Review of Income and Wealth
Series 53, Number 2, June 2007
doi: 10.1111/j.1475-4991.2007.00228.x

POVERTY COMPARISONS WITH ABSOLUTE POVERTY LINES ESTIMATED FROM SURVEY DATA

BY KENNETH R. SIMLER*

Poverty Reduction Group, The World Bank, Washington, DC

AND

CHANNING ARNDT

Department of Agricultural Economics, Purdue University, West Lafayette, Indiana

The objective of measuring poverty is usually to make comparisons over time or between two or more groups. Common statistical inference methods are used to determine whether an apparent difference in measured poverty is statistically significant. Studies of relative poverty have long recognized that when the poverty line is calculated from sample survey data, both the variance of the poverty line and the variance of the welfare metric contribute to the variance of the poverty estimate. In contrast, studies using absolute poverty lines have ignored the poverty line variance, even when the poverty lines are estimated from sample survey data. Including the poverty line variance could either reduce or increase the precision of poverty estimates, depending on the specific characteristics of the data. This paper presents a general procedure for estimating the standard error of poverty measures when the poverty line is estimated from survey data. Based on bootstrap methods, the approach can be used for a wide range of poverty measures and methods for estimating poverty lines. The method is applied to recent household survey data from Mozambique. When the sampling variance of the poverty line is taken into account, the estimated standard errors of Foster–Greer–Thorbecke and Watts poverty measures increase by 15–30 percent at the national level, with considerable variability at lower levels of aggregation.

1. INTRODUCTION

A principal objective of poverty measurement is to make comparisons between groups. Analysts and policymakers are generally interested less in the absolute level of poverty at a given place and time than they are in knowing how measured poverty levels compare to levels observed in other settings or at other points in time. Is poverty higher in the hills or on the coast? Did poverty decline following implementation of a poverty reduction program? These questions have gained an even higher profile in recent years. Besides the high profile Millennium Development Goal of halving world poverty by 2015, country development

Note: The authors gratefully acknowledge the helpful comments and suggestions of the journal's editor and two anonymous referees. We received constructive comments on earlier drafts of this paper from Daniel Gilligan, John Hoddinott, Dean Jolliffe, and Martin Ravallion. The authors are solely responsible for any remaining errors. This research was supported by the Mozambique Ministry of Planning and Development, with financing from Danish International Development Assistance (Danida), the United Kingdom's Department for International Development (DFID), and the Swiss Agency for Development and Cooperation (SDC). These are the views of the authors and do not necessarily reflect those of the Government of Mozambique, the World Bank, or any affiliated organization.

*Correspondence to: Kenneth R. Simler, Poverty Reduction Group, The World Bank, 1818 H Street, NW, Washington, DC 20433, USA (ksimler@worldbank.org).

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. For example, indications of a recent increase in poverty in Uganda sparked a debate about whether the national development strategy, which had steadily reduced poverty in the 1990s, needed an overhaul (Kappel *et al.*, 2005).

There are many ways to define and measure poverty, but with few exceptions the empirical basis for poverty comparisons is statistical, employing point estimates of relevant poverty measures and their associated standard errors. These are generally estimated from household survey data. Statistical tests are applied to assess whether differences or changes in poverty levels are significant. Research over the past 15 years has increasingly refined statistical inference methods for poverty measures. Kakwani (1993) develops distribution-free asymptotic standard errors for several additively decomposable poverty measures. Bishop *et al.* (1995) provide asymptotic theory for testing poverty measures decomposed by sub-group. Ravallion (1994a) examines the effect of errors in consumption data on poverty comparisons, finding that noisier data for some sub-groups can lead to re-rankings of poverty measures, with the exact nature of the re-ranking dependent upon the poverty measure used. These approaches assume that the data are generated by simple random sampling, but Howes and Lanjouw (1998) note that most poverty data come from stratified cluster sample surveys. Using data from Pakistan and Ghana, they find that when complex sample design is taken into account, estimated standard errors of FGT poverty measures increase by 26–33 percent in Pakistan and 45–64 percent in Ghana.

Preston (1995) demonstrates that the precision of poverty estimates depends not only on the sampling properties of the welfare measure, but also on the error associated with the poverty line itself. He presents standard error formulae for poverty measures that incorporate simple random sampling error in relative poverty lines (based on sample quantiles) as well as the welfare measure. He observes that the two sources of error could reinforce or offset one another. While the sampling error of the poverty line itself will increase the sampling error of the poverty measure, this increase will be dampened if the covariance of the welfare measure and the poverty line is positive. If the covariance is sufficiently positive, the net effect could be to reduce the sampling error of the poverty measure. Thus, one cannot say *a priori* whether accounting for sampling error in the poverty line will increase or reduce the precision of poverty estimates. Zheng (2001) builds on this work to develop analytical expressions for asymptotic distribution-free inference applicable to several additively decomposable poverty measures when relative poverty lines are set as percentages of mean income or percentages of quantiles, and allows for cluster sampling. In his empirical applications with relative poverty lines, Zheng (1997, 2001) finds that the sampling error of the poverty line always increases the standard error of poverty estimates.

Zheng (2001) states that the sampling variability of poverty lines is only relevant for relative poverty measures, asserting that absolute poverty lines are not estimated from sample survey data. However, a review of the absolute poverty literature shows that absolute poverty lines are routinely estimated from sample survey data, especially in low income countries over the past 10–20 years. For example, influential articles on estimation of absolute poverty lines have been based on survey data (Greer and Thorbecke, 1986; Ravallion and Bidani, 1994). Similarly, a recent "how to" manual by the World Bank Institute (World Bank, 2006) emphasizes the use of survey data for determining poverty lines, citing examples from empirical work in a wide range of low income countries. Yet, none of the absolute poverty studies consider the effect of poverty line variability on poverty estimates. This raises the possibility that the precision of most estimates of absolute poverty has been overstated.

This paper is structured as follows. Section 2 articulates the argument for incorporating the statistical error associated with absolute poverty lines in the calculation of standard errors of poverty measures. Section 3 proposes a method for estimating the sampling error of absolute poverty lines estimated from survey data. Based on bootstrapping, the method is extremely general and can be applied to a wide range of poverty lines and poverty measures. Section 3 also describes the Mozambican survey data used in the empirical analysis. Section 4 uses the methods and survey data to provide an initial estimate of the magnitude of the change in the standard errors of poverty measures when the sampling properties of the poverty lines are taken into consideration. Section 5 summarizes and concludes. For the case of Mozambique, we find that in the large majority of instances traditional methods underestimate standard errors for poverty measures, often substantially. We conclude that the variance of the poverty line should be accounted for when calculating standard errors of poverty measures and recommend the bootstrap approach as a practical means for doing so.

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 non-poor? Following Zheng (2000), one can formally write a generic additively decomposable poverty measure P as

$$P = \int_0^z \phi(y, z) dF(y), \tag{1}$$

where y is a money-metric welfare measure, z is a monetary poverty line, and ϕ is a poverty function that is decreasing in y, increasing in z, and homogenous of degree 0. The poverty line can be considered as the expenditure function that corresponds to a reference level of utility, u_z, which defines the poverty threshold, or

$$z = e(p, x, u_z), \tag{2}$$

where x is a vector of commodities consumed and p is the corresponding price vector.

It is possible to make interpersonal welfare comparisons over space or time by defining money-metric utility, or what Blackorby and Donaldson (1987) call the welfare ratio, as $y^* = y/e(p, x, u_z)$. The poverty measure can then be written in terms of money-metric utility as the definite integral

$$P = \int_0^1 \phi(y^*, 1) dF(y^*). \tag{3}$$

It bears noting that when the poverty line is estimated from sample survey data, y^* is the ratio of two random variables whose distribution functions are not known. Computing the variance of (3), where y^* is an argument, thus poses some challenges, which we discuss in greater detail in Section 3.

In empirical work, there are numerous options available with regard to constructing the welfare metric, setting the poverty lines, and computing poverty measures. For the purposes of the present illustration we use total consumption per capita (Deaton and Zaidi, 2002), Cost of Basic Needs (CBN) poverty lines (Ravallion, 1994b, 1998), and the poverty measures proposed by Watts (1968) and Foster *et al.* (1984).

Consider the Foster–Greer–Thorbecke (FGT), or P_α, class of poverty measures. At the household level, the general form of the FGT measure for household j can be written

$$P_{\alpha,j} = \left(\frac{z - y_j}{z}\right)^\alpha \cdot I, \quad \alpha \geq 0, \tag{4}$$

where I is an indicator function that takes the value 1 if $y_j < z$ and the value 0 if $y_j \geq z$. Poverty in a population of n households is the weighted mean of (4) over all households, with the number of members in each household (h_j) as the weights, or

$$P_\alpha = \frac{\sum_{j=1}^{n} h_j P_{\alpha,j}}{\sum_{j=1}^{n} h_j}. \tag{5}$$

The poverty headcount index and poverty gap index are obtained when $\alpha = 0$ and 1, respectively. In the case of non-self-weighting sample surveys, which is the typical source of poverty data, sample weights (or expansion factors) w_j must be employed to arrive at an unbiased estimator of individual-level poverty measures, written

$$P_\alpha = \frac{\sum_{j=1}^{n} w_j h_j P_{\alpha,j}}{\sum_{j=1}^{n} w_j h_j}. \tag{6}$$

By using h_j as weights, equations (5) and (6) assume that poverty is distributed equally within the household. Although this may be a strong assumption, it is difficult to avoid because individual-specific information on the welfare metric is rarely available. If such data were available then poverty could be measured by equation (6), but with j indexing individuals instead of households and $h_j = 1$.

Howes and Lanjouw (1998) define the estimator of the poverty measure in (6) as $\pi_\alpha = t/p$, where p (the denominator in (6)) is the sample estimate of the population size. The numerator, t, is the sample estimate of "total poverty," whose definition depends on the particular poverty measure in question. For example, when $\alpha = 0$ the numerator is the sample estimate of the number of persons below the poverty line. Likewise, when $\alpha = 1$, multiplying the numerator by the poverty line yields the sample estimate of the aggregate monetary poverty gap. Howes and Lanjouw (1998) show that under fairly weak assumptions that also conform well to the non-self-weighting stratified multiple stage cluster sampling procedures that are common among household living standards surveys, a Taylor series expansion provides a consistent estimator of the variance of π_α. More specifically, for survey stratum k, cluster c, and n_k cluster samples drawn in the survey sample, a consistent estimator of the variance of π_α is:

$$\text{(7)} \qquad V\hat{a}r(\pi_\alpha) = \frac{1}{p^2}\left[V\hat{a}r(t) + \pi_\alpha^2 V\hat{a}r(p) - 2\pi_\alpha C\hat{o}v(t, p)\right],$$

where

$$\text{(8)} \qquad V\hat{a}r(t) = \sum_{k=1}^{K} V\hat{a}r(t_k) = \sum_{k=1}^{K} \frac{1}{n_k(n_k - 1)} \sum_{c=1}^{n_k} (t_{kc} - t_k)^2,$$

$$\text{(9)} \qquad V\hat{a}r(p) = \sum_{k=1}^{K} V\hat{a}r(p_k) = \sum_{k=1}^{K} \frac{1}{n_k(n_k - 1)} \sum_{c=1}^{n_k} (p_{kc} - p_k)^2,$$

and

$$\text{(10)} \qquad C\hat{o}v(t, p) = \sum_{k=1}^{K} C\hat{o}v(t_k, p_k) = \sum_{k=1}^{K} \frac{1}{n_k(n_k - 1)} \sum_{c=1}^{n_k} (t_{kc} - t_k)(p_{kc} - p_k).$$

The crux of our argument goes back to equation (1). Whereas the welfare metric y 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. Standard absolute poverty analyses ignore this variance component, leading to incorrect estimates of the precision of poverty measures, and potentially misleading poverty comparisons over time and space.

The intuition of the argument is seen in Figure 1. In both panels of the figure, the horizontal axis is the welfare measure, the vertical axis is the proportion of the population, the dark curved line is the empirical cumulative density function (CDF) of the welfare measure (truncated at the upper end to focus on the region near the poverty line), and the vertical line labeled z is the poverty line. The dotted

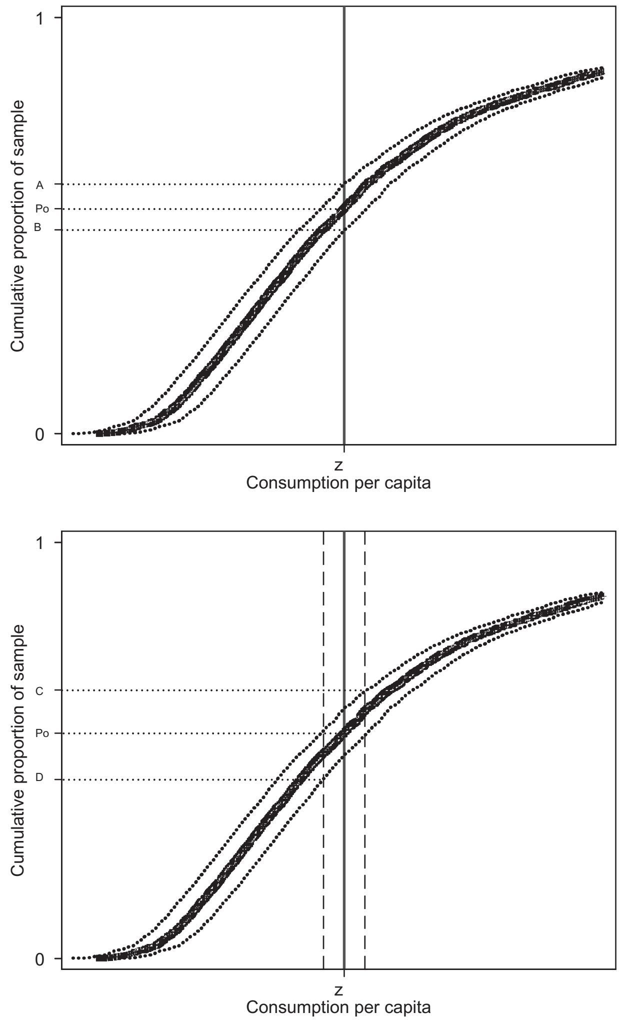

Figure 1. Illustration of Poverty Line Error's Contribution to Poverty Estimate Error

lines on either side of the CDF are an indicative confidence interval for the cumulative density of the welfare measure. The point estimate of the poverty headcount, $\hat{P}_0$, is read from the vertical axis, at the level where the poverty line intersects the CDF. If the poverty line is assumed to be fixed, then the confidence interval for $\hat{P}_0$ is the interval AB on the vertical axis, corresponding to where the upper and lower bounds of the CDF confidence interval intersect the poverty line.[1] In the lower panel of Figure 1, the assumption of a fixed poverty line is relaxed, and the dashed vertical lines represent the confidence interval around the poverty line. The potential confidence interval around $\hat{P}_0$ expands, as shown by the interval

[1]This is something of an oversimplification, but does capture the essence of the idea and is used here to illustrate the argument.

CD along the vertical axis of the lower panel. In practice, the overall effect on the precision of the estimate of $\hat{P}_0$ will depend on the joint distribution of the two random variables. For example, with strong positive correlations between the poverty line and the welfare metric the estimate of $\hat{P}_0$ could in fact be more precise when the variance of the poverty line is taken into account, as noted by Preston (1995).

Figure 1 also illustrates that whether or not the poverty line variance is included in the poverty measure's standard error, the precision of the poverty estimate also depends on the location of the poverty line. Poverty lines that are closer to the mode of the distribution, where the CDF has its steepest slope, will also tend to generate less precise poverty estimates. It should be noted that analogous illustrations could be employed for higher orders of the FGT poverty measures, for example, by using the poverty deficit curve (Ravallion, 1994b; Deaton, 1997) in place of the CDF for the estimate of the average poverty gap, P_1.

3. Data and Methods

This section describes our approach to incorporating the sampling error of the poverty line in estimates of standard errors of poverty measures using household data from Mozambique as a case study. Before describing the approach to calculating standard errors specifically, we first describe the data collection process, the definition of the welfare metric, and the setting of poverty lines. This is presented in some detail because the individual steps determine not only the point estimates of the poverty lines and poverty measures, but also the bootstrapped estimates of their standard errors.

Data Collection

We use data from the 2002–03 national Household Budget Survey in Mozambique, also known by its Portuguese abbreviation IAF (Inquérito aos Agregados Familiares sobre Orçamento Familiar). Additional details about the survey may be found in INE (2004). The survey was carried out from July 2002 through June 2003, visiting 8,700 households throughout the country. The sample had 21 strata: separate rural and urban strata for each of Mozambique's ten provinces, plus one for the capital city of Maputo. A two-stage procedure was used to select sample households. Within each stratum, primary sampling units (PSUs) (already defined on the basis of the 1997 Census) were selected with probability proportional to size. One month before the launch of the survey, the survey teams carried out a complete listing of all households in each of the 857 selected PSUs. In the second stage, households were randomly selected within each PSU, with 12 households per urban PSU and 9 households per rural PSU. The survey was limited to households residing in private residences, thus excluding those living in institutions (e.g. prisons, boarding schools, military barracks), diplomatic residences, and the homeless. The complex sampling structure implies unequal probability of selection across PSUs, so sampling weights were calculated as the inverse of the probability of selection.

The content of the 2002–03 IAF is similar to that of other household budget surveys conducted in low income countries. Households were visited by interviewers at least three times over a seven day period. On the first visit, the interviewer and household completed the module on general household characteristics, and collected consumption information on food and selected common nonfood items with reference to the preceding day (purchases, consumption from home production, and in kind transfers received). On subsequent visits other parts of the questionnaire were completed (monthly expenditures, annual expenditures, and income), as well as daily consumption information for the period since the previous interview.

Definition of the Welfare Metric

The approach used to calculate consumption follows closely the one described by Deaton and Zaidi (2002) and Deaton and Grosh (2000), drawing from several modules of the IAF. It measures the total value of consumption of food and nonfood items (including purchases, home-produced items, and gifts received), as well as imputed use-values for owner-occupied housing and household durable goods. Market purchases were valued at the price paid, whereas non-market purchases were valued at the prevailing market price in the area at that time. The only two significant omissions from the consumption measure—both because of lack of data—are consumption of commodities supplied by the public sector free of charge (or the subsidized element in such commodities) and consumption of home produced services. For example, an all-weather road, or a public market, or a public water tap, presumably enhances the well-being of the people who use those facilities. Similarly, home produced services, such as cooking and cleaning, also add to welfare. These are often not captured by household surveys.

Food prices tend to follow a seasonal pattern, which implies that the purchasing power of a given amount of money varies during the year. For example, to acquire the same amount of food, a given household might have to spend twice as much in January as it spends in June. If the household consumed the same amount in real (quantity) terms in those months, it would appear to have a higher standard of living in January in nominal monetary terms. To avoid this kind of inconsistency, 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 to take these price fluctuations into account.

As larger households tend to have higher subsistence requirements than smaller households, we divide total household consumption by household size and use consumption per capita in our poverty comparisons. Alternative normalizations or equivalency scales exist, but the per capita scale is sufficient for the purposes of the present analysis. Adapting the method to other equivalence scales is straightforward.

Setting Poverty Lines

Poverty lines were set using the CBN approach (Ravallion, 1994b). Mozambique is a large country with poorly developed infrastructure and markets. High transactions costs, combined with wide variation in agro-climatic conditions and production costs, lead to wide spatial and temporal variation in the prices of basic goods. In particular, differences in relative prices across space and time affect not only the total cost of acquiring basic needs, but also the composition of the basic needs bundle, as households adjust their consumption patterns in response to differences in relative prices.

As absolute poverty lines are supposed to represent the cost of achieving the same standard of living across the domain of comparisons, it is necessary to establish region-specific poverty lines. To define the poverty lines, the country was divided into 13 regions, based on an aggregation of the 21 survey strata that preserved the distinction between rural and urban areas, but grouping adjacent strata with similar characteristics (especially food prices and consumption patterns) if they had relatively few observations.

For each poverty line region, the food poverty line is constructed by an iterative procedure that determines the 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). The food poverty line—expressed in monetary cost per person per day—is the region-specific cost of meeting the caloric requirements when consuming a food bundle comprising goods that the poor in the region actually consume.[2] It bears emphasizing that the food bundle is not determined by an externally imposed least-cost diet, but rather by the food consumption characteristics of poor households as recorded in the survey, which contributes to the sampling error of the poverty line.

The decision to allow the basic needs food bundles to vary by region was driven by the large differences in relative food prices across the 13 poverty line regions, and corresponding consumer behavior consistent with cost minimization. Within the 13 poverty line regions, relative prices and consumption patterns are fairly homogeneous. Ravallion (1998) and Tarp *et al.* (2002) present arguments that allowing the food bundle to vary by region can result in more consistent poverty comparisons than using a fixed national bundle. Recent poverty studies that use region-specific poverty bundles and prices include Tarp *et al.* (2002), Mukherjee and Benson (2003), Gibson and Rozelle (2003), Ravallion and Lokshin (2006), Datt and Jolliffe (2005), and Arndt and Simler (2005). Note that the same arguments in favor of allowing the bundle to vary over space can also be applied to comparisons over time.

The relevant food bundles and associated prices were estimated for relatively poor households using the iterative procedure described by Ravallion (1998). All households were ranked in descending order by nominal consumption per capita,

[2]The typical food bundle of the poor may contain more or less calories than the requirement for that region. This bundle is then proportionally scaled up or down until it yields exactly the pre-established caloric requirement, and the cost of this rescaled bundle at region-specific prices determines the food poverty line for that region. Also, it is recognized that food energy is only one facet of human nutrition, and that adequate consumption of other nutrients, such as protein, iron, vitamin A, and so forth, is also essential for a healthy and active life. However, like most multipurpose household surveys, the information on food consumption in the IAF data set is not sufficiently detailed to permit estimation of the intake and absorption of other nutrients. Use of energy requirements alone is also well established in the poverty measurement literature (Greer and Thorbecke, 1986; Ravallion, 1994b, 1998).

with the bottom X percent identified as the relatively poor. The cutoff point may be considered as a preliminary estimate of the poverty headcount, and can be chosen based on past poverty assessments or other information. Preliminary poverty line bundles were constructed using the consumption patterns of the relatively poor, and the nominal consumption values converted to real terms (i.e. taking into account region-specific differences in the cost of acquiring the basic needs bundles). Households were then re-ranked using this first approximation of consumption per capita in real terms; households in regions with high (low) price levels are poorer (richer) than indicated by nominal consumption, and thus move down (up) when ranked in real terms. Revised food poverty line bundles were constructed, producing a second estimate of food poverty lines, by which the households were re-ranked again. The iterative process continues until it converges, meaning that the same, or nearly the same, sub-sample of households appears below the cutoff point on successive iterations. We experimented with several starting values, ranging from 40 to 65 percent, and found that all tended to converge on 48 percent (the poverty headcount ratio), with convergence occurring after four or five iterations. This implies that the poverty headcount estimate is robust to the choice of population sub-group that is used to construct the initial food poverty line bundles.

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. In practice, the average daily food energy requirement varies little across the 13 regions, averaging approximately 2,150 kilocalories per person.

Whereas physiological needs provide the conceptual underpinning of the food poverty lines, no similar basis is readily available for defining nonfood needs. In virtually all settings, even very poor households allocate a sizeable proportion of their total consumption to nonfood items, such as shelter and clothing. We estimate the nonfood poverty line by examining the proportion of total consumption allocated to nonfoods among those households whose total expenditure is approximately equal to the region-specific food poverty line (Ravallion, 1994b, 1998; Ravallion and Bidani, 1994). The logic is that if a household's total consumption is only sufficient to purchase the minimum amount of calories using a food bundle typical for the poor, any expenditure devoted to nonfoods is clearly a basic need, as it is displacing expenditure on basic food items. Specifically, we estimate the nonfood component of the poverty line as the average nonfood budget share of households whose total consumption is between 80 and 120 percent of the food poverty line, using a triangular kernel to give more weight to those households closer to 100 percent of the food poverty line.

Estimating Poverty Measures and Their Standard Errors

After calculating the consumption variable and estimating the region-specific poverty lines, obtaining point estimates of FGT poverty measures for the population and sub-groups requires nothing more than application of equation (6) to the survey data. Obtaining consistent estimates of the standard errors of the poverty measures is less obvious.

It should be clear from the description of constructing the poverty lines that the poverty lines, as well as the welfare metric, are built from a series of estimates of population characteristics from the sample survey data. Food energy requirements are based on survey estimates of the population's age and sex distributions. The expenditure patterns that determine the basic needs food bundles are also estimates that are subject to sampling error, as are the nonfood budget shares that determine the nonfood poverty line. Similarly, the prices used to estimate the cost of the basic needs bundles come from the survey. In this light, it seems difficult to justify the common assumption that the poverty lines are not a source of sampling error in poverty estimates.

Given the complexity of the construction of the poverty lines, deriving standard errors of the poverty measures analytically is intractable, so we estimate them via a bootstrapping procedure. Bootstrapping is a general means of generating consistent estimates of an estimator's sampling distribution when an analytical solution cannot be derived or requires unreasonable assumptions (Efron, 1979; Efron and Tibshirani, 1993). It is based on repeated (J times) samples, drawn with replacement, of size K from the original sample data, of size N, where $K \leq N$. As the original sample size, N, increases, the bootstrap approach converges to Monte Carlo for fixed K. The primary assumption behind the bootstrap is that the distribution of the observed sample is a good approximation of the distribution of the population.

In our application, the bootstrap samples are drawn in a manner that mimics the stratified cluster sample design of the IAF survey. That is, within each stratum, K clusters are randomly drawn, with replacement, where K is also the number of primary sampling units in the stratum (i.e. $K = N$). When a cluster is drawn, all of the households in that cluster are drawn. Because the bootstrap sampling is done with replacement, each cluster (and household) may appear one or more times in a given bootstrap sample, or not at all. The estimated poverty lines, poverty headcount, poverty gap, and Watts index are calculated for each bootstrap sample. The process is repeated $J = 1{,}000$ times. The standard deviation of a poverty measure over the 1,000 bootstrap replications is an estimator of the standard error of that poverty measure. The point estimates of the poverty measures are calculated from the original, non-bootstrapped sample (Efron and Tibshirani, 1993).

The process of estimating the poverty lines and poverty measures in each bootstrap replication is summarized in Table 1, which is divided into three columns. The first column lists processes that can be undertaken prior to the bootstrap loop. The calculation of *nominal* consumption per capita for each household occurs at this step as this measurement is (almost entirely) independent of the particular sample drawn.[3] The second column contains processes undertaken within the bootstrap loop. These are the steps described earlier for calculating the

[3]In the Mozambique case, hedonic regressions were used to impute use-values for owner-occupied housing. Obviously, the value obtained then depends upon the sample. Nevertheless, 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.

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 an estimator of the 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	Nonfood budget share and poverty line	
	Total region-specific poverty lines	
	Poverty measures	

poverty lines and the point estimates for the poverty measures for each bootstrap sample. The third column shows post-bootstrap processing, which is simply the calculation of the standard deviations of the poverty measures over the bootstrap replications.

4. Results

The 13 region-specific food, nonfood, and total poverty lines are shown in Table 2. The variation in the cost of basic needs is considerable across regions. Some general patterns are evident, such as the higher poverty lines in urban areas of a given province or province grouping, and the tendency for the poverty lines to increase (within urban and rural zones) as one moves down the list, which is roughly ordered from northern provinces to southern provinces.[4] Table 2 also shows the estimated standard errors of the total poverty line, estimated via the bootstrap process described earlier with 1,000 replications. The poverty line standard errors range from 3 to 19 percent of the point estimates, with most of them between 5 and 10 percent.

Table 3 presents estimates of the poverty headcount index at the national level and for several sub-national groupings. The national headcount ratio is 48 percent,

[4]It should be noted that these poverty lines, and the poverty measures presented in Tables 3 and 4, differ from the official poverty lines reported elsewhere (MPF, 2004; Arndt and Simler, 2005). The official poverty lines include a relatively novel entropy estimation adjustment to ensure that the basic needs food bundles satisfy revealed preference conditions across regions and over time. While we believe that revealed preference consistent poverty lines yield superior poverty measures, we elect to omit the revealed preference adjustment procedure in this presentation in order to focus on a commonly used approach for measuring poverty. It is straightforward to include the revealed preference adjustment procedure in the calculation of standard errors. With the Mozambique data the resulting standard errors are on average only slightly smaller than the results presented here. These results are available upon request.

TABLE 2

Region-Specific Food, Nonfood, and Total Poverty Lines for Mozambique, 2002–03

Poverty Line Region	Poverty Line (Meticais per person per day)			
	Food	Nonfood	Total	Standard Error of Total Poverty Line[1]
Rural Niassa and Cabo Delgado	4,756	1,532	6,288	282
Urban Niassa and Cabo Delgado	7,717	2,838	10,555	1,164
Rural Nampula	2,752	913	3,665	399
Urban Nampula	3,749	1,370	5,119	982
Rural Sofala and Zambézia	3,548	1,195	4,743	302
Urban Sofala and Zambézia	5,902	2,177	8,079	750
Rural Tete and Manica	6,937	1,456	8,393	598
Urban Tete and Manica	9,656	3,575	13,231	1,056
Rural Inhambane and Gaza	5,438	1,930	7,368	497
Urban Inhambane and Gaza	6,613	3,025	9,638	762
Rural Maputo Province	12,584	5,385	17,969	1,755
Urban Maputo Province	13,741	7,810	21,551	1,467
Maputo City	13,211	8,022	21,232	694

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

and ranges from 30 percent in Nampula province to 76 percent in Maputo province. The column showing standard errors without poverty line error uses the Howes and Lanjouw (1998) method described in Section 2, which includes complex sample design effects and is the method used most often in the current literature. At higher levels of aggregation, such as the national level or estimates for rural and urban areas, the standard errors are 2–3 percent of the point estimate. As sample size decreases with further disaggregation, the standard errors reach as high as 12 percent of the point estimates, although some of the provincial estimates are still fairly precise (e.g. Inhambane and Maputo provinces).

The next to last column of Table 3 shows the standard errors including the sampling error of the poverty lines, as estimated using the bootstrap procedure described in the preceding section. These standard errors are larger in all instances, despite the possibility of poverty line error offsetting the error in the welfare measure that Preston (1995) described. As seen in the rightmost column, the standard error of the national headcount is 18 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 a negligible amount in Niassa province to over 60 percent in Gaza province. On average, including the poverty line sampling error increases the estimated standard errors of the sub-national poverty headcount estimates by about 20 percent.

Table 4 shows the same set of results for the poverty gap index. At each level of aggregation the standard errors of the poverty gap index are larger relative to the point estimate than is observed for the headcount index. This is

TABLE 3

ESTIMATES OF POVERTY HEADCOUNT INDEX (P_0) AND STANDARD ERRORS, MOZAMBIQUE 2002–03

Region	Sample Size	Headcount Index	Standard Error: Without Poverty Line Error	Standard Error: With Poverty Line Error[1]	Ratio of Standard Errors
National	8,700	0.4796	0.0128	0.0151	1.18
Rural	4,695	0.4586	0.0165	0.0201	1.22
Urban	4,005	0.5239	0.0231	0.0252	1.09
Regions					
Northern	2,310	0.3977	0.0237	0.0322	1.36
Central	3,100	0.4456	0.0223	0.0259	1.16
Southern	3,290	0.6381	0.0146	0.0218	1.49
Provinces					
Niassa	816	0.4559	0.0501	0.0503	1.00
Cabo Delgado	738	0.5708	0.0355	0.0401	1.13
Nampula	756	0.3047	0.0349	0.0492	1.41
Zambézia	733	0.3514	0.0428	0.0443	1.04
Tete	756	0.7080	0.0377	0.0434	1.15
Manica	816	0.5853	0.0412	0.0465	1.13
Sofala	795	0.3093	0.0280	0.0360	1.29
Inhambane	753	0.7509	0.0250	0.0333	1.33
Gaza	786	0.4709	0.0266	0.0429	1.61
Maputo Province	828	0.7591	0.0277	0.0303	1.10
Maputo City	923	0.5804	0.0325	0.0339	1.04

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

TABLE 4

Estimates of Poverty Gap Index (P_1) and Standard Errors, Mozambique 2002–03

Region	Sample Size	Poverty Gap Index	Standard Error		
			Without Poverty Line Error	With Poverty Line Error[1]	Ratio of Standard Errors
National	8,700	0.1754	0.0058	0.0074	1.27
Rural	4,695	0.1644	0.0079	0.0094	1.20
Urban	4,005	0.1986	0.0101	0.0130	1.29
Regions					
Northern	2,310	0.1160	0.0081	0.0133	1.63
Central	3,100	0.1627	0.0102	0.0119	1.17
Southern	3,290	0.2709	0.0102	0.0156	1.53
Provinces					
Niassa	816	0.1266	0.0122	0.0150	1.22
Cabo Delgado	738	0.1796	0.0162	0.0177	1.09
Nampula	756	0.0846	0.0126	0.0208	1.65
Zambézia	733	0.1017	0.0161	0.0159	0.99
Tete	756	0.3361	0.0268	0.0298	1.11
Manica	816	0.2439	0.0283	0.0322	1.14
Sofala	795	0.0785	0.0094	0.0125	1.33
Inhambane	753	0.3519	0.0233	0.0325	1.39
Gaza	786	0.1421	0.0121	0.0199	1.65
Maputo Province	828	0.3612	0.0201	0.0258	1.28
Maputo City	923	0.2364	0.0159	0.0172	1.08

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

consistent with Kakwani's (1993) observation that the precision of FGT poverty measures (measured as the standard error divided by the point estimate) tends to decrease for higher levels of α, a finding that is corroborated by the results of Howes and Lanjouw (1998). Comparing the standard errors estimated with and without poverty line sampling error we see that in Zambézia province, including the poverty line sampling error marginally reduces the total standard error of the poverty gap estimate. For all other estimates, the poverty line error increases the standard error of the poverty gap estimate, in some cases by as much as two-thirds. On average, the inclusion of poverty line sampling error increases the standard errors of the poverty gap estimates by about 30 percent, considerably more than the increase observed for the poverty headcount index.

As noted earlier, the issue of poverty line variance is not limited to the FGT class of measures, and the method presented here is readily adapted to other poverty measures. The Watts index (Watts, 1968) is one of the earliest summary measures of poverty. Although it is not as widely used as the FGT class of poverty measures, it has been noted for its favorable theoretical properties (Zheng, 1993), and has also received considerable attention recently in the "pro-poor growth" literature (see, for example, Ravallion, 2004). The Watts index, W, may be written as

$$W = \int_0^H \log[z/y(p)]dp, \tag{11}$$

where H is the headcount index and $y(p)$ is the quantile function, which is the inverse of the cumulative distribution function $p = F(x)$ at the p-th quantile. The Watts index was estimated using the Mozambican data, with the results presented in Table 5. These are qualitatively similar to the poverty gap results in Table 4. At the national level, the standard error for the Watts index is 30 percent larger when the poverty line error is included. In most cases, incorporation of the poverty line error increases the standard error of the estimates of the Watts index, varying widely from marginally higher in Inhambane province to more than double for the northern region.

In four instances (Tete, Manica, and Maputo provinces, plus Maputo City) including the poverty line error reduces the standard error of the Watts index. One instance of reduction was already noted for the poverty gap. As discussed earlier, this can occur if the welfare measure and the poverty line have a large positive covariance. Intuitively, one expects the covariance to be positive because a richer (poorer) sample would yield poverty lines reflecting more expensive (inexpensive) items in the reference consumption bundle. In these five instances, the covariance is sufficiently large to more than offset the additional variance contributed by the poverty line sampling error.

How important is the increase in standard errors of the estimated poverty measures when poverty line sampling error is included? One way of assessing this is to put it in the context of the existing literature. As indicated earlier, Howes and Lanjouw (1998) found that accounting for sample stratification and clustering increased the standard errors of estimated FGT poverty measures by 26–33 percent in Pakistan and 45–64 percent in Ghana. Adding the poverty lines as a

TABLE 5

Estimates of Watts Index and Standard Errors, Mozambique 2002–03

Region	Sample Size	Watts Index	Standard Error		
			Without Poverty Line Error	With Poverty Line Error[1]	Ratio of Standard Errors
National	8,700	0.2585	0.0100	0.0130	1.30
Rural	4,695	0.2455	0.0139	0.0172	1.24
Urban	4,005	0.2859	0.0158	0.0195	1.23
Regions					
Northern	2,310	0.1521	0.0112	0.0234	2.09
Central	3,100	0.2476	0.0178	0.0238	1.34
Southern	3,290	0.4101	0.0194	0.0212	1.09
Provinces					
Niassa	816	0.1702	0.0173	0.0249	1.43
Cabo Delgado	738	0.2382	0.0231	0.0277	1.20
Nampula	756	0.1088	0.0171	0.0326	1.91
Zambézia	733	0.1374	0.0239	0.0338	1.42
Tete	756	0.5625	0.0554	0.0512	0.92
Manica	816	0.3845	0.0552	0.0507	0.92
Sofala	795	0.1023	0.0133	0.0189	1.42
Inhambane	753	0.5573	0.0496	0.0515	1.04
Gaza	786	0.1871	0.0176	0.0238	1.35
Maputo Province	828	0.5639	0.0374	0.0352	0.94
Maputo City	923	0.3445	0.0254	0.0232	0.91

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

source of error increases the standard errors of the national-level poverty estimates in Mozambique by 18–30 percent. This suggests that accounting for poverty line sampling error may be nearly as important quantitatively as accounting for complex sample design, although results from other countries, and using alternative methods 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, as in the present example, in which the complex sample design was also included in estimating the poverty line error. Even though the impact of incorporating the variability of the poverty line might not be quite as dramatic as the effect of complex sample design found by Howes and Lanjouw (1998), there is no good reason to consistently overstate the precision of the poverty headcount by 15–20 percent, and the poverty gap or Watts indices by an even greater margin.

5. Conclusions

Poverty reduction is a fundamental objective of economic development, and reducing poverty is a major focus of governments, international financial institutions, and non-governmental and community-based organizations. The success of policies, programs, and development lending is increasingly judged in terms of poverty reduction. There has been substantial progress over the past three decades in the measurement of poverty, with the development of additively decomposable measures that reflect not only the number of poor persons, but also the depth and severity of poverty for sub-groups of the population. As most poverty estimates come from sample survey data, the statistical properties of poverty measures and appropriate inference procedures are important for evaluating the precision of poverty estimates and the statistical significance of poverty comparisons.

Studies of relative poverty have observed that there is sampling error associated with both the welfare metric and relative poverty lines calculated from the survey data. The recognition of poverty lines' sampling error has not extended to absolute poverty lines, even though they are also routinely estimated from sample survey data. 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, using CBN poverty lines and FGT and Watts poverty indices as an illustration. The approach is based on bootstrap methods that can be similarly applied to other methods of setting poverty lines (such as the Food Energy Intake approach) and to other 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 20–30 percent, with the standard errors increasing by up to 65 percent for some sub-groups, with similar results for the Watts index. Thus, to be considered statistically significant, changes in poverty levels need to be larger than previously believed.

Are there circumstances in which one can safely ignore the sampling error of poverty lines, and treat them as fixed constants, without sampling error that contributes to the error of the poverty measures? In our view, the only situation would be the case of poverty lines that are determined exogenously, without reference to survey data. As absolute poverty lines are supposed to 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, and accepting a measurable loss in precision by estimating poverty lines from available data, the latter has clear advantages.

Poverty analysts are increasingly employing stochastic dominance approaches to make robust poverty comparisons across a range of plausible poverty lines, rather than a single set of poverty lines pegged to a somewhat arbitrary level of utility (Atkinson, 1987; Davidson and Duclos, 2000). This is intuitively appealing, and avoids the need to make the (usually unconvincing) claim that the poverty line divides the population into discrete states of poor and non-poor. However, while stochastic dominance approaches usefully sidestep the issue of the point estimates of poverty lines, they do not necessarily avoid the issue of the variance of poverty lines. To make interpersonal welfare comparisons when the cost of acquiring basic needs varies over time or space, nominal consumption must be deflated by cost of living indices (Ravallion, 1998). Establishing a common welfare metric is typically accomplished by computing the welfare ratio, which is nominal consumption divided by the relevant poverty line, or y^*/z (Blackorby and Donaldson, 1987). If these poverty lines or cost of living indices are estimated from survey data, then the associated sampling error should be included in the confidence interval around the empirical CDFs. To the extent that the poverty line error increases the total error of the poverty estimates, the confidence regions around each CDF will be wider, and it will become more difficult to reject a null hypothesis of no dominance.[5] Adapting the methods presented in this paper to stochastic dominance approaches to poverty comparisons is an area for future research.

References

Arndt, C. and K. Simler, "Estimating Utility Consistent Poverty Lines," Food Consumption and Nutrition Division Discussion Paper No. 189, International Food Policy Research Institute, Washington, DC, 2005.

Atkinson, A. B., "On the Measurement of Poverty," *Econometrica*, 55(4), 749–64, 1987.

Bishop, J. A., K. V. Chow, and B. Zheng, "Statistical Inference and Decomposable Poverty Measures," *Bulletin of Economic Research*, 47(4), 329–40, 1995.

Blackorby, C. and D. Donaldson, "Welfare Ratios and Distributionally Sensitive Cost-Benefit Analysis," *Journal of Public Economics*, 34(3), 265–90, 1987.

Datt, G. and D. Jolliffe, "Poverty in Egypt: Modeling and Policy Simulations," *Economic Development and Cultural Change*, 53(2), 327–46, 2005.

Davidson, R. and J.-Y. Duclos, "Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality," *Econometrica*, 68(6), 1435–65, 2000.

Deaton, A., *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*, Baltimore, The Johns Hopkins University Press for the World Bank, 1997.

Deaton, A. and M. Grosh, "Consumption," in M. Grosh and P. Glewwe (eds), "Designing Household Survey Questionnaires for Developing Countries: Lessons from Ten Years of LSMS Experience," World Bank, Washington, DC, 2000.

[5]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.

Deaton, A. and S. Zaidi, "Guidelines for Constructing Consumption Aggregates for Welfare Analysis," Living Standards Measurement Study Working Paper No. 135, World Bank, Washington, DC, 2002.
Efron, B., "Bootstrap Methods: Another Look at the Jackknife," *Annals of Statistics*, 7(1), 1–26, 1979.
Efron, B. and R. Tibshirani, *An Introduction to the Bootstrap*, Chapman and Hall, New York, 1993.
Foster, J. E., J. Greer, and E. Thorbecke, "A Class of Decomposable Poverty Measures," *Econometrica*, 52(3), 761–6, 1984.
Gibson, J. and S. Rozelle, "Poverty and Access to Roads in Papua New Guinea," *Economic Development and Cultural Change*, 52(1), 159–86, 2003.
Greer, J. and E. Thorbecke, "A Methodology for Measuring Food Poverty Applied to Kenya," *Journal of Development Economics*, 24(1), 59–74, 1986.
Howes, S. and J. O. Lanjouw, "Does Sample Design Matter for Poverty Rate Comparisons?" *Review of Income and Wealth*, 44(1), 99–109, 1998.
INE (Instituto Nacional de Estatística), "Inquérito aos Agregados Familiares Sobre Orçamento Familiar—Relatório Final," Instituto Nacional de Estatística, Maputo, 2004.
Kakwani, N., "Statistical Inference in the Measurement of Poverty," *Review of Economics and Statistics*, 75(4), 632–9, 1993.
Kappel, R., J. Lay, and S. Steiner, "Uganda: No More Pro-poor Growth?" *Development Policy Review*, 23(1), 27–53, 2005.
MPF (Ministry of Planning and Finance, International Food Policy Research Institute, Purdue University), "Poverty and Well-Being in Mozambique: The Second National Assessment," Ministry of Planning and Finance, Maputo, 2004.
Mukherjee, S. and T. Benson, "The Determinants of Poverty in Malawi, 1998," *World Development*, 31(2), 339–58, 2003.
Preston, I., "Sampling Distributions of Relative Poverty Statistics," *Applied Statistics*, 44(1), 91–9, 1995.
Ravallion, M., "Poverty Rankings Using Noisy Data on Living Standards," *Economics Letters*, 45(4), 481–5, 1994a.
———, *Poverty Comparisons*, Harwood Academic Press, Chur, Switzerland, 1994b.
———, "Poverty Lines in Theory and Practice," Living Standards Measurement Study Working Paper No. 133, World Bank, Washington, DC, 1998.
———, "Pro-Poor Growth: A Primer," World Bank Policy Research Working Paper No. 3242, World Bank, Washington, DC, 2004.
Ravallion, M. and B. Bidani, "How Robust is a Poverty Profile?" *World Bank Economic Review*, 8(1), 75–102, 1994.
Ravallion, M. and M. Lokshin, "Testing Poverty Lines," *Review of Income and Wealth*, 52(3), 399–422, 2006.
Sen, A., "Poverty: An Ordinal Approach to Measurement," *Econometrica*, 44(2), 219–31, 1976.
Tarp, F., K. R. Simler, C. Matusse, R. Heltberg, and G. Dava, "The Robustness of Poverty Profiles Reconsidered," *Economic Development and Cultural Change*, 51(1), 77–108, 2002.
Watts, H. W., "An Economic Definition of Poverty," in D. P. Moynihan (ed.), *On Understanding Poverty*, Basic Books, New York, 1968.
WHO (World Health Organization), "Energy and Protein Requirements," Technical Report Series No. 724, Geneva, 1985.
World Bank, *Introduction to Poverty Analysis*, World Bank Institute, Washington, DC, http://siteresources.worldbank.org/PGLP/Resources/PovertyManual.pdf, 2005 (accessed August 1, 2006).
Zheng, B., "An Axiomatic Characterization of the Watts Index," *Economics Letters*, 42(1), 81–6, 1993.
———, "Statistical Inference for Poverty Measures with Relative Poverty Lines," Luxembourg Income Study Working Paper No. 167, Luxembourg, 1997.
———, "Poverty Orderings," *Journal of Economic Surveys*, 14(4), 427–66, 2000.
———, "Statistical Inference for Poverty Measures with Relative Poverty Lines," *Journal of Econometrics*, 101(2), 337–56, 2001.