Review of Income and Wealth
Series 47, Number 1, March 2001

HOW TO COMPARE APPLES AND ORANGES: POVERTY MEASUREMENT BASED ON DIFFERENT DEFINITIONS OF CONSUMPTION

BY JEAN OLSON LANJOUW

Yale University and the Brookings Institution

AND

PETER LANJOUW

the World Bank and Vrije Universiteit, Amsterdam

Poverty rates calculated on the basis of household consumption expenditures are routinely compared across countries and time. The surveys which underlie these comparisons typically differ in the types of food and non-food expenditures included, often in ways which are easily overlooked by analysts. With several examples we demonstrate that these commonly occurring variations in expenditure definitions can give rise to marked differences in poverty rates where there are no real differences in well-being. We show that one approach to calculating poverty lines, used with the headcount measure of poverty, can allow comparisons based on data with different definitions of consumption. In addition to allowing comparative poverty analysis using existing survey data, the results suggest that poverty monitoring could be done effectively at lower cost by alternating detailed expenditure surveys with far more abbreviated surveys.

1. INTRODUCTION

Goals as diverse as the econometric analysis of long-run growth or the design of targeted welfare policies call for comparable measures of poverty for sets of countries, population subgroups, locations or time-periods. When comparing poverty across data sources, analysts are sometimes not careful enough, and at other times, too careful. On the one hand, poverty rates based on different indicators are often treated as comparable when they are not. On the other, efforts to make poverty measures comparable by fielding similarly detailed, and thus expensive, surveys across regions and time-periods may be unnecessary.

A number of steps are routinely taken in order to make sets of poverty measures more comparable. For example, purchasing power parity adjustments have been introduced to correct for varying costs of living across countries, within countries and over time.[1] Those constructing more recent income and expenditure

Note: This paper was completed while both authors were visiting the Vrije Universiteit, Amsterdam and we thank the university for its support. We also thank seminar participants at the Tinbergen Institute, Amsterdam, and University of Namur for useful discussions, and Gaurav Datt, Angus Deaton, Jean Drèze, Francisco Ferreira, Jesko Hentschel, Stephen Howes and Martin Ravallion, for comments and suggestions. The views expressed in this paper are those of the authors and should not be interpreted to reflect those of the World Bank or any of its affiliates. All remaining errors are our own.

[1]See, for example, Summers and Heston, 1988 and 1991; and Bidani and Ravallion, 1994. Price indices sometimes take account of the fact that the inflation adjustments necessary to compare poverty may be quite different from those to compare average welfare levels. See, for example, Rocha, 1993; the Government of India, 1993; and Deaton and Tarozzi, 2000.

databases for country-level analyses have recognized the importance of national coverage, and consistency in the reporting unit, the treatment of non-wage income and taxation (see, for example, Deininger and Squire, 1996). There is also recognition that measures based on income and consumption indicators are not interchangeable. However, given one of those indicators, its underlying *definition* is often simply assumed to be the same across data sets. We show here that restricting attention to welfare measures based on consumption expenditures is not, by itself, sufficient to ensure comparability.[2] Surveys vary widely and in subtle ways, from differences in the formulation of questions or choices of recall periods to differences in the method used to impute categories of expenditure. Ignoring this inconvenient truth is not innocuous. We show here that what would appear to be small differences in definitions can be the source of surprisingly large differences in measured poverty. In so far that the unwitting analyst typically assumes that a common definition of consumption is being applied, these data-driven differences in measured poverty will be attributed to real factors.

In some cases there is a simple remedy to the problem of different definitions. We demonstrate that, if certain assumptions hold, there exists a method of calculating a poverty measure, the headcount rate, which is entirely robust to variation in the comprehensiveness of the consumption definition. This result requires: (1) that different components of expenditure follow a regularity akin to Engel's Law; (2) that expenditure patterns are stable across groups or time periods being compared; and (3) that there is no mismeasurement in the data (or that the degree of noise is the same both across indicators and across the groups being compared). Although these assumptions are clearly strong, we apply the result in a variety of empirical settings and find that for many purposes it works remarkably well. It is also simple to apply, even in the absence of specific information about the components of expenditure included in various data sets.

This finding raises the following question: If comparisons can be made regardless of the comprehensiveness of the data, do we need to collect detailed data? Full-scale surveys are very expensive to field.[3] As a result, such survey data sets are either relatively small or costly. Our theoretical and empirical results suggest that in certain situations, depending upon the specific objective, one could purposefully field surveys that are non-identical by design and still make valid comparisons. For example, when tracking poverty over time, one might consider being less "careful" and interspersing periodic full-scale household surveys with more abbreviated, low-cost surveys.

Whether one can use the approach described here to compare poverty rates across groups using different expenditure indicators depends on whether the

[2]Our focus in this paper, and our result in terms of achieving comparability across non-identical definitions of welfare, pertains to consumption rather than income as the indicator of well-being. The attraction of consumption over income as a basis for evaluating poverty has been widely discussed (see, for example, Atkinson, 1989 and 1991; Deaton, 1997; and Ravallion, 1994a). In the United States, shifting from an income-based analysis of poverty to one based on consumption has been advocated recently by the Panel on Poverty and Family Assistance, convened by the Joint Economic Committee of Congress (National Research Council, 1995), and also forcefully urged in Jorgensen (1998).

[3]The cost of a World Bank Living Standards Measurement Survey (LSMS) covering two to five thousand households ranges between $300,000 to $1.5 million, depending on population, local geographic conditions, availability of expertise, equipment, and so on.

assumptions listed above fail badly, given the purpose at hand. In our empirical examples it does not appear that variation in the degree of mismeasurement at different levels of consumption aggregation is a crucial issue. It is unlikely, however, that the assumption of uniform patterns of expenditure will hold across countries at different levels of development. For example, the relative consumption of energy is far higher in more developed economies. Even within groups of high- or low-income countries, differences in relative prices, cultural preferences, or needs will lead to different expenditure patterns. Thus our approach is not going to be very helpful at resolving problems of comparability at the level of cross-country analysis. For policy purposes, however, the focus will more often be on changes in poverty over time or across subgroups within a given country. Here stable expenditure patterns, at least in the short run, are more plausible. Usefully, when considering differences in poverty across population subgroups (as in a poverty profile) the reasonableness of the assumptions can be checked from an initial full-scale survey.

Using household survey data from several developing countries, the next section demonstrates that one can make misleading comparisons of poverty, even when the consumption aggregates used seem to be very similar. Section 3 shows that, of the various techniques frequently employed to develop poverty lines, there is one which will yield headcount poverty rates which are invariant to the indicator employed (under the assumptions listed above). We then consider mismeasurement of household expenditures. Some components of consumption are better measured than others, which may create a rationale for wanting a more (or less) comprehensive measure of expenditures. We give some guidance for making this decision. In Section 4 we demonstrate that the invariance result holds in a wide variety of empirical settings, and that its application is very successful in overcoming the problems presented in Section 2. In the last part of Section 4 we examine a poverty profile. A profile gives poverty rates for different subgroups of the population, indicating the extent to which groups with certain characteristics are more likely to be poor (for example, the unemployed versus the employed). Such profiles are particularly useful for understanding welfare conditions and the effects of policies. Because headcount measures are the basis of most profiles in practice, the robustness result may also be useful in their construction. The final section concludes the paper.

2. Comparing Poverty Under Varying Consumption Definitions

The following examples demonstrate the magnitude of the problem.

Example 1: El Salvador—Variation in Food Expenditure Definitions

In El Salvador, during the period between July and September 1994, a household survey (the *Encuesta de Hogares de Propositos Multiples*) was fielded with a coverage of 4,220 households. The survey was the subject of an experiment: two non-overlapping samples were drawn from the same sampling frame and were administered different consumption modules. A short module asking about the consumption of eighteen food items and six non-food items was completed for

3,182 households. A long module inquiring into the consumption of seventy-two food items and twenty-five non-food items was completed for 1,038 households. The two modules were both after a full definition of consumption—and differed only in the extent to which they aggregated consumption items.

The two sub-samples were drawn from the same frame and were explicitly intended to be identical in all respects. Scott and Jolliffe (1995) show that in terms of location of residence, household size, income levels, education, and so on, one cannot reject the hypothesis that the households in the two sub-samples were drawn from the same underlying population. But in terms of consumption levels these two sub-samples varied. Table 1 shows that average consumption levels by

TABLE 1

PER CAPITA MONTHLY EXPENDITURE: LONG AND SHORT QUESTIONNAIRES

Percentiles	Long*	Short
10th	124.97 (3.46)	94.83 (1.77)
20th	193.60 (1.81)	153.35 (1.08)
30th	242.20 (1.79)	204.62 (0.96)
40th	296.78 (1.64)	255.24 (1.05)
50th	358.54 (2.34)	315.25 (1.12)
60th	445.25 (3.68)	382.19 (1.59)
70th	575.17 (4.31)	483.18 (2.29)
80th	730.99 (4.37)	627.97 (3.36)
90th	992.79 (9.65)	864.83 (11.78)
Top	2090.5 (84.78)	2225.4 (99.24)

Source: Republic of El Salvador: *Encuesta de Hogares de Propositos Multiples*, 1994-*III*.

Note: *Estimated standard errors are in parentheses and take into account sampling design.

decile in the two sub-samples differ markedly, with households covered by the short module consuming significantly less than those covered by the long module. (Tests of equality of the percentile averages across surveys give p-values less than 0.1 for the top class and less than 0.05 for all others.) If one were to measure poverty by applying a single poverty line across the two sub-samples, one would conclude that poverty among those covered by the short module is much greater than among those covered by the long module. Because both samples are representative of the same underlying population, this is clearly incorrect.

Example 2: Ecuador—Changes Over Time

In Ecuador, the proportion of people living below the poverty line in 1994 was calculated to be 52 percent in a study of poverty undertaken by the World Bank (World Bank, 1996).[4] Adjusting this poverty line for inflation, and then calculating the poverty rate from a new household survey in 1995, the proportion in poverty appeared to decline significantly from 52 to 45 percent. This was surprising, because it occurred against a backdrop of very sluggish economic growth and no obvious new policies aimed at poverty reduction.

[4]The poverty line that yielded this poverty rate was one of several which the World Bank study employed to gauge the sensitivity of poverty rates to the precise location of the line. The incidence of poverty endorsed as "official" was 35 percent.

Although the 1994 and 1995 Ecuador surveys were both high-quality household surveys, based on the World Bank's Living Standards Measurement Survey (LSMS) model and employing a very similar sampling structure, the definition of the consumption aggregate was not identical in the two years. Some modifications to the questionnaire had taken place: the number of food items broken out increased from seventy-three to ninety-four between 1994 and 1995; and several additional non-food consumption items were added. Possibly, the changes in survey design produced a (misleading) appearance of a drop in poverty.

In both of the preceding practical examples, there is a presumption that the groups being compared are similar. In the first because the samples were drawn randomly from the same population, and in the second because the samples were drawn from two adjacent time periods, between which there had been no expectation of a marked change in poverty.

Example 3: Ecuador, Nepal, Brazil—Variation in Non-Food Expenditure Definitions

In Table 2 we demonstrate the effect of ignoring changes in consumption definitions in the clearest way by making comparisons across groups which we know to be not only similar but identical, because they are the same households. Each column shows a set of headcount poverty rates for the indicated country, calculated from an LSMS-style household survey. For each country a single poverty line was developed based on the consumption definition provided in the first row—food and (only) basic non-food expenditure. The headcount rate calculated using this definition is treated as the numeraire. Each subsequent row gives the relative size of the headcount rated when based on the consumption aggregate indicated. Clearly if one expands the items included in consumption while at the same time leaving the poverty line unchanged, measured poverty falls. Our interest, therefore, is in variations in non-food expenditure which could be difficult to determine and which are often overlooked by analysts. The consumption definitions actually employed in different countries easily span the range described in

TABLE 2

THE HEADCOUNT USING A CONSTANT POVERTY LINE WITH ALTERNATIVE CONSUMPTION AGGREGATIONS

Consumption Aggregate	Ecuador	Nepal	Brazil
Food spending plus basic non-food spending	1.00	1.00	1.00
Food plus basic non-food spending including energy and education spending	0.85	0.91	0.89
Above with actual or imputed water expenditures*	0.81	n/a	n/a
Above with actual or imputed value of housing services*	0.70	0.77	0.65
Above with imputed value of owned consumer durables	0.68	0.76	n/a

Sources: *Encuesta de Condiciones de Vida*, 1994 (Instituto Nacional de Estatisticos y Censo, Ecuador); Nepal Living Standards Survey, 1995/96 (Central Bureau of Statistics, Nepal); *Pesquisa Sobre Padoes de Vida*, 1996/97 (Fundacao Instituto Brasileiro de Geografia e Estatistica, Brazil).

Note: *Imputations were carried out for those households which did not report usable expenditures (see Hentschel and Lanjouw, 1996).

Table 2, and as the Ecuador example above illustrates, within countries the definitions also typically change over time. Even with this restriction, the fall in calculated poverty is dramatic as more detailed components are included. Holding the poverty line constant, the headcount rate using the most comprehensive consumption aggregate is at most three-quarters of the rate calculated from an aggregate including food and only basic non-food expenditures. As in the two real examples above, Table 2 clearly demonstrates that comparisons of poverty across varying definitions of consumption when holding a poverty line fixed in real terms would be very misleading.

3. Poverty Measurement with Varying Consumption Definitions: Theory

In this section we show that there exists an approach to deriving a poverty line as an explicit function of the consumption definition, such that poverty rate estimates do not vary due to differing underlying consumption definitions. We consider a range of poverty measures. The approach builds on Engel's Law and assumes deterministic expenditure patterns. However, the effect of allowing for deviations from the assumed expenditure relationships, or measurement error in the data, is discussed at the end of the section.

We set out the simplest case of moving from per-capita food expenditure, F, alone as the welfare indicator to a more comprehensive measure, Y, which includes food plus non-food expenditures: $Y = F + NF$.[5] Let food be a continuous and monotonically increasing function of total expenditure, $F = f(Y)$, and similarly for non-food spending, $NF = g(Y)$. It will be convenient below to define the inverse function $Y = f^{-1}(F) = k(F)$, with $k(0) = 0$. We invoke Engel's Law in assuming $k' > 0$ and $k'' \geq 0$—total per-capita expenditures rise at a non-decreasing rate with food. This implies that the budget share devoted to food is constant or declines as total expenditures increase.

Let z denote the per-capita food poverty line, which is the cost of purchasing a minimum food basket. Typically, this line provides for an energy intake considered to be just adequate given the consumption patterns of low-income households. Let Z be the final per-capita poverty line, which allows for a certain amount of "essential" non-food spending over and above what is necessary to meet a person's nutritional requirements. Because it is difficult to set objective criteria as to what constitutes essential non-food consumption, and because few surveys collect information on actual quantities of non-food items purchased, the non-food component of the final poverty line is usually obtained by scaling up, by some factor, the food poverty line.[6] In practice, procedures differ in how this

[5]Our discussion and the calculations presented below are in per-capita terms, unadjusted by equivalence or economies of scale parameters. The approach would generalize to an expenditure function which included demographic characteristics. One would obtain a set of poverty lines linked to household characteristics.

[6]There are some examples where, rather than scaling up the food poverty line, a final poverty line is reached by specifying explicitly which non-food items should be added to the food poverty line (see Marnie and Micklewright, 1993, for an example with reference to Uzbekistan). This approach is generally regarded as ad-hoc and there is certainly no reason to expect poverty rate calculations to be robust to differences in aggregation of the consumption indicator when this method is used.

is done. In the poverty line developed by Orshansky for the United States, the basic food poverty line was scaled up by a factor of three, based on the empirical observation that in the United States approximately 75 percent of the average household's budget was spent on non-food items (Orshansky, 1963, 1965). As pointed out by Deaton (1997) the choice of this scalar was quite arbitrary and is not terribly intuitive.

Ravallion (1994a, 1998) proposes two alternatives, both of which differ from the Orshansky approach in that the determination of required non-food expenditure is based on the expenditure patterns of the poorer members of the population. The first, "austere" approach entails finding the amount normally spent on non-food items by those households whose *total* expenditure, Y, is just equal to the *food* poverty line, and adding this amount to the food poverty line. The idea is that because these households are sacrificing essential food consumption in order to acquire a certain number of non-food items, they must view these items as essential. The second, "upper bound" approach is to scale up the food poverty line by the amount spent on non-food by households whose actual food expenditures equal the food poverty line, z.

It is only the upper bound method that yields robust comparisons. To see this, first note that in this case:

$$Z = k(z). \tag{1}$$

Let $I = F$, Y be a welfare measure and $l = z, Z$, the corresponding poverty line. If $h(F)$ is the probability density function of food spending in the sample, then measured poverty is:

$$P(I, l) = \int_0^{\infty} p(I, l)h(F)\, dF \tag{2}$$

where $p(I, l)$ is a household poverty indicator. The change in measured poverty in moving from one consumption definition to another is then:

$$P(Y, Z) - P(F, z) = \int_0^{\infty} [p(Y, Z) - p(F, z)]h(F)\, dF. \tag{3}$$

We examine how estimated poverty changes with the consumption definition, using the Foster–Greer–Thorbecke (1984) class of poverty indicators. We consider separately the common headcount (the FGT indicator with a parameter of 0) and the remaining poverty indicators (taking parameter values greater than or equal to one).

Headcount:
$$p(I, l) = \begin{cases} 1 & \text{if } I < l; \\ 0 & \text{else.} \end{cases}$$

Since $k' > 0$, $F < z$ if and only if $Y = k(F) < k(z)$. Therefore, $p(Y, Z) - p(F, z) = 0$ for all F.

Hence, the headcount ratio does not change as the consumption definition changes—the same individuals are considered poor.

FGT Measures ($\alpha \geq 1$):
$$p(I, l) = \begin{cases} [1 - I/l]^{\alpha} & \text{if } I < l; \\ 0 & \text{else.} \end{cases}$$

If $F \geq z$ then again, $Y = k(F) \geq k(z)$ and $p(Y, Z) - p(F, z) = 0$. If $F < z$ then $Y < k(z)$ and

$$p(Y, Z) - p(F, z) = \left[1 - \frac{Y}{k(z)}\right]^{\alpha} - \left[1 - \frac{F}{z}\right]^{\alpha}. \tag{4}$$

The non-increasing food share implies:

$$\frac{F}{Y} \geq \frac{z}{k(z)} \tag{5}$$

so

$$\frac{F}{z} \geq \frac{Y}{k(z)} \tag{6}$$

and $p(Y, Z) - p(F, z) \geq 0$.

The FGT measures may rise as the consumption definition expands. As in the case of the headcount, the same individuals are considered poor under either definition of consumption. However, the relative distance between those poor and the poverty line may increase as the consumption aggregate becomes more comprehensive.

Thus, under the given set of assumptions about expenditure patterns, there exists a method for setting the poverty line which will yield non-varying estimates of the incidence of poverty when different definitions of consumption are applied. We have also seen that, while the approach described here leaves the headcount unchanged as the consumption aggregate expands, indicators of poverty in the FGT class which are sensitive to depth are not similarly robust. In particular, they are likely to increase as the consumption definition expands. A similar argument would hold for inequality measures. If Engel's Law holds over most of the expenditure distribution and continues to hold with the addition of non-food components to the aggregate measure, inequality may increase with a more comprehensive definition of consumption.

The exposition in this section has been in terms of just two definitions of consumption expenditures: food and a composite of both food and non-food expenditures. If one is interested in comparing two different definitions of consumption restricted to food, or both including non-food items, the foregoing analysis may also apply. One requires that there be, for each consumption definition, a behavioral regularity analogous to Engel's Law regarding the share of the expenditure going to some set of items common to both. The case of El Salvador discussed below is an example comparing two different food expenditure definitions, where a set of staple foods has the required relation to other food purchases.

Aggregation and Error

The derivation of the robustness result above hinges crucially on strong assumptions—that expenditures follow an Engel relationship and that they are accurately measured. In fact, households' observed expenditures will deviate from that predicted by an Engel model both because of mismeasurement and because

of variation in households' expenditure patterns. For example, a household may get particular enjoyment from housing and spend more on it than the "typical" household with the same total expenditure.

Consider first mismeasurement. The introduction of noise means that poverty rates may no longer be invariant as the consumption definition expands. If, for example, observed expenditure (in logs) is equal to the true value plus an independently and normally distributed error term, then the density function of observed expenditure has the same shape as that of true expenditure but with fatter tails. As Ravallion (1988) has shown, if the poverty line is located to the left of the mode of these two distributions, the incidence of poverty measured from observed expenditure will be higher than that measured from true expenditure, and vice versa.[7] Thus, although it would not matter in the absence of error, one might want a more (or less) comprehensive indicator if expenditures differ in how accurately they are observed. Anand and Harris (1989) argue that durable non-food expenditures are likely to be measured with more error than food expenditure, and suggest that the latter, less comprehensive, consumption aggregate should be used in poverty analysis. Lusardi (1996), on the other hand, notes that in the United States the extent of measurement error in food consumption seems to be a critical problem. If measurement errors are at least partially independent across components, there is a trade-off between the benefits from aggregating across components (so that errors can cancel each other out) and the potential danger from adding further imprecisely measured components.

To get some intuition about magnitudes, let $F = sY$ and $NF = (1-s)Y$ be the levels of food and non-food predicted by Engel's Law for a household with total expenditure, Y. Let $f = F\varepsilon$ and $\eta f = NF\eta$ be observed food and non-food expenditures, respectively, with

$$\varepsilon \sim N(1, \sigma_\varepsilon^2), \qquad \eta \sim N(1, \sigma_\eta^2).$$

Observed total income is $y = Y[s\varepsilon + (1-s)\eta]$. It is straightforward to show that households are more likely to be appropriately classified as poor or non-poor based on y rather than f alone iff

$$\sigma_\varepsilon^2 > s^2\sigma_\varepsilon^2 + (1-s)^2\sigma_\eta^2 + s(1-s)\sigma_{\varepsilon,\eta}. \tag{7}$$

Consider the simplest case where the measurement errors are independent, $\sigma_{\varepsilon,\eta} = 0$.[8] Then the condition simplifies to y being preferred to f iff:

$$\left[\frac{1+s}{1-s}\right]\sigma_\varepsilon^2 > \sigma_\eta^2, \tag{8}$$

or $\mathrm{Var}(f/F) > \mathrm{Var}(y/Y)$.[9]

[7]Poverty measures that belong to the FGT class are overestimated regardless of where the poverty line is relative to the mode of the distribution (Ravallion, 1988; see also, 1994b).

[8]A positive correlation between measurement errors would diminish the advantages of aggregation, but one might expect a low correlation. Questions about household expenditures are often answered by different family members—in LSMS surveys interviewers are generally instructed to ask questions of those who are best placed to answer them (Grootaert, 1986). In addition, at least some of the non-food expenditures which can be included in the consumption aggregate consist of imputed expenditures for housing, consumer durables, or certain public services, rather than household responses (see for example, Hentschel and Lanjouw, 1996).

[9]See Lanjouw and Lanjouw (1996) for details. For a similar result see Deaton and Zaidi (1999).

The term in square brackets is always greater than one, which means that it may be advantageous to use non-food expenditure information even if it is somewhat noisier than food expenditure information. This is the benefit of aggregation. The data for Ecuador suggest that non-food information can, in fact, be substantially noisier and still be useful. There the term in square brackets is equal to approximately four (using a food share of 0.60, see Table 3, rows 1 and 2). In other words, as long as the variance in the error associated with non-food expenditure is no more than four times that associated with food, aggregating these two components into a comprehensive measure of consumption provides a better indicator of welfare than food alone.

In addition to the effects of mismeasurement, our robustness result will also not hold in general if households deviate from the assumed deterministic relationship between different components of expenditure. If deviations are random, like the measurement error described above, then they will have the same effect as that error, although in this case it is always the more comprehensive measure which best captures household welfare. But at the subgroup level deviations may well be systematic—for example, better educated parents tend to spend relatively more on the education of their children, at every level of income, than less educated parents. When deviations are systematic, the effect on comparisons across groups of increasing the comprehensiveness of the consumption measure to include such expenditures can be marked, and again the more comprehensive measure best captures the relative welfare of different households. We will see an example of both situations below.

4. Implementation in Practice: Illustrations

Given the reality of error, will the robust approach to poverty measurement be successful in addressing the types of problems illustrated in Section 2? We saw in the first example, from El Salvador, a household survey where two very different food consumption modules were given to non-overlapping samples from the same underlying population. However, with reference to four of the items included in the official food poverty line basket the two consumption modules coincide exactly: corn tortilla, bread, beans and rice. Because these are staples, expenditures on these goods are likely to have an Engel curve relation to more comprehensive measures of expenditure.

To implement our robust approach, we define an abbreviated food poverty line based on only these four items. Average expenditure on these four items by households in the bottom 40 percent of the income distribution was 82.1 colones per person per month, which we take as our food poverty line, z.[10] The robust final poverty lines, Z, derived from this (abbreviated) food poverty line are 575 colones per person per month for those households covered by the short consumption module, and 667 colones per person per month for those covered by the long consumption module. Each line is calculated non-parametrically by

[10]This is purely for the purpose of illustration. A serious attempt to measure poverty in El Salvador using this dataset would require more care in the specification of the abbreviated food poverty line.

taking average total consumption among sample households with food expenditure within 1 percent of z, within 2 percent of z, in increasing bands to within 5 percent of z. The total poverty line, Z, is then the average of these values (see Equation 1).

From the theory in the previous section, we would expect that our approach would yield estimates of the headcount rate which is the same for the two samples (given that we cannot reject the assumption that these are drawn from the same population on the basis of a range of indicators aside from consumption). This is indeed the case; 72 percent of the population is poor in El Salvador irrespective of the consumption definition being used.

In our second example in Section 2 we described two household surveys in Ecuador, fielded in 1994 and 1995 respectively. We implement the robust approach described above by specifying a food poverty line based on the subset of food items included in both surveys. This yields two different final poverty lines corresponding to 105,550 sucres and 181,402 sucres per capita per month in 1994 and 1995, respectively. The difference is not only the result of inflation over the period, but also reflects the fact that the 1995 survey embodies a more comprehensive consumption definition than the 1994 survey. On the basis of these poverty lines, the incidence of poverty in Ecuador increased from 52 to 56 percent between 1994 and 1995 (although this increase is not statistically significant). This stands in sharp contrast to the observation that poverty *fell* (from 52 to 45 percent) when only inflation is adjusted for—and accords far more closely with the general view that living standards did not improve between 1994 and 1995.[11]

We next return to the three-country example discussed in Section 2—Ecuador, Nepal and Brazil—to see whether the drops in estimated poverty with increasingly comprehensive data shown in Table 2 are successfully dealt with using the robust approach to poverty line construction. It should be emphasized that this controlled empirical examination is important because the assumptions required to derive the robustness result of Section 3 are unlikely ever to be fully satisfied. One can see this for Ecuador in Figure 1, which shows parametric and non-parametric estimates of the Engel curve and associated 95 percent confidence bounds. There is a good deal of variation around the curves and the food share even increases at low levels of income (as in many other datasets; for a survey, see Thomas, 1986). The main reason for this departure from Engel's Law is a scatter of households with moderate incomes but close to zero food expenditure—which suggests that there is substantial mismeasurement of food expenditure.[12]

Nevertheless, headcount poverty rates calculated as suggested in Section 3 are remarkably stable across different expenditure definitions. Table 3 gives detailed results for Ecuador in the first three columns, including average expenditure at each level of aggregation, the robust poverty line level and the corresponding headcount rates. The final two columns give headcounts calculated in a similar way for Nepal and Brazil. No pair of consumption definitions leads to poverty

[11]Demery and Mehra (1998) have successfully implemented the approach described in this paper to overcome rather more serious problems of definitional changes across survey years in Ghana.

[12]One explanation is that the survey recording period may not correspond to the period in which households purchase food. (See Lanjouw and Lanjouw, 1996, for a proof that an inappropriate recording period can lead to a non-monotonic Engel curve.)

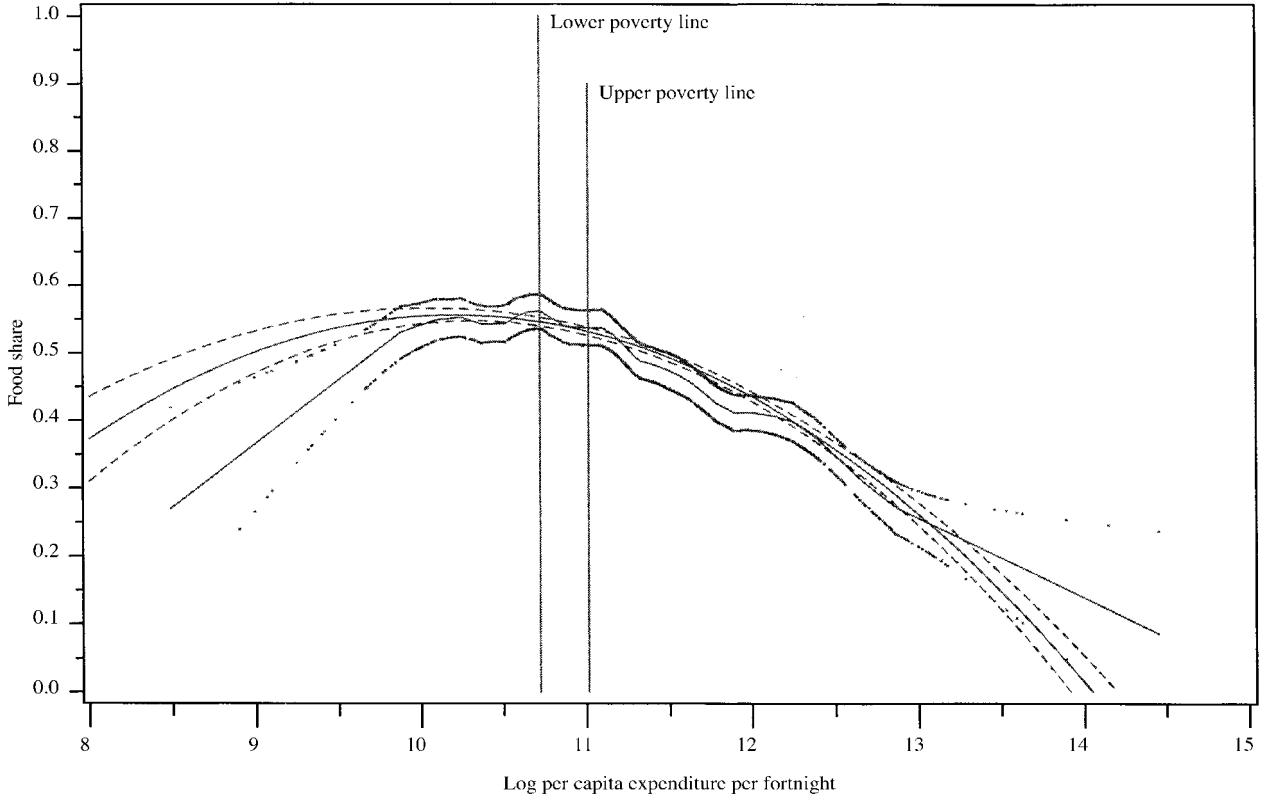

Figure 1. Food Share Regression for Ecuador—1994

Legend: Quadratic Engel curve with 95 percent confidence interval and non-parametric regression curve with 95 percent confidence interval. The non-parametric regression applies the normal kernel estimator with a bandwidth of 0.1917 and a c-value of 0.9599.

TABLE 3

USING THE ROBUSTNESS RESULT: THE HEADCOUNT USING ALTERNATIVE CONSUMPTION AGGREGATIONS

	Ecuador			Nepal	Brazil
Consumption Aggregate	Average Fortnightly Per Capita Consumption	Robust Poverty Line	Headcount	Headcount	Headcount
Food spending	36,917	30,728	0.50 (0.02)[1]	0.44	0.71
Food spending plus basic non-food spending	61,600	44,057	0.53 (0.02)	0.43	0.67
Food plus basic non-food spending including energy and education spending	69,390	47,843	0.52 (0.02)	0.44	0.67
Above with actual or imputed water expenditures[2]	71,696	50,012	0.52 (0.02)	n/a	n/a
Above with actual or imputed value of housing services[2]	80,992	55,884	0.53 (0.03)	0.45	0.68
Above with imputed value of owned consumer durables	84,315	56,775	0.52 (0.03)	0.45	n/a

Sources: See Table 2.

Notes:

[1]Standard errors (in parentheses) take into account stratification and clustering in the surveys (see Howes and Lanjouw, 1998).

[2]Imputations were carried out for those households which did not report usable expenditures (see Hentschel and Lanjouw, 1996).

rates that are more than three percentage points apart, even including comparisons where one consumption definition includes only food. The implication is that, in the absence of great changes in spending patterns, one could successfully monitor temporal changes in poverty at the country level based on a survey of just food expenditures.

What happens if alternative approaches are used? Table 4 gives headcount rates for Ecuador using, first, a constant poverty line and then a line calculated using the "austere" alternative of scaling (see Section 3). It then gives the poverty rate calculated using the robust poverty line but using an FGT poverty measure (with a parameter value of 2) rather than a headcount. The final column gives Gini coefficients calculated using the distribution of the indicated consumption aggregates over the whole population.

Contrary to the stable estimates found using the robust approach, when using the "austere" method of poverty line construction the incidence of poverty falls dramatically as the consumption definition expands: from 50 percent when only food expenditures are used to as low as 35 percent when the full consumption aggregate for Ecuador is applied. Even between consumption aggregates which include both food and non-food spending, the decline in measured poverty is as large as ten percentage points.

On the other hand, contrary to our theoretical findings, the FGT2 measure, like the headcount, happens also to remain stable with aggregation, when using

TABLE 4

OTHER MEASURES USING ALTERNATIVE CONSUMPTION AGGREGATIONS—ECUADOR

Consumption Aggregate	Constant Poverty Line Headcount	Austere Poverty Line[1] Headcount	FGT2	Gini Coefficient
Food spending	n/a	0.50 (0.01)[2]	0.10 (0.02)	0.359
Food spending plus basic non-food spending	0.53 (0.02)	0.45 (0.02)	0.10 (0.008)	0.420
Food plus basic non-food spending including energy and education spending	0.45 (0.02)	0.40 (0.02)	0.10 (0.007)	0.421
Above with actual or imputed water expenditures[3]	0.43 (0.02)	0.39 (0.02)	0.09 (0.007)	0.415
Above with actual or imputed value of housing services[3]	0.37 (0.02)	0.36 (0.02)	0.09 (0.007)	0.420
Above with imputed value of owned consumer durables	0.36 (0.02)	0.35 (0.02)	0.09 (0.007)	0.430

Source: *Encuesta de Condiciones de Vida*, 1994 (Instituto Nacional de Estatisticos y Censo, Ecuador).

Notes:

[1]The constant poverty line corresponds to the robust poverty line derived from a food and basic non-food expenditure definition.

[2]Standard errors (in parentheses) take into account stratification and clustering in the surveys (see Howes and Lanjouw, 1998).

[3]Imputations were carried out only for those households which did not report usable expenditures (see Hentschel and Lanjouw, 1996).

the robust method of scaling the poverty line. Under the assumptions in Section 3 we would expect it to increase with aggregation. This empirical result is likely to be due to the mismeasurement seen in Figure 1: little food expenditure is attributed to some middle income households, causing the depth of poverty measured by F alone to be overstated.

Considering the full distribution of expenditure as captured in measured inequality we do see the expected increase. The Gini coefficient ranges from 0.359 if only food expenditures are taken into account (including an imputed value for home-consumed food production) to a Gini of 0.430 when the fullest definition of consumption is employed. Clearly the definition of consumption can also have an important influence on estimated inequality. The difficulty of ensuring the comparability of indicators, particularly across countries, implies that comparisons of inequality should also be approached cautiously.[13]

The Poverty Profile

Does the success of the robust approach when making comparisons of aggregate poverty carry over to population subgroups? Poverty profiles are a useful policy tool because they reveal differences in the relative poverty of certain subgroups of the population. They are constructed by calculating the incidence of poverty (or some other measure of poverty) for population subgroups defined by a range of household characteristics using a common (price-adjusted) poverty

[13]Atkinson and Brandolini, 1999, find a similar sensitivity of inequality measures to data definitions in comparisons among OECD countries.

line.[14] Because profiles, like poverty rates, are based on the household indicators, $p(Y, Z)$, they too remain unchanged with different consumption definitions if the aggregation consistent poverty measure is used, household expenditure patterns follow Engel's Law, the patterns are the same across groups, and there is no measurement error.

TABLE 5

HOUSEHOLD CHARACTERISTICS AND THE RISK OF POVERTY*

Household Characteristics	Food Spending	Food plus Basic Non-Food	Food plus Non-Food (with Education and Energy)	Expenditure Including Imputed Water	Expenditure Including Imputed Water and Housing	Expenditure with Imputed Water, Housing and Durables
Average risk of poverty	**0.50**	**0.53**	**0.52**	**0.52**	**0.53**	**0.52**
Sierra	**0.50**	**0.49**	**0.47**	**0.47**	**0.47**	**0.46**
Costa	**0.49**	**0.49**	**0.48**	**0.48**	**0.49**	**0.49**
Oriente	**0.67**	**0.69**	**0.69**	**0.69**	**0.70**	**0.70**
Rural	**0.55**	**0.60**	**0.60**	**0.61**	**0.63**	**0.63**
Urban	**0.46**	**0.41**	**0.38**	**0.38**	**0.37**	**0.36**
Black and white TV	0.52	0.50	0.49	0.49	0.49	0.48
Color TV	0.33	0.27	0.22	0.22	0.21	0.19
Bicycle	0.49	0.45	0.42	0.42	0.41	0.40
Refrigerator	0.39	0.33	0.29	0.28	0.27	0.26
House with mud walls	0.54	0.53	0.55	0.55	0.56	0.56
House with dirt floor	0.65	0.67	0.69	0.70	0.73	0.74
House with wood walls	0.60	0.68	0.65	0.67	0.72	0.72
Telephone connection	**0.30**	**0.21**	**0.14**	**0.14**	**0.09**	**0.09**
Networked electricity	0.48	0.47	0.45	0.45	0.45	0.44
Networked water	0.44	0.40	0.36	0.36	0.35	0.34
Waste disposal	0.43	0.37	0.34	0.34	0.32	0.31
Sewage removal	0.45	0.41	0.38	0.38	0.37	0.36
Head with no education	0.61	0.68	0.67	0.68	0.69	0.70
Primary educated head	0.55	0.57	0.56	0.57	0.58	0.57
Secondary educated head	0.40	0.35	0.30	0.30	0.30	0.29
Tertiary educated head	**0.28**	**0.18**	**0.14**	**0.13**	**0.11**	**0.10**
Indigenous head	0.76	0.76	0.76	0.76	0.77	0.78

Source: *Encuesta de Condiciones de Vida*, 1994 (Instituto Nacional de Estatisticos y Censo, Ecuador).

Note: *Figures indicate the headcount rate of poverty among households with the row characteristic and calculated with the column consumption aggregate.

A set of simple poverty profiles for Ecuador is presented in Table 5. Each column gives a profile based on the indicated definition of consumption using the robust approach to setting the poverty line in conjunction with expenditure patterns in the population. There is a remarkable degree of stability in headcount poverty estimates across definitions of consumption for many of the subgroups. Estimated poverty is stable across all consumption aggregates for the three main sub-regions of the country. For groups defined by urban versus rural residence we find that poverty rates are robust across all consumption aggregates which include any form of non-food consumption. Poverty rates based on food consumption alone do differ substantially, however, from those which include non-food spending. Between these groups, there are clearly differences in the relative

[14]One can also estimate the relationship between the household indicator of poverty and various household characteristics simultaneously (on the basis of a probit or logit model, for example), but we shall confine our illustrations to separate comparisons of poverty and household characteristics, in turn.

importance of food in total expenditure, with urban residents spending more on non-food goods at a given level of total expenditure. Because these groups deviate *systematically* from a population-based Engel curve relation in this respect, one would not want to make poverty comparisons across them using a common poverty line, basing the comparison only on food expenditures. This conclusion applies also for other subgroups. In a few cases, such as households with a telephone connection and those with tertiary education, deviations from the population-based Engel relationship persist over even broader definitions of consumption.

Taken together, these poverty profiles suggest that a simple consumption definition comprising food and only basic non-food expenditures may be sufficient to make subgroup comparisons as accurate as those based on more detailed information. It follows that to track changes in the poverty profile over time one could periodically field surveys which collect only this basic consumption information (assuming, as in the case of aggregate poverty, no changes in expenditure patterns over time). This is encouraging, but one need not rely on this example. In any given situation, an initial full-scale survey can be checked, as here, to determine what level of consumption detail is necessary for comparisons across the population subgroups of interest to policymakers.

5. Concluding Remarks

This paper has argued that comparisons of poverty rates derived from varying definitions of consumption are potentially misleading. We have shown that the magnitude of error introduced can be substantial, even when definitions appear to resemble each other closely and would often be treated as the same in practice.

Under certain assumptions, theory indicates how measured poverty will evolve as the definition of consumption is expanded to include more items and services. These assumptions are: that the relationship between two definitions of consumption follows Engel's Law; that consumption patterns are the same across groups being compared; and that there is no measurement error. Given these, we show how one can make comparisons that are robust to alternative definitions of consumption by using the headcount measure and the specific approach to setting a poverty line outlined in Section 3.

The assumptions underpinning the theoretical results in this paper are clearly strong, and are unlikely ever to hold empirically. Nevertheless, we have shown on the basis of a series of data sets from very different countries that the approach is remarkably effective in practice. This finding should be useful to those wanting to use sets of poverty measures, in econometric studies or for policy design, but who are confronted with the very real problem of having data from surveys of varying design.

Our robustness result points to the possibility of monitoring poverty using abbreviated, low-cost survey—that is, purposefully creating data which are not comparable in order to lower the costs of collection. The approach one could take is as follows. In the first stage, a full-sized expenditure survey is fielded. This survey is analyzed to set a baseline poverty rate, and to discover what abbreviated

consumption definition, combined with the aggregation-consistent approach to setting the poverty line, yields the same headcount rate as with the most comprehensive consumption definition. It can also be checked, at this stage, whether poverty rates for the specific population subgroups of interest to policymakers are also stable between the full and the abbreviated consumption definition. Subsequent surveys are then fielded which gather information only on those consumption components necessary for the abbreviated definition. The headcount rates stemming from these simpler and cheaper surveys can then be compared to the rates arising from the initial full-scale survey.

The approach does have limitations. It is important to stress that it relies heavily on stable consumption patterns. Eventually, as the interval between the first stage survey and subsequent monitoring surveys widens, the underlying assumption that consumption patterns have remained unchanged becomes less tenable, and it becomes necessary to field another full-scale household survey. Similarly, one can imagine circumstances where, for example, due to drought, war, or other fundamental changes in relative prices between food and non-food goods, the relationship between food consumption and total resources could change rapidly. In such circumstances, it is only possible to monitor poverty by fielding comprehensive (and identical) surveys on a regular basis, holding the poverty line constant in real terms.

References

Anand, S. and C. Harris, Food and the Standard of Living, in J. Drèze and A. Sen (eds), *The Political Economy of Hunger*, *Vol 1: Entitlement and Well-Being*, Clarendon Press, Oxford, 1989.

Atkinson, A. B., *Poverty and Social Security*, Harvester-Wheatsheaf, Hemel Hempstead, 1989.

———, Comparing Poverty Rates Internationally: Lessons from Recent Studies in Developed Countries, *the World Bank Economic Review*, 5(1), 1991.

Atkinson, A. B. and A. Brandolini, Promise and Pitfalls in the Use of "Secondary" Data-Sets: Income Inequality in OECD Countries, mimeo, Nuffield College, Oxford University, UK, 1999.

Bidani, B. and M. Ravallion, How Robust is a Poverty Profile?, *the World Bank Economic Review*, 8, 1994.

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

Deaton, A. and A. Tarozzi, Price and Poverty in India, mimeo, Research Program in Development Studies, Princeton University, 2000.

Deaton, A. and S. Zaidi, Guidelines for Constructing Consumption Aggregates for Welfare Analysis, mimeo, Development Economics Research Group, the World Bank, 1999.

Deininger, K. and L. Squire, A New Data Set Measuring Income Inequality, *the World Bank Economic Review*, 10(3), 1996.

Demery, L. and K. Mehra, Measuring Poverty Over Time: Dealing with Uncooperative Data in Ghana, mimeo, Poverty Reduction and Economic Management Unit, the World Bank, 1998.

Foster, J., J. Greer and E. Thorbecke, A Class of Decomposable Poverty Measures, *Econometrica*, 52, 1984.

Government of India, *Report of the Expert Group on Estimation of Proportion and Number of Poor*, Government of India Planning Commission, New Delhi, 1993.

Grootaert, C., Measuring and Analyzing Levels of Living in Developing Countries: An Annotated Questionnaire, Living Standards Measurement Survey Working Paper No. 24, the World Bank, 1986.

Hentschel, J. and P. Lanjouw, Constructing an Indicator of Consumption for the Analysis of Poverty: Principles and Illustrations with Reference to Ecuador, Living Standards Measurement Survey Working Paper No. 124, the World Bank, 1996.

Howes, S. and J. O. Lanjouw, Does Sample Design Matter for Poverty Comparisons?, *Review of Income and Wealth*, Series 44, No. 1, 99–109, 1998.

Jorgensen, D., Did We Lose the War on Poverty?, *Journal of Economic Perspectives*, 12(1), 1998.

Lanjouw, J. O. and P. Lanjouw, Comparing Poverty With Non-Identical Consumption Aggregates: Theory and Illustrations from Ecuador and Pakistan, mimeo, PRDPH, the World Bank, 1996.
Lusardi, A., Permanent Income, Current Income and Consumption: Evidence from Two Panel Data Sets, *Journal of Business and Economic Statistics*, 14(1), 1996.
Marnie, S. and J. Micklewright, A Note on the Subsistence Minimum Income Calculated by the Ministry of Labour, Uzbekistan, mimeo, European University Institute, Florence, 1993.
National Research Council, *Measuring Poverty: A New Approach*, National Academy Press, Washington, DC, 1995.
Orshansky, M., Children of the Poor, Social Security Bulletin, Vol. 26, Department of Social Security, United States Government, 1963.
———, Counting the Poor: Another Look at the Poverty Profile, Social Security Bulletin, Vol. 28, Department of Social Security, United States Government, 1965.
Ravallion, M., Expected Poverty Under Risk-Induced Welfare Variability, *The Economic Journal*, 98, 1988.
———, *Poverty Comparisons*, Harwood Academic Publishers, Chur, 1994a.
———, Poverty Rankings Using Noisy Data on Living Standards, *Economics Letters*, 45, 1994b.
———, Poverty Lines in Theory and Practice, mimeo, Development Economics Research Group, the World Bank, 1998.
Rocha, S., Poverty Lines for Brazil: New Estimates from Recent Empirical Evidence, mimeo, IBGE/POF, Brazil, 1993.
Scott, K. and D. Jolliffe, A Comparison of Alternative Consumption Modules in the El Salvador Encuesta de Propositos Multiples, mimeo, PRDPH, the World Bank, 1995.
Summers, R. and A. Heston, A New Set of International Comparisons of Real Product and Price Levels Estimates for 130 Countries, 1950–1985, *Review of Income and Wealth*, 1–25, 1988.
———, The Penn World Table (Mark 5): An Extended Set of International Comparisons, 1950–1988, *Quarterly Journal of Economics*, 106, 1991.
Thomas, D., *The Food Share as a Welfare Measure*, PhD Dissertation, Princeton University, 1986.
World Bank, *Ecuador Poverty Report*, Report No. 14533-EC, 1996.