The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

Poverty comparisons with endogenous absolute poverty lines

Kenneth R. Simler
Research Fellow
Food Consumption and Nutrition Division
International Food Policy Research Institute
2033 K Street, NW
Washington, DC 20006-1002 USA
k.simler@cgiar.org
Tel: +1.202.862.5636
Fax: +1.202.467.4439

Channing Arndt*
Principal Advisor
National Directorate of Studies and Policy Analysis
Ministry of Planning and Development
Maputo, Mozambique
and
Associate Professor
Department of Agricultural Economics
Purdue University
carndt@purdue.edu
Tel: +258 82 314 6920

Contributed paper prepared for presentation at the International Association of Agricultural Economists conference, Gold Coast, Australia
August 12-18, 2006.

Copyright 2006 by Simler and Arndt.

Poverty comparisons with endogenous absolute poverty lines

1. Introduction

A principal objective of poverty measurement is to make comparisons between groups. Did poverty decline following implementation of a poverty reduction program? Is poverty higher in the hills or on the coast? These questions have become ever more important in recent years. Besides the high-profile Millennium Development Goal of halving world poverty by 2015, country development programs and donor support are increasingly driven by the Poverty Reduction Strategy Paper (PRSP) process, which requires close monitoring of poverty levels and detectable progress in reducing poverty.

There are many ways to define and measure poverty, but with few exceptions the empirical basis for poverty comparisons is statistical, using estimates from household survey data. Research over the past 15 years has increasingly refined statistical inference methods for poverty measures (Kakwani, 1993; Bishop et al., 1995; Ravallion, 1994a; Howes and Lanjouw, 1998). All of the methods used for absolute poverty lines tacitly assume that the source of statistical error is in the welfare metric, e.g., income, expenditure, or consumption. That is, the welfare metric is treated as a random variable, but the poverty line is treated as a fixed constant.

In fact, poverty lines are often estimated from the same survey data as the welfare measure, and thus are also random variables. This has been recognized in the relative poverty literature, where poverty lines are often computed directly from empirical income distributions, such as one-half of median income. The two sources of error (the welfare metric and the poverty line) could reinforce or offset one another, such that accounting for sampling error in the poverty line could increase or reduce the precision of poverty estimates (Preston, 1995). In Zheng’s (1997, 2001) empirical applications the sampling error of the poverty line always increases the standard errors of poverty estimates.

This paper brings these two strands of the literature together to provide a method for more accurately assessing the precision of estimates of absolute poverty, leading to more reliable poverty comparisons. It argues that like relative poverty lines, absolute poverty lines that are estimated from sample data (which is the norm) have a sampling error that needs to be included in the standard errors for poverty measures. We present a bootstrapping procedure for estimating the sampling error of absolute poverty lines, and assess its effect on the standard error of Foster-Greer-Thorbecke (1984) poverty measures. We use recent household survey data from Mozambique to explore the impact on poverty comparisons.

The remainder of the paper is structured as follows. Section 2 considers challenges in estimating poverty and assessing the precision of estimated poverty measures. This is followed by a description of the methods and data in section 3. Section 4 presents empirical results. Section 5 summarizes and concludes, including remarks about the scope for wider application of this procedure.

2. Estimating poverty

The measurement of poverty poses two fundamental questions (Sen, 1976). First, how does one identify the poor among the total population? Second, how does one aggregate information on individuals and households into a scalar measure of poverty? The first question has two components, namely, how do we measure individual welfare and, using this same metric, how do we determine the threshold that separates the poor from the nonpoor?

These elements are illustrated clearly in commonly used poverty measures. For example, at the household level, the general form of the Foster-Greer-Thorbecke (FGT) measure can be written as:

$$P_{\alpha,j} \quad = \quad \max\left[0\,,\left(\frac{z-y_j}{z}\right)^{\alpha}\right] \quad , \quad \alpha \geq 0 \tag{1}$$

where, y_j is a money-metric welfare measure for household j and z is the poverty line. An aggregate scalar measure of poverty in the population, P_α, is obtained as the weighted mean of (1) over all households. The weights are the number of members in each household (h_j), and survey sample weights (or expansion factors) w_j, so that an unbiased estimator of poverty in the population is:

$$P_\alpha = \frac{\sum_{j=1}^{n} w_j\, h_j P_{\alpha,j}}{\sum_{j=1}^{n} w_j\, h_j}\,. \tag{2}$$

The crux of our argument goes back to equation (1). Whereas the welfare metric y_j is treated as a random variable with a sampling error, the absolute poverty line z is routinely treated as a fixed constant, even though it is also estimated from the survey data. Ignoring this variance component leads to incorrect estimates of the precision of poverty estimates, and potentially misleading poverty comparisons over time and space.

The methods presented in this paper can be applied to any poverty line that is estimated statistically, including the Cost of Basic Needs (CBN) approach (Ravallion, 1998) and the Food Energy Intake (FEI) method (Greer and Thorbecke, 1986). Similarly, they are not limited to FGT poverty measures. In the empirical application in this paper, we focus on one method for estimating poverty lines (CBN), and two poverty measures in the FGT class (the headcount and poverty gap).

3. Data and methods

We use data from the 2002–03 Mozambique Household Budget Survey, also known by its Portuguese abbreviation IAF (for Inquérito aos Agregados Familiares sobre Orçamento Familiar) (see INE, 2004 for additional details). The survey was conducted from July 2002 through June 2003. A stratified three-stage cluster sampling procedure was used to select 8,700 sample households in 857 enumeration areas (EA). The unequal probability of selection across EAs requires the use of sampling weights, which are calculated as the inverse of the probability of selection.

The welfare metric is consumption per capita, following the approach described by Deaton and Zaidi (2002). Sensitivity analysis with adult equivalence scales altered the ranking of households, but did not affect the aggregate poverty measures reported here. As food prices tend to follow a seasonal pattern, an intra-survey temporal food price index was developed from the survey data, and all nominal values of food consumption were adjusted by the index.

The CBN approach was used to set poverty lines. As relative prices of basic foods vary widely in Mozambique, we allowed both the reference food bundles and the price vectors to vary by region (Ravallion, 1998; Tarp et al., 2002). Thirteen poverty line regions were defined based on an aggregation of the 21 survey strata, preserving the distinction between rural and urban areas, but grouping adjacent strata with similar characteristics if they had relatively few observations. For each poverty line region, the food poverty line is constructed by determining the composition and caloric content of the typical diet of the poor in that region, the average cost (at local prices) of a calorie when consuming that diet, and the food energy intake requirements for the reference population (the poor). Caloric requirements for moderately active individuals, disaggregated by age and sex, were obtained from the World Health Organization

(WHO, 1985). Average per capita requirements were allowed to vary by poverty line region, reflecting differences in the average household composition across regions.

The relevant food bundles and associated prices were estimated for relatively poor households using the iterative procedure described by Ravallion (1998). To ensure that the region-specific bundles were of comparable quality we employed revealed preference tests and an entropy estimation procedure to adjust the composition of the bundles such that they satisfy revealed preference conditions and retain the maximum information content inherent in the original estimated bundles (see Arndt and Simler, 2005). The nonfood component of the poverty line was estimated non–parametrically, using the weighted mean nonfood budget share among those households whose total expenditure is approximately equal to the region-specific food poverty line (Ravallion, 1994b, 1998).

After calculating the welfare metric and region-specific poverty lines, equation (2) yields point estimates of FGT poverty measures for the population and sub-groups. Obtaining consistent estimates of the standard errors of the poverty measures is less obvious, because the poverty lines, as well as the welfare metric, are built from a series of estimates of population characteristics from the survey data (e.g., expenditure patterns that determine the basic needs food bundles, age and sex distributions that determine food energy requirements). Given this complexity, estimating standard errors of the poverty measures analytically is intractable, so we use bootstrap methods (Efron, 1979; Efron and Tibshirani, 1993). The bootstrap samples mimic the stratified cluster sample design of the IAF survey. The estimated poverty lines, poverty headcount, and poverty gap are calculated for each bootstrap sample, with 1,000 replications.

Table 1 summarizes the process of estimating the poverty lines and poverty measures. The first column lists components of *nominal* consumption for each household, which is done prior to the bootstrap loop, as this measurement is largely independent of the particular sample drawn.[1] The second column contains processes undertaken within the bootstrap loop, including the estimation of poverty lines and poverty measures for each replication. The third column shows post-bootstrap processing, which is simply the calculation of the standard deviations of the poverty lines and measures over the 1,000 replications.

4. Results

Table 2 presents the 13 region-specific poverty lines. The variation in the cost of basic needs is considerable across regions, with costs tending to be higher in urban areas and southern provinces. Table 2 also shows the bootstrap-estimated standard errors of the total poverty line. These range from 4 to 14 percent of the point estimates, with most between 4 and 8 percent.

Table 3 presents estimates of the headcount index at the national and sub-national levels. The national headcount ratio is 54 percent, ranging from 36 percent in Sofala province to 81 percent in Inhambane province. The column showing standard errors without poverty line error uses the Howes and Lanjouw (1998) approach, which includes complex sample design effects and is the method used most often in the current literature. At higher levels of aggregation (e.g., national, rural, urban), the standard errors are 2 to 4 percent of the point estimate. As sample size decreases with disaggregation, the standard errors reach as high as 11 percent of the point estimates.

---

[1] Hedonic regressions were used to impute use-values for owner-occupied housing. Although the value obtained depends upon the sample, nominal use-values (rent foregone) for owner-occupied housing is in principle observable at the household level. The poverty line, in contrast, is not. Based on this distinction, we elect to treat estimates of use-value for owner occupied housing as data.

The next to last column of Table 3 shows the bootstrapped standard errors that include the sampling error of the poverty lines. These standard errors are larger in all instances but two. As seen in the rightmost column, the standard error of the national headcount is 27 percent higher when poverty line sampling error is included. For other levels of aggregation, including the poverty line as a source of variation increases the standard error of the headcount estimate from effectively zero to more than 33 percent in Gaza province. On average, including the poverty line sampling error increases the standard errors of the poverty headcount by about 15 percent.

Table 4 shows the poverty gap results. At each level of aggregation the standard error of the poverty gap is larger (relative to the point estimate) than in Table 3, consistent with Kakwani's (1993) observation that the precision of FGT poverty measures tends to decrease for higher levels of α. On average, the inclusion of poverty line sampling error increases the standard errors of the poverty gap estimates by about 17 percent.[2]

How important is the increase in standard errors of the estimated poverty measures when poverty line sampling error is included? To put it in the context of the existing literature, Howes and Lanjouw (1998) found that accounting for sample stratification and clustering increased the standard errors of estimated FGT poverty measures by 26 to 33 percent in Pakistan and 45 to 64 percent in Ghana. Adding the poverty lines as a source of error increases the standard errors of the national-level poverty estimates in Mozambique by 27 to 29 percent. This suggests that accounting for poverty line sampling error may be nearly as important quantitatively as accounting for complex sample design. Results from other countries, and using alternative methods

[2] Similar results are obtained for the FGT P_2 index (available from the authors upon request).

of setting the poverty lines, would be needed before drawing a firm conclusion in this regard. It should also be noted that there is no conflict between incorporating sample design and including poverty line error. Rather, it is advisable to do both.

5. Conclusions

Poverty reduction is a fundamental objective of economic development, and the success of policies, programs, and donor support is increasingly judged in terms of poverty reduction. As most poverty estimates come from sample survey data, the statistical properties of poverty measures are important when making poverty comparisons.

Although relative poverty studies have noted the sampling error associated with both the welfare metric and relative poverty lines calculated from survey data, this recognition has not extended to absolute poverty lines. This paper addresses this gap by proposing a general method for including the sampling error of poverty lines in the standard error of poverty measures. The approach is based on bootstrap methods that can be similarly applied to various methods of setting poverty lines (e.g., CBN, FEI) and to various poverty measures. Using recent data from Mozambique, we estimate that accounting for the sampling error of poverty lines increases the standard errors of FGT poverty measures by an average of about 15 percent, with the standard errors increasing by up to 34 percent for some sub-groups.

Are there circumstances in which one can safely ignore the sampling error of poverty lines, and treat them as fixed constants? In our view, the only such situation would be poverty lines that are determined exogenously, without reference to survey data. As absolute poverty lines should reflect the same standard of living across the domain of comparisons, and the cost of acquiring basic needs inevitably varies spatially and temporally, it is highly improbable that one could divine utility-consistent poverty lines without reference to data. Given a choice between arbitrarily specifying poverty lines that are certain to be utility-inconsistent to an unknown degree, or accepting a measurable loss in precision by estimating poverty lines from available data, the latter has clear advantages.

Stochastic dominance approaches, which make poverty comparisons across a range of plausible poverty lines (Atkinson, 1987), are also not automatically exempt from considering the sampling error of poverty lines when making statistical inferences. Poverty lines are not only a dividing line (admittedly artificial) between the poor and nonpoor, but also serve as cost of living indices, permitting interpersonal welfare comparisons when the cost of acquiring basic needs varies over time or space (Ravallion, 1998). If stochastic dominance analyses use poverty lines or cost of living indices estimated from survey data to facilitate comparisons, then the associated sampling error should be included in the confidence interval around the empirical cumulative distribution function, which will affect the precision of poverty comparisons.[3] Adapting the methods presented in this paper to stochastic dominance approaches to poverty comparisons is an area for future research.

[3] Likewise, because the dollar-a-day poverty line is based in part on statistically estimated purchasing power parity (PPP) calculations, it is not immune from the poverty line sampling error described in this paper.

6. References

Arndt, C. and K. Simler, 2005. Estimating utility-consistent poverty lines. FCND Discussion Paper No. 189. International Food Policy Research Institute, Washington, D.C.

Atkinson, A. B., 1987. On the measurement of poverty. *Econometrica*. 55, 749–764.

Bishop, J. A., K. V. Chow, and B. Zheng, 1995. Statistical inference and decomposable poverty measures. *Bulletin of Economic Research*. 47, 329–340.

Deaton, A. and S. Zaidi, 2002. Guidelines for constructing consumption aggregates for welfare analysis. LSMS Working Paper No. 135. World Bank, Washington, D.C.

Efron, B., 1979. Bootstrap methods: Another look at the jackknife. *Annals of Statistics*. 7, 1–26.

Efron, B. and R. Tibshirani, 1993. *An introduction to the bootstrap*. Chapman and Hall, New York.

Foster, J. E., J. Greer, and E. Thorbecke, 1984. A class of decomposable poverty measures *Econometrica*. 52, 761–766.

Greer, J. and E. Thorbecke, 1986. A methodology for measuring food poverty applied to Kenya. *Journal of Development Economics*. 24, 59–74.

Howes, S. and J. O. Lanjouw, 1998. Does sample design matter for poverty rate comparisons? *Review of Income and Wealth*. 44, 99–109.

INE (Instituto Nacional de Estatística), 2004. Inquérito aos Agregados Familiares Sobre Orçamento Familiar – Relatório Final. Instituto Nacional de Estatística, Maputo.

Kakwani, N., 1993. Statistical inference in the measurement of poverty. *Review of Economics and Statistics*. 75, 632–639.

Preston, I., 1995. Sampling distributions of relative poverty statistics. *Applied Statistics*. 44, 91–99.

Ravallion, M., 1994a. Poverty rankings using noisy data on living standards. *Economics Letters*. 45, 481–485.

———, 1994b. *Poverty comparisons*. Harwood Academic Publishers, Chur, Switzerland.

———, 1998. Poverty lines in theory and practice. LSMS Working Paper No. 133. World Bank, Washington, D.C.

Sen, A., 1976. Poverty: An ordinal approach to measurement. *Econometrica*. 44, 219–231.

Tarp, F., K. R. Simler, C. Matusse, R. Heltberg, and G. Dava, 2002. The robustness of poverty profiles reconsidered. *Economic Development and Cultural Change*. 51, 77–108.

WHO (World Health Organization), 1985. Energy and protein requirements. Technical Report Series No. 724. Geneva.

Zheng, B., 1997. Statistical inference for poverty measures with relative poverty lines. Luxembourg Income Study Working Paper No. 167, Luxembourg.

———, 2001. Statistical inference for poverty measures with relative poverty lines. *Journal of Econometrics*. 101, 337–356.

Table 1
Outline of calculations included and excluded from bootstrap procedure

Data collected or calculated before applying bootstrap	Calculations included in the bootstrap loop	Post-bootstrap calculations
Household food and nonfood consumption expenditure	Identification of poorest households	Standard deviation of estimated poverty measures over all replications as consistent estimator of standard error of poverty measures
Value of consumption of home-produced items	Average household composition and calorie requirements per person	
Value of transfers received	Intra-survey temporal price index	
Use-value of durable assets	Composition and cost of food poverty line bundles	
Use-value of owner-occupied housing	Bundles that satisfy revealed preference conditions	
	Nonfood budget share and poverty line	
	Total region-specific poverty lines	
	Poverty measures	

Table 2
Region-specific poverty lines, Mozambique 2002–03

	Poverty line (Meticais per person per day)			
Poverty line region	Food	Nonfood	Total	Standard error of total poverty line[1]
Rural Niassa and Cabo Delgado	5,434	1,665	7,099	274
Urban Niassa and Cabo Delgado	7,540	2,690	10,231	1,082
Rural Nampula	4,471	1,501	5,972	425
Urban Nampula	4,853	1,807	6,661	947
Rural Sofala and Zambézia	4,155	1,318	5,473	330
Urban Sofala and Zambézia	6,591	2,183	8,775	671
Rural Tete and Manica	5,629	1,304	6,933	482
Urban Tete and Manica	7,145	2,545	9,690	714
Rural Inhambane and Gaza	6,614	2,394	9,008	388
Urban Inhambane and Gaza	7,264	3,457	10,721	467
Rural Maputo Province	11,801	4,963	16,764	1,246
Urban Maputo Province	11,898	6,398	18,296	644
Maputo City	12,224	7,291	19,515	519

Source: Authors' calculations from the 2002–03 IAF.
[1] Estimated by bootstrapping with 1,000 replications.

Table 3
Estimates of poverty headcount index and standard errors, Mozambique 2002–03

Region	N	Headcount index	Standard error: Without poverty line error	Standard error: With poverty line error[1]	Ratio of standard errors
National	8,700	0.5407	0.0136	0.0173	1.27
Urban	4,005	0.5147	0.0225	0.0259	1.15
Rural	4,695	0.5529	0.0168	0.0206	1.23
Northern	2,310	0.5528	0.0257	0.0321	1.25
Central	3,100	0.4551	0.0240	0.0282	1.18
Southern	3,290	0.6654	0.0135	0.0167	1.24
Niassa	816	0.5211	0.0544	0.0553	1.02
Cabo Delgado	738	0.6315	0.0341	0.0366	1.07
Nampula	756	0.5261	0.0382	0.0482	1.26
Zambézia	733	0.4455	0.0460	0.0500	1.09
Tete	756	0.5980	0.0422	0.0416	0.99
Manica	816	0.4355	0.0411	0.0409	1.00
Sofala	795	0.3613	0.0276	0.0350	1.27
Inhambane	753	0.8068	0.0216	0.0240	1.11
Gaza	786	0.6014	0.0260	0.0347	1.33
Maputo Province	828	0.6927	0.0283	0.0296	1.05
Maputo City	923	0.5360	0.0309	0.0315	1.02

Source: Authors' calculations from the 2002–03 IAF.
[1] Estimated by bootstrapping with 1,000 replications.

Table 4
Estimates of poverty gap index and standard errors, Mozambique 2002–03

			Standard error		
Region	N	Poverty gap	Without poverty line error	With poverty line error[1]	Ratio of standard errors
National	8,700	0.2051	0.0065	0.0084	1.29
Urban	4,005	0.1969	0.0097	0.0118	1.22
Rural	4,695	0.2090	0.0084	0.0102	1.21
Northern	2,310	0.1949	0.0114	0.0153	1.34
Central	3,100	0.1603	0.0110	0.0129	1.17
Southern	3,290	0.2913	0.0099	0.0118	1.19
Niassa	816	0.1583	0.0150	0.0169	1.13
Cabo Delgado	738	0.2162	0.0168	0.0187	1.11
Nampula	756	0.1953	0.0178	0.0229	1.29
Zambézia	733	0.1400	0.0194	0.0218	1.12
Tete	756	0.2630	0.0249	0.0267	1.07
Manica	816	0.1678	0.0274	0.0257	0.94
Sofala	795	0.1067	0.0107	0.0133	1.24
Inhambane	753	0.4221	0.0221	0.0244	1.10
Gaza	786	0.2062	0.0135	0.0164	1.21
Maputo Province	828	0.3111	0.0186	0.0205	1.10
Maputo City	923	0.2086	0.0148	0.0165	1.11

Source: Authors’ calculations from the 2002–03 IAF.

[1] Estimated by bootstrapping with 1,000 replications.