Measurement Study

OLS: Ordinary Least Squares

PIE: Poverty Incidence Error

PPP: Purchasing Power Parity

SDA: Structural Dimension of Adjustment survey

Poverty Accuracy Measures

Note: As a convenient shorthand, all households living below the extreme poverty line are referred to as “very poor” or “VP”; all households living above the extreme poverty line are referred to as “not very-poor” or “NVP.” The technical term "not very-poor" may tend to obscure the fact that many of these households would be considered poor or very poor even by the standards of many developing countries, and desperately poor by the standards of developed countries like the United States.

Total Accuracy = sum of correctly predicted VP plus correctly predicted NVP, expressed as a percentage of the total sample.

Poverty Accuracy = correctly predicted VP as a percentage of total “true” VP.

Undercoverage = “true” VP incorrectly predicted as NVP, expressed as a percentage of total “true” VP.

Leakage = “true” NVP incorrectly predicted as VP, expressed as a percentage of total “true” VP.

Poverty Incidence Error (“PIE”) = difference between the number of households predicted to be VP and the number of households who are “true” VP, expressed as percentage of the total sample.

Balanced Poverty Accuracy Criterion (“BPAC”) = Poverty Accuracy minus the absolute difference between Undercoverage and Leakage