Estimating Utility Consistent Poverty Lines

Channing Arndt and Kenneth R. Simler

African Development and Poverty Reduction: The Macro-Micro Linkage

Forum Paper 2004

13 - 15 October 2004
Lord Charles Hotel, Somerset West, South Africa

Development Policy Research Unit
Trade and Industrial Policy Strategies
Cornell University

Sponsors

Cornell
gtz
Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ) GmbH
USAID
IDRC CRDI
Friedrich Ebert Stiftung

Estimating Utility Consistent Poverty Lines

July 2004

Channing Arndt,
Associate Professor
Purdue University
Department of Agricultural Economics
403 W. State Street
West Lafayette, IN 47907-2056
carndt@purdue.edu
phone: 765 494 4304

Kenneth R. Simler,
Research Fellow
International Food Policy Research Institute
Food Consumption and Nutrition Division
2033 K Street, NW
Washington, DC 20006-1002

Abstract:

The “Cost of Basic Needs” (CBN) approach to drawing consumption based poverty lines is widely applied and lays credible claim to being the best practice for estimating poverty measures. Unfortunately, a growing mass of evidence indicates that poverty estimates obtained under the CBN approach are often demonstrably utility inconsistent. Here, we introduce an information theoretic approach for estimating utility consistent poverty lines. An example of the approach is provided for the case of Mozambique. The approach represents a powerful addition to the poverty analyst’s toolkit and enhances the attractiveness of the CBN approach for practical poverty measurement problems.

JEL Classification codes: C81, D12, O12.

1. Introduction

Poverty reduction has become one of the primary objectives of development assistance, and changes in poverty have become central yardsticks by which development assistance and accompanying government action are measured. The stakes are often high. In many poor countries, failure to reduce measured poverty levels over the medium term would very likely catalyze widespread calls for major policy reform as well as country program reviews by donor organizations including a reassessment of the overall level of assistance. In this environment, reasonably precise measures of poverty levels and changes in poverty levels through time are obviously desirable.

Not surprisingly, a voluminous literature exists on appropriate ways to measure poverty and a large variety of approaches have been proposed. This article focuses specifically on the "Cost of Basic Needs" (CBN) approach to drawing consumption based poverty lines described by Ravallion (1994, 1998). This approach is widely applied and lays credible claim to being the best practice for estimating income (or consumption) poverty measures such as the P_α class of measures developed by Foster, Greer, and Thorbecke (1984). Nevertheless, a growing mass of evidence indicates that poverty estimates obtained under the CBN approach are often demonstrably utility inconsistent (Ravallion and Lokshin 2003; Tarp et al. 2002).

In this article, we introduce an information theoretic approach for estimating utility consistent poverty lines, and provide an example of the approach for the case of Mozambique. The article is structured as follows. Section 2 briefly reviews the issue of utility consistency in the context of poverty lines and presents the revealed preference tests that have been applied recently to test for utility consistency. Section 3 provides a brief review of information theory, with a focus on entropy estimation. Section 4 presents the estimator in general form. Section 5 discusses application of the approach to the case of Mozambique. We conclude, in section 6, that the information theoretic approach presented here represents a powerful addition to the poverty analyst's toolkit and enhances the attractiveness of the CBN approach for practical poverty measurement problems.

2. Utility Consistency and the CBN Approach

Consumption Bundles, Relative Price Shifts, and Utility

In the CBN approach, poverty lines are obtained on the basis of the cost of a bundle of goods that correspond with an arbitrarily chosen minimum level of well being, usually based on nutritional requirements and basic nonfood needs. The exact number of distinct bundles to employ and their composition remain a vexing problem. Tarp et al. (2002) demonstrate that where large differences in relative prices across regions exist and preferences are not Leontief (or nearly so), use of a common bundle across all regions leads to poverty lines that are typically higher than they should be and the strong possibility of re-ranking in poverty comparisons (e.g., utility inconsistency).[1] The same argument holds for a constant bundle through time rather than across space when relative prices vary through time.

The crux of the issue can be illustrated in a highly simplified context using the case of relative price changes through time. Figure 1 depicts a consumer whose preferences are defined on only two goods, c_1 and c_2. Assume that we perfectly observed this utility maximizing consumer at a point in time, say 1996. Line M_0 represents the expenditure necessary to attain utility level U in 1996. If U represents the chosen minimum level of well being, then M_0 depicts the poverty line for 1996. Suppose that we wish to re-estimate the poverty line for 2002. If relative prices have changed between 1996 and 2002, then consumption patterns will change in accordance with preferences. The expenditure necessary to achieve the same utility level, or the appropriate new poverty line in 2002, is depicted by line M_1. Using the original basket of goods and 2002 prices to estimate the 2002 poverty line (M_2) overstates the expenditure necessary to maintain the same utility level.

The situation is exactly analogous if relative prices vary across space for a given point in time. Generally, when relative prices vary either through time or across space, the link between the cost of any single bundle and any given level of welfare deteriorates. The exact degree of deterioration depends upon preferences and the extent of shifts in relative prices. In terms of Figure 1, if the consumer in question actually spent between levels M_1 and M_2 in 2002, she would be better off in utility terms but measured as worse off using the constant bundle poverty line, i.e., one based on the quantities of the 1996 bundle (c_1^{96}, c_2^{96}).

The obvious solution to dealing with the shortcomings of a common bundle is to develop multiple bundles, each associated with a different time period (or region), in order to account for the shifts in consumption patterns brought about by the differences in relative prices through time (or across space). Multiple bundles have been estimated by, for example, Datt, Jolliffe and Sharma (2001); Gibson and Rozelle (2003); Ravallion and Lokshin (2003); and Tarp et al. (2002). Though conceptually straightforward, the practical difficulties associated with choosing multiple bundles that *yield the same (or nearly the same) level of well being* are considerable. To grapple with this issue, non-parametric tests for utility consistency have come to be applied to the problem of setting poverty lines (Ravallion and Lokshin 2003; Gibson and Rozelle 1999, 2003). These tests are considered in the next sub-section.

Revealed Preference Tests for Utility Consistency

The notion of revealed preferences originates in microeconomic theory. The idea is to apply the restrictions on rational consumer behavior postulated in microeconomic theory without imposing any specific form for preferences on individual behavior. Revealed preference restrictions rely on the assumption that consumers prefer consuming more rather than less (non-satiation) (Varian, 1992). To begin, assume identical consumers who prefer more to less and whose preferences are defined on I [$i \in I$] commodities with each consumer living in a distinct spatial domain, r, among the set of spatial domains R [$r \in R$]. We instruct these consumers to spend the minimum necessary in order to attain the same arbitrary level of utility. Prices potentially differ across spatial domains. Under these conditions, the following revealed preference conditions will hold:

$$\sum_i p_{ir} q_{ir'} \geq \sum_i p_{ir} q_{ir} \quad \forall\, r, r' \tag{1}$$

where r' represents an alias index for the set of spatial domains R [$r, r' \in R$] and the variables p and q represent prices and quantities, respectively.

The logic behind this set of restrictions is as follows. For a given spatial domain r, the representative consumer in r has the opportunity to choose any bundle that delivers the required utility level. As this consumer is rational and prefers more to less, she will choose a cost minimizing bundle, q_r.[2] By cost minimization, the cost of any other bundle that delivers the same level of utility, evaluated at prices p_r, must be at least as much as the cost minimizing bundle. Since, by assumption, the selected bundle in each spatial domain delivers the same utility level, condition (1) must hold for all possible pairs of spatial domains.

It is illuminating to consider the contrary case: suppose the inequality in (1) fails to hold for some region pair (r and r'), yet preferences are still identical across regions. When both bundles are evaluated at prices p_r, a rational consumer would only choose a more expensive bundle, q_r, over an alternative bundle, $q_{r'}$, if the chosen bundle yields higher utility. In this case, the maintained hypothesis of constant utility levels across bundles is rejected; in other words, the bundle, q_r, is revealed preferred to the alternative bundle, $q_{r'}$. Failure of condition (1) is also possible if preferences are not identical across regions. Also, note that comparison of bundles through time rather than across space as presented above is completely analogous.

The application of conditions such as those outlined in (1) to a group of bundles designed to represent the standard of living associated with a poverty line for a particular spatial domain and/or at a particular time is both straightforward and highly attractive. As Ravallion and Lokshin (2003) point out, the revealed preference conditions in (1) are necessary but not sufficient conditions for identical preferences. It is possible that the representative consumers mentioned above could have different preferences, yet all revealed preference conditions could still be satisfied. Rather than identical preferences, satisfaction of revealed preference conditions indicates the existence of a coherent preference set that rationalizes the observed behavior (Varian, 1992).

For the purposes of poverty analysis, this is arguably a highly attractive feature of revealed preference conditions. Meeting revealed preference conditions does not impose identical preferences across regions or through time. However, for the conditions to be satisfied, the observed behavior must be consistent with a set of rational preferences. Consistency with some arbitrary unknown preference set would appear to be a minimum condition for making the welfare comparisons fundamental to poverty analysis. Revealed preference conditions require this minimum and no more.

Empirical Relevance of the Problem of Choosing Utility Consistent Bundles

To summarize, utility consistency of the CBN approach is not guaranteed. In particular, the use of a single basic needs consumption bundle will generate inconsistent regional comparisons when relative prices vary across space, and inconsistent inter-temporal comparisons when relative prices change over time. The use of multiple bundles represents the obvious alternative; however, assuring that the different bundles give the same (or nearly the same) level of well-being presents a real concern. Fortunately, revealed preference conditions provide a convenient means of testing for utility consistency. This revealed preference approach to testing bundles for utility consistency has been applied recently by Ravallion and Lokshin (2003) for Russia; Gibson and Rozelle (1999, 2003) for Papua New Guinea; and MPF et al. (2004) for Mozambique.[3]

Unfortunately, the available information on the results of revealed preference tests indicate that failures of revealed preference conditions are widespread despite the best efforts to develop comparable bundles. For example, an analysis by Ravallion and Lokshin (2003) divides Russia into 23 spatial domains with a separate bundle developed for each domain. This results in 506 (23 x 22) revealed preference conditions.[4] These conditions fail nearly half the time. Only six of the 253 distinct pairs of bundles are mutually consistent, meaning that the revealed preference conditions are satisfied both when region A is compared with region B and when region B is compared with region A. Gibson and Rozelle (1999) experience similar difficulties in Papua New Guinea. The original bundles developed for Mozambique and presented in section 5 also frequently violate revealed preference conditions and generate relatively few mutually consistent pairs.

From this point, the existing literature becomes significantly less helpful. Ravallion and Lokshin (2003) point out the possibility that scalar corrections to the bundles could result in satisfaction of the revealed preference conditions, and they construct a test to determine whether feasibility with respect to revealed preference conditions might be obtained through scalar corrections. However, the possibility of a set of scalar corrections is rejected for their case (Russia). Gibson and Rozelle (1999) opt to aggregate regions until revealed preference conditions are satisfied (or nearly so). Both approaches fall well short of a general procedure for estimating CBN bundles that meet revealed preference conditions.

When confronted with failure of revealed preference conditions despite every effort at careful construction of the bundles, we posit that the poverty analyst has, in fact, reached a familiar juncture in experimental science. The analyst has used all available information to construct the bundles, yet the bundles fail to meet the minimum conditions necessary for making valid welfare comparisons. As all information has already been incorporated into the construction of the original bundles, no additional information exists, by definition, on which to base adjustment of the bundles in order to conform to revealed preference conditions.

This situation is a familiar one in diverse areas of enquiry. Consider the following examples. Despite careful collection of all available data following a year of economic activity, the national accountant is invariably confronted with a situation where the raw data fail to respect basic macroeconomic identities. Having carefully collected data on the motion of particles, the physicist notes that these data do not completely respect accepted rules of motion. A crime lab receives a photo of the license plate of a car fleeing the scene of a robbery. Unfortunately, the image is blurry and the license plate cannot be read. In all of these cases, the signal, despite the best attempts at observation, is noisy and fails to conform with what is required to be true. National accounts must respect basic macroeconomic identities. In order to make credible inferences, the physicist's data must conform to basic laws of motion. The license plate photo does not conform to what must be true about the license plate in question (e.g., the shape is rectangular with particular dimensions, the letters and numbers written on the plate are in block form, etc.).

Also, in all of these cases, relevant additional information would, without doubt, be useful. However, information is, at least for practical purposes, finite. At some point, the analyst is forced to obtain a coherent picture from the available information. Increasingly, entropy estimation is being employed to accomplish this objective. The following section briefly introduces entropy estimation approaches.

3. Entropy Estimation

Entropy approaches to estimation are motivated by "information theory" and the work of Shannon (1948), who defined a function to measure the uncertainty, or entropy, of a collection of events, and Jaynes (1957a; 1957b), who proposed maximizing that function subject to appropriate consistency relations, such as moment conditions. The maximum entropy (ME) principle and its sister formulation, minimum cross entropy (CE), are now used in a wide variety of fields to estimate and make inferences when information is incomplete, highly scattered, and/or inconsistent (Kapur and Kesavan 1992). The basic philosophy of entropy estimation is to use all available information and no more.

In economics, the ME principle has been successfully applied to a wide and rapidly growing range of estimation problems. Theil (1967) provides an early investigation of information theory in economics. Mittelhammer, Judge, and Miller (2000) provide a recent text book treatment which is focused more tightly on the ME principle and its relationships with more traditional estimation criteria such as maximum likelihood. Golan (2002) edits a special issue of the *Journal of Econometrics* focused on cutting-edge applications of the entropy principle.

In general, information in an estimation problem using the entropy principle comes in two forms: (1) information (theoretical or empirical) about the system that imposes constraints on the values that the various parameters to be estimated can take; and (2) prior knowledge of likely parameter values. In the first case, the information is applied by specifying constraint equations in the estimation procedure. In the second, the information is applied by specifying a discrete prior distribution and estimating by minimizing the entropy distance between the estimated and prior distributions—the minimum cross entropy (CE) approach. The prior distribution does not have to be symmetric and weights on each point in the prior distribution can vary. If the weights in the prior distribution are equal (e.g., the prior distribution is uniform), then the CE and ME approaches are equivalent.

Considerable work has been done to bring the general regression model into information theoretic frameworks (Golan, Judge, and Miller 1996). The framework also supports statistical inference (Imbens 1997). However, for most applications, the real power of the framework is that it makes efficient use of scarce information in estimating parameters.[5] As stated by physicists Buck and McAulay (1991, p. ix), "the intention is to give a way of extracting the most convincing conclusions implied by given data and any prior knowledge of the circumstance."

4. The Entropy Estimator

While the philosophy of entropy estimation and the properties of entropy estimators fill volumes, the actual empirical application of the entropy procedure is relatively simple. To obtain utility consistent bundles, we set up the following constrained minimization problem.

$$\underset{q_{ir},s_{ir}^{ent}}{Min} \quad \sum_r \sum_i s_{ir}^{ent} \ln\left(\frac{s_{ir}^{ent}}{s_{ir}^{orig}}\right) \tag{2}$$

Subject to:

$$\sum_i p_{ir} q_{ir'} \quad \geq \quad \sum_i p_{ir} q_{ir} \quad \forall\, r, r' \quad r \neq r' \tag{2a}$$

$$s_{ir}^{ent} \sum_{i'} p_{i'r} q_{i'r} \quad = \quad p_{ir} q_{ir} \quad \forall\, i, r \tag{2b}$$

$$s_{ir}^{ent} \geq 0\,,\; s_{ir}^{ent} \leq 1,\; q_{ir} \geq 0 \quad \forall\, i, r \tag{2c}$$

where notation is carried forward from condition (1), which has now become constraint (2a). New or revised notation is as follows:

s_{ir}^{ent} budget shares of the adjusted bundle,

s_{ir}^{orig} budget shares of the original bundle,

r, r' index of location in space or time, and

i' an alias index for i.

Equation (2) is the minimum cross entropy objective. The problem treats the expenditure shares from the original bundle as providing prior information on consumption patterns in the region. Prior information in the minimum cross entropy framework classically comes in the form of probabilities associated with a discrete distribution. In the problem at hand, the expenditure shares are viewed as the probability that an arbitrarily small quantity of currency will be devoted to the purchase of good i. The objective minimizes the entropy distance between the adjusted and original budget shares. The composition of the bundle—and consequently, the values of the budget shares—may need to be altered in order to meet revealed preference conditions. Overall, the optimization problem seeks to find, for each region r, vectors of quantities that satisfy revealed preference conditions and that preserve, to the greatest degree possible, the information content in the original budget shares.

The objective (2) is strictly convex. If all constraints were linear, the optimal solution would be a unique global minimum. However, as constraint equation (2b) is non-linear, the possibility of local optima must be admitted unless a proof of uniqueness, within the relevant range of the solution set, can be devised. Pending a proof, standard numerical methods associated with non-linear estimation (such as alternative starting values for variables and alternative algorithms for finding optimal solutions) should be employed in order to avoid local optima.

Two additional comments are worthwhile. First, for the minimization problem presented in (2), the minimum possible value for the objective function is zero (Golan, Judge, and Miller, 1996). This occurs when the adjusted shares equal the original shares. Consequently, the scalar correction procedure suggested by Ravallion and Lokshin (2003) will be the solution to the minimization problem in (2), if feasible, as it leaves the original budget shares intact. Second, other constraints can easily be added. For example, the analyst might wish to require that each bundle exactly meets estimated calorie requirements for each location.[6]

We turn now to an application of the approach presented above to the case of Mozambique. For the sake of brevity, many details of the procedure applied to Mozambique are not discussed here. Full details on the poverty analysis for Mozambique can be found in MPF et al. (2004).

5. Estimating Poverty Lines for Mozambique

Background

In 1996-97, the Mozambican National Institute of Statistics conducted the first nationally representative household consumption survey (HCS) in Mozambique. Analysis of the survey indicated a poverty headcount of about 69.4 percent at the national level, with poverty more prevalent in rural than urban areas (MPF/UEM/IFPRI, 1998). In 2002-03, a second nationally representative HCS was undertaken. Two primary objectives guided the design, implementation, and analysis of the 2002-03 HCS. The first objective was to provide the best possible picture of poverty and well-being in the year 2002-03. The second objective was to provide a sound basis for comparison with the 1996-97 survey. In order to satisfy the second objective, the approach and methods employed for analysis of the two surveys were quite similar.

During the interval between surveys, published information on the evolution of macro-aggregates allowed for the possibility of rapid reductions in poverty. Real GDP per capita grew by a cumulative 62% between 1996 and 2002. Real consumption per capita registered a slightly lower but still impressive cumulative growth of 50% over the same period. Furthermore, national accounts data indicated that all major sectors, including agriculture, contributed to this growth. In addition, a Core Welfare Indicators Questionnaire (CWIQ) survey undertaken in 2000–01 pointed to impressive gains in poverty correlates such as bicycle ownership and radio ownership (INE 2001). Nevertheless, substantial debate persisted regarding the participation of the poor in the observed economic growth. The 2002-03 HCS was expected to contribute substantially to, if not settle, this debate.

Basic Approach

The 2002-03 HCS contained detailed information on expenditure for a random sample of 8700 households. The sample represented the nation, rural and urban zones, and each province plus Maputo City. An important feature of the survey was an explicit attempt to be representative in time as well as space. Data collection for the survey began in July 2002 and finished in June 2003 (covering a complete agricultural season). Available indicators point to a high level of information quality.

To measure poverty, analyses of both the 1996-97 and 2002-03 surveys employed the CBN approach (Ravallion 1994, 1998). As prices (both relative prices and price levels) and consumption patterns vary drastically across space in Mozambique, 13 spatial domains were identified in 1996-97. These same spatial domains were employed in 2002-03. In each domain, a basket of food products that satisfied basic calorie needs was identified using information on the age and sex composition of the household and the recorded consumption patterns of poorer households. The cost of this basket, valued at average prices from within each domain, represented the food poverty line for each domain. A nonfood poverty line was obtained for each domain by calculating the share of food expenditures for households whose total food and nonfood consumption per capita was near the food poverty line. The total poverty line was obtained as the sum of the food and the nonfood poverty lines.

Fixed Bundles and Substitution Effects

In 2002-03, two methods were employed to obtain the food basket. The first method simply involved using the same baskets from 1996-97; this is referred to here as the fixed bundle approach. Since this method is easy, it is ideal if it can be expected to function reasonably well. Unfortunately, the data indicate that relative prices for food products shifted substantially between 1996-97 and 2002-03 and that consumers adjusted their food consumption patterns in response to these shifts in relative prices. As mentioned in section 2, the fixed bundle approach could be expected to overstate poverty rates in this environment.

The potential magnitude of the overstatement is worth exploring. The fixed bundle approach indicated a national poverty headcount ratio of 63.2%. One way to consider the potential magnitude of the implications of the assumption of Leontief preferences implicit in the fixed bundles approach (when relative prices have changed) is to assume an alternative preference structure that permits substitutability. For this purpose, Cobb-Douglas preferences are an attractive choice, as the functional relationship is well known and easy to apply. The Cobb-Douglas functional form probably overstates substitutability between some commodity pairs and understates it for others. Nevertheless, the assumption that consumers have Cobb-Douglas preferences likely provides a reasonably robust indicator of the potential importance of substitution effects.

A calibrated Cobb-Douglas utility function can be obtained easily from the 1996-97 results for each of the 13 spatial domains. With these functions in hand, a new food poverty line for 2002–03 was obtained by minimizing the cost of achieving the same utility level as observed in 1996-97 under prices prevailing in 2002-03. The nonfood poverty line was then obtained using the methodology described above.

Under the assumption of Cobb-Douglas preferences, the measured poverty rate falls to 52.1% at the national level. This represents a substantial decline from the 63.2% rate obtained using the fixed bundle approach. As preferences are almost surely not Cobb-Douglas, the result is not a viable poverty estimate. Nevertheless, the result does confirm that accounting for price variation and consumer response, as in the flexible bundle approach, could result in measured poverty rates substantially below the levels implied by the fixed bundles approach.

Flexible Bundles and Revealed Preference Conditions

To account for quantity response to shifts in relative prices, new bundles within each domain were derived using the iterative approach described in Ravallion (1994), which was also the procedure applied to the 1996-97 HCS. The food bundles developed for each spatial domain were then tested to see if they satisfied regional and temporal revealed preference conditions. The results of the regional revealed preference tests are presented in Table 1. These regional revealed preference tests pass only slightly better than half the time, and with only six mutually consistent pairs out of 78 possible. Temporal conditions also failed in some instances although not as frequently. Note in particular that failures are consistently observed across row 3 (Nampula rural) indicating a low quality bundle in that domain. Consistent failures are also observed in columns 7, 8, 11, 12, and 13 (Manica, Tete, and Maputo) indicating relatively high quality bundles in those domains. Overall, poverty lines calculated on the basis of these original flexible bundles would be utility inconsistent.

Some practical issues associated with conducting the revealed preference tests for the particular case of Mozambique merit mentioning. First, the revealed preference tests are conducted using only food bundles. Price and quantity information is employed only for the food bundle that generates the food poverty line. As mentioned earlier, the nonfood poverty line is calculated on the basis of nonfood consumption shares for households living near the food poverty line. Difficulties associated with identifying quantities and prices of nonfood goods drive this choice. As a result, the revealed preference tests are conducted on a sub-set of the expenditures that comprise the poverty line, albeit an important subset since food represents around three quarters of total expenditure for those living at or somewhat below the poverty line.[7]

Second, consumption patterns do vary widely across the 13 spatial domains used in Mozambique. To conduct the regional revealed preference tests, one must determine the cost of the bundle chosen in region r' using prices in region r. In some instances, some commodities from the bundle of region r' are very rarely or not at all produced/traded in region r. Hence, information on prices for region r is lacking. In these instances, we assume that the price prevailing in region r is the maximum observed price across all spatial domains.

Estimating Revealed Preference Consistent Food Bundles

To obtain consistent food bundles across space and through time, we solve the following constrained minimization problem.

$$\underset{q_{ir}^{02},\, s_{ir}^{ent}}{Min} \qquad \sum_{r}\sum_{i} s_{ir}^{ent} \ln\left(\frac{s_{ir}^{ent}}{s_{ir}^{orig}}\right) \tag{3}$$

Subject to:

$$\sum_{i} p_{ir}^{02} q_{ir'}^{02} \quad \geq \quad \sum_{i} p_{ir}^{02} q_{ir}^{02} \quad \forall\, r, r' \quad r \neq r' \tag{3a}$$

$$\sum_{i} p_{ir}^{02} q_{ir}^{96} \quad \geq \quad \sum_{i} p_{ir}^{02} q_{ir}^{02} \quad \forall\, r \tag{3b}$$

$$\sum_{i} p_{ir}^{96} q_{ir}^{02} \quad \geq \quad \sum_{i} p_{ir}^{96} q_{ir}^{96} \quad \forall\, r \tag{3c}$$

$$s_{ir}^{ent} \sum_{i'} p_{i'r} q_{i'r} \quad = \quad p_{ir} q_{ir} \quad \forall\, i, r \tag{3d}$$

$$\sum_{i} calpg_{i} q_{ir} \quad = \quad cal \quad \forall\, r \tag{3e}$$

$$s_{ir}^{ent} \geq 0\,,\; s_{ir}^{ent} \leq 1\,,\; q_{ir} \geq 0 \quad \forall\, i, r \tag{3f}$$

where basic notation is carried forward from equation set (2) with some revisions. [8] Notation is as follows:

r, r'	indices of spatial domains,
i,i'	indices of commodities,
s_{ir}^{ent}	budget shares of the adjusted bundle from 2002,
s_{ir}^{orig}	budget shares of the original bundle from 2002,
$calpg_{i}$	calories per quantity unit for each commodity i, and
cal	calorie requirement for all bundles.

In addition, price and quantity variables now have temporal superscripts.

Practical choices also impinge upon the actual estimation. First, note that calorie requirements are imposed with equality (constraint (3e)).[9] More importantly, the spatial domains in Maputo City and Maputo Province (domains 11, 12, and 13) are excluded from the revealed preference conditions that compare bundles across space (constraint 3a). This choice reflects the large differences in mode of living that exist in Maputo. In particular, cash income and cash requirements are clearly more important in Maputo than in other regions. Also, opportunities to earn cash are more readily available. As a result, household members in Maputo City, including lower income households, more frequently work outside the home and more frequently purchase services, such as prepared food or milled grain, than in other regions where these services are overwhelmingly produced within the home. The method employed to measure consumption counts all expenditures made outside the home but ignores services produced and consumed at home, such as food preparation.[10]

This characteristic of the data helps to explain the nearly complete failure of revealed preference conditions in the Maputo columns of Table 1. The failures might not be nearly so complete if the value of home produced services implicit in the bundles from other regions were included at Maputo shadow values on labor. As the data do not permit estimation of this value, we choose to simply exclude Maputo from the spatial revealed preference conditions. This choice is also consistent with the analysis from 1996-97 where revealed preference conditions for Maputo also failed drastically.

Table 2 shows spatial revealed preference conditions for the 10 spatial domains included in the adjustment procedure. Diagonal elements of the table provide the adjusted food poverty line estimate for each spatial domain. The adjusted bundles for all regions (including Maputo) also respect the temporal revealed preference conditions specified in equations (3b) and (3c).

Results

Table 3 compares poverty headcount results for the original and the adjusted bundles. Poverty headcount ratios for 1996-97 and for the fixed bundle approach in 2002-03 are presented as well. Comparing the two sets of flexible bundles (original and adjusted), one observes that estimated poverty rates at the national level climb by about six percentage points when utility consistency is imposed. The increase is driven strongly by substantial adjustments to the bundle in region 3 (Nampula rural) – a very populous and primarily rural region. For those with experience in Mozambique, the adjusted bundle results are far more credible. The original estimate for Nampula appears very low and fails to concord with other available indicators such as asset ownership.[11] The increase in poverty rates in Nampula under the adjusted bundle is instrumental in raising the rural poverty rate above the urban poverty rate, which is also a more credible result. In addition, the adjustment procedure significantly reduces estimated poverty rates in Manica, which is consistent with other available welfare correlates for Manica.[12]

Despite generating substantial changes in estimated poverty rates for some provinces, the qualitative story in terms of directions of movement compared with the levels estimated for 1996-97 is fairly robust. Under both the original flexible bundle and the adjusted flexible bundle approaches, poverty rates declined substantially nationwide. Rates declined more rapidly in rural than in urban zones. Overall, the correlation in measured changes in poverty headcounts relative to the estimates from 1996-97 for the two estimators by province is about .85 indicating that both estimators tend rather strongly to point in the same direction on a provincial basis.[13] The utility consistent estimates derived from the adjusted bundles present, in our view, the preferred estimates.

6. Summary and Conclusions

With the focus of international development resources increasingly turned toward poverty reduction, the demand for reliable empirical estimates of poverty levels has grown dramatically. Governments and donors are particularly interested in making valid poverty comparisons across space and through time. In fluid developing country environments, generating these comparisons is challenging.

We present an approach for estimating utility consistent poverty lines building on the approaches developed by Ravallion (1994, 1998) and Ravallion and Lokshin (2003). The basic approach, motivated by information theory, has been applied to numerous conceptually similar problems in empirical science across a variety of disciplines. Even though the philosophical roots of information theory and the links between information theory and other estimation criteria fill volumes, the actual practical application of the approach is quite straightforward.

For the case of Mozambique, the approach permitted estimation of poverty measures based on utility consistent poverty lines. These poverty measures are preferred over fixed bundle estimates because of substantial changes in relative prices and consumption patterns over the period 1996-97 and 2002-03. Relative to the original flexible bundle, the adjusted bundles provide more credible results, with the larger adjustments conforming to available information on correlates of poverty such as asset ownership. Utility consistency substantially enhances the credibility of the results obtained and should facilitate welfare comparisons in the future.

Overall, we conclude that the approach for estimating utility consistent bundles represents a powerful addition to the poverty analyst's toolkit and enhances the attractiveness of the CBN approach for practical poverty measurement problems.

7. Endnotes

---

[1] It bears emphasizing that in addition to serving as dividing lines between poor and nonpoor households, poverty lines are typically used as deflators that map nominal consumption to real, thus establishing the comparability of the welfare measure across population sub-groups.

[2] The i subscripts have been dropped. The vector, q_r, is (I x 1). An analogous notational shift is made for prices.

[3] Note that these are all settings where relative prices can be expected to vary substantially across space because of one or more factors, including large distances, ecological diversity, poor infrastructure, and weak market integration .

[4] The test when r and r' refer to the same region is trivial and is omitted.

[5] Golan, Judge, and Miller (1996) show that the ME/CE approach is an "efficient" information processing rule, as described by Zellner (1988).

[6] Note that, if a calorie constraint is added, the scalar correction approach of Ravallion and Lokshin (2003) will no longer be feasible.

[7] Revealed preference conditions then implicitly impose separability between food and non-food expenditures.

[8] If the original share, s_{ir}^{orig}, was equal to zero, then the adjusted share was set equal to zero. In the limit, $x \ln(x/x) = 0$ as x approaches zero. Numerically, these terms were simply dropped from the objective.

[9] The bundles are adjusted based on constant calorie requirements for each spatial domain. Once the adjusted bundles are obtained, they are rescaled to meet calorie requirements for each spatial domain.

[10] This bias in consumption surveys is essentially the same as the bias in the national accounts, which exists in all countries. Home produced/consumed goods, such as agricultural products, are valued while home produced services, such as cooking and cleaning, are not.

[11] For example, in 2002–03 households in Nampula province ranked below the national average in bicycle ownership, slightly above the national average in radio ownership, and below the national average in number of meals consumed the day prior to the beginning of interviews. These indicators correlate with measured poverty rates (MPF et al. 2004).

[12] Households in Manica in 2002-03 ranked well above the national average in bicycle ownership, radio ownership, and meals per day.

[13] The fixed bundle analysis also tells a similar qualitative story in terms of directions of movements relative to 1996-97 at the provincial level. The correlation in the changes in provincial headcounts between the adjusted bundle measure and the fixed bundle measure is 0.87.

8. References

Buck, B. and V. A. MacAulay eds. 1991. *Maximum Entropy in Action.* Clarendon Press, Oxford.

Datt, G., D. Jolliffe and M. Sharma. 2001. A profile of poverty in Egypt. African Development Review 13(2): 202–237.

Foster, J., J. Greer and E. Thorbecke. 1984. A class of decomposable poverty measures. Econometrica. 52: 761-765.

Gibson, J. and S. Rozelle. 1999. Results of the household survey component for the 1996 poverty assessment for Papua New Guinea. Discussion Paper, Poverty and Human Resources Division, The World Bank, Washington, D.C.

Gibson, J., and S. Rozelle. 2003. Poverty and access to roads in Papua New Guinea. Economic Development and Cultural Change. 52(1): 159–186.

Golan, A. 2002. Information and Entropy Econometrics—An Editor's View. Journal of Econometrics 107(1–2): 1-15.

Golan, A., G. Judge and D. Miller. 1996. "Maximum Entropy Econometrics: Robust Estimation with Limited Data". (Series in The Financial Economics and Quantitative Analysis). John Wiley & Sons Ltd.

Imbens, G. 1997. One-step estimators for over-identified generalized method of moment models. Review of Economic Studies 64: 359-383.

INE (Instituto Nacional de Estatística). 2001. Questionário de indicadores básicos de bem-estar: relatório final. Maputo: Instituto Nacional de Estatística.

Jaynes, E.T. 1957a. Information theory and statistical mechanics II. Physics Review 108: 171-190.

Jaynes, E.T. 1957b. Information theory and statistical mechanics I. Physics Review 106: 620-630.

Kapur, J.N. and H.K. Kesavan, 1992. *Entropy optimization principals with applications*. Academic Press Boston, p. 1-76.

Mittelhammer, R.C., G.G. Judge and D.J. Miller. 2000. *Econometric Foundations*. Cambridge University Press, Cambridge.

MPF/UEM/IFPRI (Mozambique, Ministry of Planning and Finance/Eduardo Mondlane University/ International Food Policy Research Institute). 1998. Understanding poverty and well-being in Mozambique: The first national assessment (1996–97). Maputo.

MPF/IFPRI/Purdue (Mozambique Ministry of Planning and Finance / International Food Policy Research Institute / Purdue University). 2004. Poverty and well-being in Mozambique: The second national assessment. Maputo.

Ravallion, M. 1994. *Poverty comparisons*. Chur, Switzerland: Harwood Academic Publishers.

———. 1998. Poverty lines in theory and practice. Living Standards Measurement Study Working Paper No. 133. Washington, D.C.: World Bank.

Ravallion, M. and M. Lokshin. 2003. On the utility consistency of poverty lines. World Bank Policy Research Working Paper No. 3157. Washington, D.C.: World Bank.

Shannon, C.E. 1948. A mathematical theory of communication. Bell System Technical Journal 27: 379-423.

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 (1): 77–108.

Theil, H. 1967. *Economics and Information Theory*. North-Holland Pub. Co., Amsterdam.

Varian, R. H. 1992. Microeconomic Analysis. Third Edition. University of Michigan. W.W. Norton & Company. New York. London.

Zellner, A., 1988. 'Optimal information processing and Bayes theorem'. *American Statistician,* 42: 278-284.

9. Figures and Tables

Figure 1: Impact of changes in relative prices.

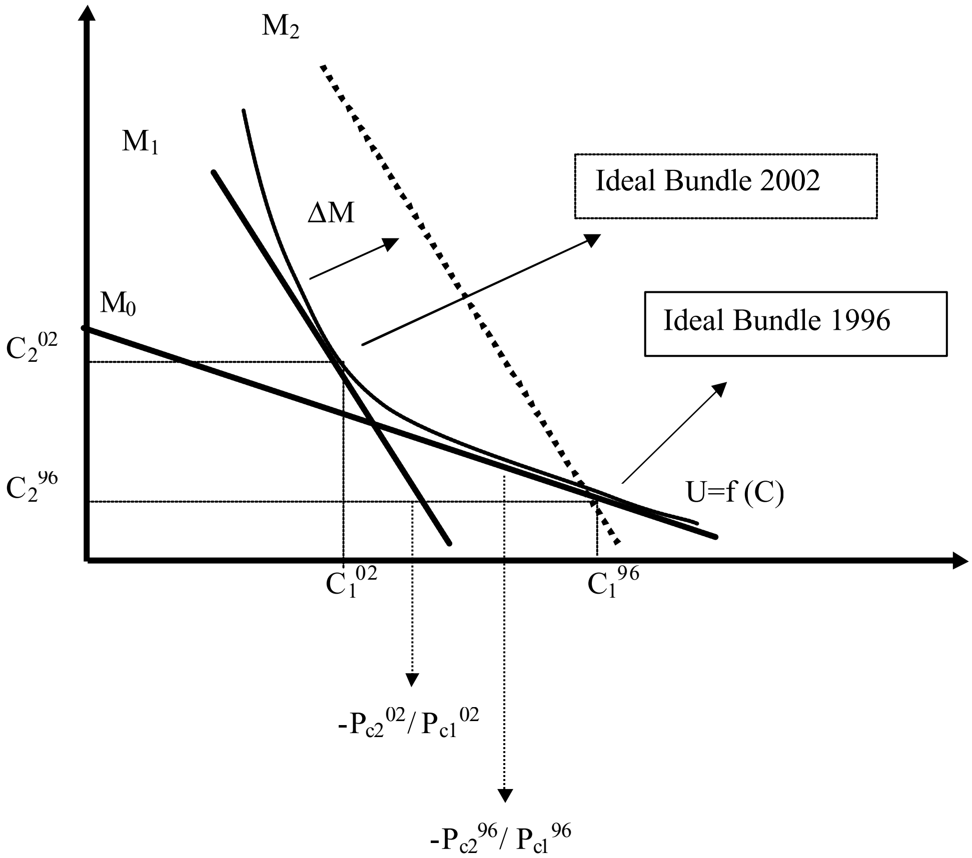

Table 1: Results of Regional Revealed Preferences Tests for Original Flexible Bundles from 2002-03.

	Region	1	2	3	4	5	6	7	8	9	10	11	12	13
1	Niassa-Cabo Delgado, rural	**4756**	**6397**	3991	*4472*	4007	**5621**	**5508**	**6330**	5580	**6250**	**6536**	**8436**	**9984**
2	Niassa-Cabo Delgado, urban	5903	**7717**	4501	5490	4922	6601	**6420**	**7599**	7090	7972	**8791**	**10409**	**10300**
3	Nampula, rural	**3500**	**4470**	**2752**	**3660**	**2907**	**4713**	**3041**	**2492**	**4703**	**3539**	**3499**	**4820**	**7099**
4	Nampula, urban	*4879*	**5853**	3542	**3749**	**3058**	**5232**	**4471**	**5956**	*5816*	**5429**	**5216**	**7833**	**7397**
5	Sofala-Zambezia, rural	**4589**	**6167**	3663	4399	**3548**	**5459**	**4768**	**5090**	**5041**	**5080**	**5691**	**7033**	**9124**
6	Sofala-Zambezia, urban	5730	**7402**	4216	5358	4446	**5902**	**6180**	**7006**	6331	*6811*	**8102**	**8177**	**9389**
7	Manica-Tete, rural	6770	8770	4741	7210	5090	7741	**6937**	**9584**	9608	10260	**12430**	*15311*	**11361**
8	Manica-Tete, urban	7737	9813	5646	7079	6058	8910	7863	**9657**	9087	10128	**12221**	**13032**	**11770**
9	Gaza-Inhambane, rural	**4454**	**5813**	3389	*4014*	3577	**5601**	**4587**	**4950**	**5438**	**5932**	**10243**	**8752**	**8969**
10	Gaza-Inhambane, urban	5090	**6728**	3943	5048	4303	*6753*	**5580**	**6419**	6458	**6613**	**9812**	**9279**	**9451**
11	Maputo Province, rural	7102	10317	5677	7657	6376	9478	7291	**9532**	9663	10422	**12584**	*13772*	*13816*
12	Maputo Province, urban	8158	10971	5860	8153	7482	11599	*9158*	11329	10938	11580	*13881*	**13741**	13700
13	Maputo Cidade	7866	10626	5653	7837	7146	11458	8921	11179	10766	11433	*13501*	**13270**	**13211**

Notes: Rows refer to quantities and columns refer to prices. For example, the value 4879 in cell (4,1) refers to quantities from Nampula-urban evaluated at Niassa-Cabo Delgado, rural prices. The diagonal elements (bolded and shaded) indicate the original flexible basket food poverty lines. Off-diagonal elements represent different price-quantity combinations. Values in a column less than the value in the diagonal indicate failures of revealed preference conditions. These values are bolded. Nine pairs, of 78 possible, are mutually consistent. These 18 cells are in italics. All figures are in Meticais per person per day.

Table 2: Post Adjustment Spatial Revealed Preference Tests for 2002-03.

		1	2	3	4	5	6	7	8	9	10
1	Niassa-Cabo Delgado, rural	**5434**	***7541***	***4471***	5146	4424	6679	6137	7573	***6614***	7808
2	Niassa-Cabo Delgado, urban	5642	**7541**	***4471***	5290	4746	***6591***	6190	7355	6627	7707
3	Nampula, rural	5988	8912	**4471**	5762	4502	7804	***5628***	***7145***	7856	8297
4	Nampula, urban	7014	8900	5067	**4853**	***4155***	7312	6603	9937	7936	8359
5	Sofala-Zambezia, rural	5816	8340	4600	5486	**4155**	7162	5772	***7145***	***6614***	***7264***
6	Sofala-Zambezia, urban	6060	8209	***4471***	5836	4673	**6591**	6411	7564	6790	7666
7	Manica-Tete, rural	6087	10244	***4471***	8629	4182	8286	**5628**	9806	11301	10810
8	Manica-Tete, urban	6118	**7541**	4648	5786	4935	7003	6039	**7145**	7435	8010
9	Gaza-Inhambane, rural	5823	7553	***4471***	5380	4920	7954	5937	***7145***	**6614**	8936
10	Gaza-Inhambane, urban	5564	**7541**	***4471***	5605	4713	7468	5990	***7145***	6839	**7264**

Notes: Rows refer to quantities and columns refer to prices. The diagonal elements (bolded and shaded) indicate the adjusted flexible basket food poverty lines. Bolded values indicate binding spatial revealed preference conditions. All figures are in Meticais per person per day.

Table 3: Poverty headcount results.

	1996-97	2002-03			Difference
		Fixed Bundles	Flexible Bundles		Adj.-Orig.
			Original	Adjusted	
National	69.4	63.2	48.0	54.1	6.1
Urban	62.0	61.3	52.4	51.5	-0.9
Rural	71.3	64.1	45.9	55.3	9.4
Niassa	70.6	61.2	45.6	52.1	6.5
Cabo Delgado	57.4	72.3	57.1	63.2	6.1
Nampula	68.9	68.1	30.5	52.6	22.1
Zambezia	68.1	58.6	35.1	44.6	9.4
Tete	82.3	71.6	70.8	59.8	-11.0
Manica	62.6	60.2	58.5	43.6	-15.0
Sofala	87.9	48.4	30.9	36.1	5.2
Inhambane	82.6	80.1	75.1	80.7	5.6
Gaza	64.6	58.6	47.1	60.1	13.1
Maputo Prov	65.6	66.9	75.9	69.3	-6.6
Maputo City	47.8	45.5	58.0	53.6	-4.4

Note: The headcount results are presented by province rather than by spatial domain.