Economic Systems Research, Vol. 15, No. 2, June 2003

Routledge
Taylor & Francis Group

Household Surveys, Consumption, and the Measurement of Poverty

ANGUS DEATON

ABSTRACT *Household surveys are playing an increasingly important role in the measurement of poverty and well-being around the world. The Living Standards Measurement Study, which was begun in the World Bank under the guidance of Graham Pyatt in 1979, has played an important role in this movement. Its surveys are widely used within the Bank to measure consumption-based poverty, and survey data are now the exclusive basis for the global poverty counts. This paper discusses a number of unresolved issues in using consumption-based surveys for measuring well-being, including the choice of a money-metric versus welfare-ratio approach, the collection of suitable price information, the effects of measurement error on estimation, and methods for correcting per capita consumption for the demographic structure of the household.*

KEYWORDS: *Household surveys; consumption; poverty*

1. Introduction

The World Bank's Living Standard Measurement Study (LSMS) began under Graham Pyatt's guidance in 1979. The LSMS thrives to this day, and has been involved in collecting more than 40 household surveys throughout the world. These surveys are used routinely by the World Bank to measure poverty, and the results are an integral part of World Bank policymaking. That the LSMS should have thrived and prospered through more than two decades of continual reorganization and convulsion at the Bank is a remarkable testimony to its founders, particular their clear understanding of the dangers of policymaking in a data vacuum. Twenty years ago, as now, the World Bank focused on poverty as its central mission, but for many of the intervening years, poverty was not of great interest, and senior World Bank staffers saw little need for data to document the importance of market-based reforms. The policy of getting prices right requires little in the way of data support. Yet even through those years the LSMS prospered, because its designers had conceived it broadly enough that its data were useful in many different ways, not only for the original purpose of measuring poverty, but also for the analysis and understanding of the mechanisms that sustain it.

Angus Deaton, Woodrow Wilson School of Public and International Affairs, 347 Wallace Hall, Princeton University, Princeton, New Jersey 08544, USA. E-mail: deaton@princeton.edu. Some of this material previously appeared in my joint work with Salman Zaidi, 'Guidelines for constructing consumption aggregates for welfare analysis', and I am grateful to him for permission to use the work here. I also thank Martin Ravallion for useful discussions on these issues.

ISSN 0953-5314 print; 1469-5758 online/03/020135-25 DOI: 10.1080/0953531032000091144

In 1979, these developments were a long way ahead, and no LSMS data were to be collected until the surveys in Peru and Côte d'Ivoire in the mid-1980s. Yet there was a great deal of preparatory work to be done and Graham asked me to prepare a conceptual paper on the use of household survey data for measuring welfare. I did this, very much under Graham's guidance and with his intellectual support, and my report became one of the first LSMS working papers. I argued that consumption, not income, should be the basis for our measurements, and suggested that the concept of 'money metric utility' was an appropriate basis for linking theory and measurement. Although many country household surveys had routinely collected consumption data and based their poverty estimates on them—India being the prime example—many other countries used income, not consumption. The battle for consumption is not yet entirely won, although even in Latin America, long a bastion of income surveys, there is increasing interest in moving to a consumption basis. However, the LSMS won the battle for consumption based-measures within the World Bank, which now routinely uses consumption whenever possible, not only for its country work, but also for its periodic counts of the number of people in the world who live in households whose consumption per head is less than $1 (or $2) a day at international purchasing power prices.

In this paper, which builds on and updates my original LSMS working paper, I bring together some of the new insights that have come from the two subsequent decades of research on these topics. I start in Section 2 with the original money metric utility concept, and discuss an alternative approach based on 'welfare ratios'. This leads into a discussion of price indexes in Section 3, and how they can be measured within the survey framework. Price indexes are an extraordinarily important component of poverty measurement, and are frequently not given the attention that they deserve. But the price level is not the only factor that needs to be taken into account. Household measures of expenditure, even real expenditure, need to be assessed against household needs in order to derive the individual measures of living standards on which policy must be based. This is the task of measures of household needs, often simple headcounts, but sometimes more sophisticated counts in which different people are given different weights. Recent work on these weights, or 'equivalence scales', is briefly discussed in Section 4. Finally, Section 5 presents a discussion of ways to explore the uncertainty underlying our measures. Poverty lines are always arbitrary, at least in part, and our data, even with instruments such as the LSMS, are subject to substantial errors in measurement.

2. Theory of the Measurement of Welfare

In this section, I discuss briefly the theoretical basis for the consumption-based measure of welfare. My concern here is a fairly narrow one, focusing on an economic definition of living standards. I do not consider other important components of welfare, such as freedom, health status, life expectancy, or levels of education, all of which are related to income and consumption, but which cannot be adequately captured by any simple monetary measure. Consumption measures are limited in their scope, but are nevertheless a central component of any assessment of living standards.

An important concept is *money metric utility* (Samuelson, 1974), which measures levels of living by the money required to sustain them. I start with this in Section 2.1. An alternative approach, based on Blackorby & Donaldson's (1987) concept of *welfare ratios*, whereby welfare is measured as multiples of a poverty line, is presented in Section 2.2. Each of the money-metric and welfare-ratio approaches has its strengths and weaknesses; both start from a nominal consumption aggregate, but adjust it differently. This material updates the discussion in Deaton (1980) in one of the earliest LSMS Working Papers.

2.1. Money-metric Utility

The starting point is the canonical consumption problem in which a household chooses the consumption of individual goods to maximize utility within a given budget and at given prices. Consumer preferences over goods are thought of as a system of indifference curves, each linking bundles that are equally good, and with higher indifference curves better than lower ones. A given indifference curve corresponds to a given level of welfare, well-being, or living-standards, so that the measurement of welfare boils down to labelling the indifference curves, and then locating each household on an indifference curve. There are many ways of labelling indifference curves. One possibility would be to take some reference commodity bundle and to label indifference curves by the distance from the origin of their point of intersection with the bundle. In Figure 1, the reference quantity vector is shown as the line q^0 so that the two indifference curves *II* and *JJ* are labelled as *OA* and *OB* respectively. Instead of a reference set of quantities, we can select a reference set of prices, and calculate the amount of money needed to reach the

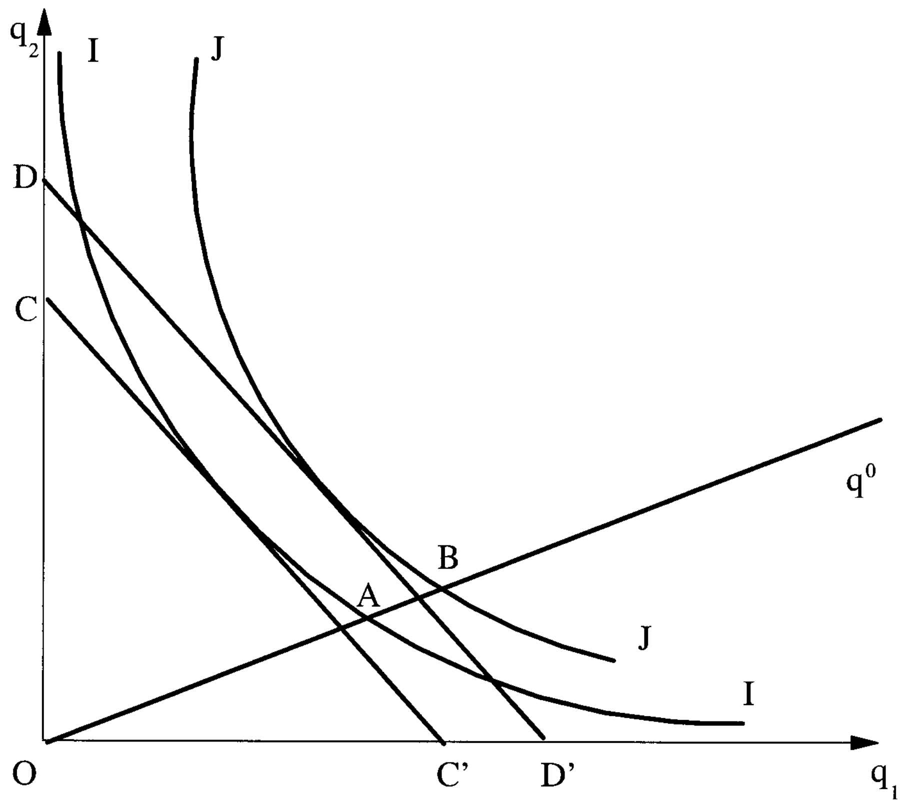

Figure 1. Two ways of labelling indifference curves.

two indifference curves; this is Samuelson's money metric utility. In the figure, money metric utility is constructed by drawing the two tangents to the indifference curves, with the slope set by the reference prices, so that the costs of reaching the curves are OC' and OD' in terms of q_1 or OC and OD in terms of q_2.

To see how this works, we introduce some notation. Write x for total expenditure, and denote by $c(u, p)$ the cost or expenditure function, which associates with each vector of prices p the minimum cost of reaching the utility level u. Since the household maximizes utility, it must minimize cost of reaching u, so that

$$c(u, p) = x \tag{1}$$

Denote by superscript h the household whose welfare we are measuring, and let p^0 denote a vector of reference prices, the choice of which we discuss below. Money metric utility for household h, denoted u_m^h, is defined by

$$u_m^h = c(u^h, p^0) \tag{2}$$

which is the minimum cost of reaching u^h at prices p^0. Note that, although utility itself is, to a large extent, arbitrary, we can label indifference curves any way we choose, as long as higher indifference curves are labelled with larger values of utility. Money metric utility is defined by an indifference curve and a set of prices, is independent of the labels, and is therefore well defined given the indifference curves.

The exact calculation of money metric utility requires knowledge of preferences. Although preferences can be recovered from knowledge of demand functions, we typically prefer some shortcut method that, even if approximate, does not require the estimation of behavioural relationships with all the accompanying assumptions, including often controversial identifying assumptions, and potential loss of credibility. The most convenient such approximation comes from a first-order expansion of $c(u^h, p^0)$ in prices around the vector of prices actually faced by the household, p^h. The derivatives of the cost function with respect to prices are the quantities consumed, a result known as Shephard's Lemma (or Roy's Identity), see for example Deaton & Muellbauer (1980, Chapter 2). In consequence, if we write q for the vector of quantities, we can approximate the cost function as follows

$$c(u^h, p^0) \approx c(u^h, p^h) + (p^0 - p^h) \cdot q^h \tag{3}$$

where the centred '$\cdot$' indicates an inner product. Since the minimum cost of reaching u^h at p^h is the amount spent, $p^h \cdot q^h$, equation (3) can be written as

$$u_m^h = c(u^h, p^0) \approx p^0 \cdot q^h \tag{4}$$

which is the household's vector of consumption items priced at reference prices. Note the convenient link with National Income Accounting Practice, in which real national product would include real consumer's expenditure, which is the sum over all consumers of their consumption valued at base prices, i.e. the sum of the right-hand side of (4) over all agents.

This equation is still not quite in convenient form for practice, since we rarely observe a complete set of *quantities* for each household, and may not even have available a complete set of reference prices. The *Paasche* price index comparing the price vectors p^h and p^0 is defined as

$$P_P^h = \frac{p^h \cdot q^h}{p^0 \cdot q^h} \tag{5}$$

so that, from equation (4), we have

$$u_m^h \approx \frac{p^h \cdot q^h}{P_P^h} = \frac{x^h}{P_P^h} \tag{6}$$

so that money metric utility can be approximated by adding up all the household's expenditures, and dividing by a Paasche index of prices.

When we are working with a single cross-sectional household survey, the price variation is less temporal than spatial; people who live in different parts of the country pay different prices for comparable goods. (If we have two surveys for the same country at different times, or if the survey is spread over months or years, the variation will be both temporal and spatial.) In industrialized countries, where transportation is easy and inexpensive, and there are integrated distribution systems for most consumer goods, spatial price variation is small, housing being the major exception. In many developing countries, however, spatial price differences can be large, in both relative and absolute prices, and it is important to take them into account. In the temporal context, a Paasche price index is one whose (quantity) weights relate to the current period, rather than the base period. In the current spatial context, the 'current period' is replaced by the 'household under consideration', whose purchases are used to weight the prices it faces relative to some base or reference prices. Perhaps the major practical point about equation (5) is that the weights for the prices differ from household to household, so that, for example, two households in the same village, buying their goods in the same markets, and facing the same prices, will have different price indexes if they have different tastes or incomes. At first sight, such a situation may seem hopelessly complicated. The transparency is restored if we think of money metric utility as equation (4), the household's consumption bundle priced at fixed prices, and if we recognize that equation (6), the deflation of nominal expenditure by a Paasche index with household specific weights, as simply a means of calculating the constant price total.

2.2. An Alternative Approach: Welfare Ratios

One of the important uses of measures of standard of living is to support policy, particularly policy where distribution is an issue. In particular, much policy is conducted on the basis that transfers of money are more valuable the lower in the distribution is the recipient. This may take the form of a focus on poverty where the poor are given more attention than the non-poor, or it may be more sophisticated, involving distributional weights that decline as we look at people with higher standards of living. Blackorby & Donaldson (1988) have shown that the use of money metric utility can cause difficulties in this context. To see the problem, start by assuming that total household expenditure (or income) x is a satisfactory measure of living standards, something that would be true if everyone faced the same prices, and everyone lived alone, or at least in households that all had the same size and composition. Monetary transfers then correspond exactly to changes in welfare, so that policymakers who are averse to inequality can work under the assumption that increases in x have a lower social marginal value the higher in the distribution is the recipient. But money metric utility is not x, but a function of x. As Figure 1 makes clear, money-metric utility is higher the higher is

x, so that more money corresponds to a higher indifference curve and standard of living. But what Blackorby and Donaldson show is that, special cases apart, money metric utility is not a *concave* function of x, that the rate at which money metric utility increases with x can be constant, decreasing, or increasing, and that, in general, which is the case depends on the choice of the reference price vector p^0. This has the effect of breaking any close link between redistributive policy and the measurement of its effects. For example, suppose that a change in policy—for example, a transfer policy—has the effect of transferring money from better-off to worse-off households, so that the distribution of money income has become more equal. However, because we do not know exactly how money metric utility is linked to money, there is no guarantee that the distribution of money metric utility has also narrowed. So we have lost the ability to monitor the distributional effects of policy, and what we get when we try will be different at different choices of reference prices p^0. Since we are often forced to use whatever prices are available to us, we may not even be able to control the outcome.

In order to avoid these problems, Blackorby & Donaldson (1987) have proposed the use of a 'welfare ratio' measure in place of money-metric utility; within the World Bank, the use of welfare ratios is reviewed by Ravallion (1998b). The basic idea is to express the standard of living relative to a baseline indifference curve. In poverty analysis, a natural (and useful) choice is the poverty indifference curve, the level of living that marks the boundary between being poor and non-poor. The welfare ratio is then the ratio of the household's expenditure to the expenditure required to reach the poverty indifference curve, both expressed in the prices faced by the household. Once again, Figure 1 can serve to illustrate. If *II* is taken to be the poverty indifference curve, and *JJ* the indifference curve we are trying to measure, then provided the two price lines are taken to illustrate *current*, not reference, prices, the welfare ratio is OD/OC or (equivalently) OD'/OC'. In terms of the cost functions, the ratio is given by

$$wr^h = \frac{c(u^h, p^h)}{c(u^z, p^h)}$$

where u^z is the utility poverty-line, the utility corresponding to the poverty indifference curve.

Unlike money metric utility, which is a money measure—the minimum amount of money needed to reach an indifference curve—the welfare ratio is a pure number—the standard of living relative to poverty. In practice, it is useful to convert the welfare ratio into a money measure, and again the obvious procedure is to multiply the ratio by the poverty line, defined as the cost of obtaining poverty utility at reference prices, $c(u^z, p^0)$. This gives the welfare ratio measure, which we denote by u_r^h.

$$u_r^h = \frac{c(u^h, p^h)}{c(u^z, p^h)} \times c(u^z, p^0) \tag{8}$$

Like the money metric utility measure, equation (8) is total expenditure x^h divided by a price index, in this case the true cost of living index for p^h versus p^0 computed at the poverty line indifference curve. This cost-of-living price index would normally be approximated by the Laspeyres index:

$$P_{Lz}^h = \frac{p^h \cdot q^z}{p^0 \cdot q^z} = \sum_{i=1}^n \frac{p_i^0 q_i^z}{p^0 q^z}\left(\frac{p_i^h}{p_i^0}\right) = \sum_{i=1}^n w_i^{z0}\left(\frac{p_i^h}{p_i^0}\right) \tag{9}$$

where q_i^z is the quantity of i consumed at the poverty line and the weights w_i^z are the shares of the budget at the poverty line indifference curve and prices p^0. Putting equations (8) and (9) together, we get an expression for the money version of the welfare ratio that corresponds to equation (6) for money metric utility:

$$u_r^h = \frac{x^h}{P_{Lz}^h} \tag{10}$$

If we compare equations (6) and (10), we see that money metric utility involves deflation of expenditure by a Paasche index of prices, while the welfare ratio measure involves deflation of expenditure by a Laspeyres price index. (The calculation of the poverty-line weights in equation (9) will be discussed in Section 3.)

In some applications, such as in comparing national price indexes at two moments of time, Paasche and Laspeyres price indexes are close to one another, either because the two sets of weights are similar in the two periods, or because relative prices are similar. In the current context, where we are most often interested in comparing prices between different places, where both weights and relative prices are often quite different, the Paasche and Laspeyres price indexes can also differ substantially, as will therefore the money metric utility and welfare ratio measures. On the theoretical side, the point to note is that the Laspeyres index in equation (10) is computed at the poverty indifference curve, so that its weights (see also equation (9)) are unaffected by changes in total expenditure of household h. As a result, u_r^h is proportional to x^h, and there is a direct link between redistributive policy and the measurement of its effects. Welfare ratios resolve the difficulties of using money-metric utility to monitor the outcomes of distributionally sensitive policies. On the empirical side, the Paasche and Laspeyres indexes will be close to one another when the price relatives are close to one another over different goods and services, or when the weights applied to them are the same at the base, in this case the poverty line, as for other households in the survey. But there is no reason to suppose that either will be true in cross-sectional surveys. Regional price differences are often markedly different across goods depending on agricultural zones or distance from the ocean, and expenditure patterns differ sharply over households of different types, or even across households that have much the same observable characteristics. In practice, as well as in theory, the money-metric and welfare-ratio approaches are likely to give different answers.

How do we choose between the two approaches to welfare measurement? As I have presented it so far, the balance seems to favour the welfare ratio approach. It is simpler to calculate, since the weights for the price index are the same for everyone, and it has a straightforward theoretical link to total expenditure, which facilitates distributional analysis. Deflation of an expenditure measure by a fixed weight Laspeyres index is a procedure that is both simple and transparent and that could be explained and defended to policymakers. For some, those benefits are likely to be decisive. Nevertheless, the welfare ratio approach is not without its own Achilles heel. As Blackorby and Donaldson show, welfare ratios do not necessarily indicate welfare correctly. It is possible for a policy to make someone better off, and yet to decrease their welfare ratio. This cannot happen for money metric utility, no matter which set of reference prices are used in the evaluation. So while money metric utility is more problematic for distributional calculations, the welfare ratio approach throws out at least some of the baby along with the bathwater.

3. Adjusting for Cost of Living Differences

Price indexes are used to aggregate a large number of individual prices into a single number, so that individual prices are the raw material for the indexes. In LSMS and other surveys, there are several possible sources for the prices, see Deaton & Grosh (2000) for further discussion of how prices can be collected and for an analysis of some of the differences between them. In brief, there are three possible sources. The first source is the survey itself, and the reports of purchases by the households surveyed. In many (but not all) surveys, households report both quantities and expenditures for most of the foods they purchase (three kilograms of rice for 5 rupees) as well as for a few non-food items where quantities are well-defined, fuels being the obvious example. Dividing expenditures by quantities gives ‘unit values’. These are affected by quality choices; someone who buys better cuts of meat will pay more per unit, but experience shows that the spatial variation of unit values is closely related to price variation. As a result, unit values provide good price information, especially when averaged over households in a cluster. An example of a large-scale application of this methodology is Deaton & Tarozzi (2000), who use more than 7 million price observations from the 43rd and 50th Rounds of India’s National Sample Survey to calculate price indexes over time, as well as across states, and between the urban and rural sector of each state.

The second source of price information is a dedicated price questionnaire, often administered in each cluster as part of a community questionnaire. The price questionnaire seeks to measure prices in the markets actually patronized by survey households and, in principle, provides a direct measure of what we need. In practice, there may be some compromise of data quality from the fact that the investigators do not actually make purchases. There are also sometimes problems of locating a wide enough range of homogeneous goods in all the relevant markets, so that it may be hard to match prices from the questionnaire with the expenditure patterns of the households in the survey. But this is the preferred source of price information when quantities are not collected from each household, and the only source for those goods, such as most non-food items, and food eaten away from home, where quantity observation is not possible in principle.

The third source of price data is ancillary data, for example from government price surveys. This is typically a source of last resort. Such data are often thin on the ground, and there will often be many households whose nearest observed price is so far away as to be irrelevant. Nevertheless, such data are sometimes the only information available, and it is usually better to use them than to make no correction at all.

Note, finally, that the situation is somewhat different depending on whether we need to compute price indexes over space or over time. In the latter case, for example when we are comparing two surveys for the same country some years apart, there will usually be available some national consumer price index that tells us by how much the general price level has changed between the two surveys. In the absence of spatial data on prices, the temporal index should be used to deflate all nominal expenditures to ensure that welfare comparisons between the two periods are not being driven by inflation.

Before turning to the details, it is useful to begin by recalling the formulas for money-metric and welfare-ratio utilities, whereby each is expressed as total expenditure deflated by a price index. For money metric utility, we have from equation (6) that

$$u_m^h \approx \frac{p^h \cdot q^h}{P_P^h} = \frac{x^h}{P_P^h} \tag{11}$$

where the Paasche price index in the denominator is given by

$$P_P^h = \frac{p^h \cdot q^h}{p^0 \cdot q^h} \tag{12}$$

Here, the weights for the price index are the quantities consumed by the household itself and therefore differ from one household to another. By contrast, welfare-ratio utility uses a Laspeyres index so that, from equation (10)

$$u_r^h = \frac{x^h}{P_{Lz}^h} \tag{13}$$

where, if we are using the poverty line as the base, the Laspeyres is given by equation (9)

$$P_{Lz} = \frac{p^h \cdot q^z}{p^0 \cdot q^z} = \sum_{i=1}^{n} w_i^{z0} \left(\frac{p_i^h}{p_i^0} \right) \tag{14}$$

Most of past practice has been based on using Laspeyres indexes for adjustment, although not always with weights tailored to the poverty line as in equation (14), and relatively little attention has been given to the calculation of the Paasche index. In this section, I focus on the calculation of equations (12) and (14) using the data from a typical LSMS survey.

3.1. Paasche Price Index

Equation (12) can be rewritten in the form:

$$P_P^h = \left(\sum w_k^h \left(\frac{p_k^0}{p_k^h} \right) \right)^{-1} \tag{15}$$

where w_k^h is the share of household h's budget devoted to good k. This formula can be calculated from expenditure data and price relatives alone. The following approximation is also sometimes useful:

$$\ln P_P^h \approx \Sigma w_k^h \ln \left(\frac{p_k^h}{p_k^0} \right) \tag{16}$$

Note that these indexes involve not only the prices faced by household h in relation to the reference prices, but also household h's expenditure pattern, something that is not true of a Laspeyres index. The distinction is an important one; to convert total expenditure into money metric utility, the price index must be tailored to the household's own demand pattern, a demand pattern that varies with the household's income, demographic composition, location, and other characteristics.

The reference price vector p^0 is inevitably selected as a matter of convenience, but should not be very different from prices actually observed. A good choice is to take the median of the prices observed from individual households (for foods and fuels, if unit values are collected) or from the community questionnaire (otherwise). Especially when using the unit values from individual records, there will be some outliers, not only for the usual reasons, but also because there are often misunderstandings about units—such as eggs being reported in dozens instead of in units. Use of medians rather than means reduces sensitivity to such accidents. The use of a national average price vector ensures that the money metric measures conform as closely as possible to national income accounting practice, as well as eliminating results that might depend on a price relative that occurs only rarely or in some particular area.

In general, even if quantities and unit values are available at the household level, this will only be the case for a limited set of goods, typically foods and perhaps some fuels. For non-foods, and perhaps some foods, price relatives will come from community questionnaires or even other regional sources, and will not be available at the household level. In such cases, we must use the price relative that seems most appropriate for each household, in which case equation (16), for example, becomes

$$\ln P_P^h = \sum_{k \in F} w_k^h \ln\left(\frac{p_k^h}{p_k^0}\right) + \sum_{k \in NF} w_k^h \ln\left(\frac{p_k^c}{p_k^0}\right) \tag{17}$$

where F denotes the set of goods (foods) for which we have individual household price relatives, and NF is the set where we do not (non-foods), and the superscript c denotes a cluster or regional price. One further refinement is likely to be useful. Because the household level unit values are likely to be noisy, and to contain occasional outliers, it is wise to replace the individual p_k^h by their medians over households in the same PSU or locality.

Analysts often want to use LSMS data for purposes other than deflating nominal consumption for each household, and to calculate some indicator of regional price levels, or of regional price levels at different times through the survey year. This can be done using either the Paasche indexes of this subsection, or the Laspeyres indexes discussed below. The most straightforward procedure is simply to take means (or better, medians) within the relevant region or season of the individual Paasche indexes as calculated above. Such indexes could be made more relevant to the poor by averaging the individual household price indexes only over those at or below the poverty line, see the next subsection for discussion of procedures. Note that when all households within a region R face the same prices, so that

$$\ln P_P^h = \sum w_k^h \ln\left(\frac{p_k^R}{p_k^0}\right) \tag{18}$$

the average of the (log) prices is given by

$$\overline{\ln P_P^R} = \sum \overline{w_k^R} \ln\left(\frac{p_k^R}{p_k^0}\right) \tag{19}$$

so that the appropriate weights for the average index are the means of the budget shares over all (or poor) households. Note that is not the same as using the weights defined as the share of aggregate purchases in aggregate total expenditure, weights that are typically used in computing consumer price indexes by statistical offices. These aggregate weights effectively weight each household, not on a 'democratic' basis, with one household or individual getting equal weight, but on a 'plutocratic' basis in which each household is weighted according to its total expenditure. Such indexes are biased towards the rich, and are especially inappropriate in poverty work.

3.2. Calculating Laspeyres Index

For researchers who wish to follow the welfare-ratio rather than money-metric approach to measuring living standards, the relevant price index is not the Paasche index (12), but the Laspeyres index (14). Because this index uses the same weights for all households, it is typically more straightforward to calculate than is the Paasche, although in both cases the hardest task is finding the price relatives, not calculating the weights. Once again, it is often useful to write the Laspeyres in terms of budget shares and price relatives so that, corresponding to equation (15), we now have

$$P_L^h = \frac{q^z \cdot p^h}{q^z \cdot p^0} = \Sigma w_k^{z^0}\left(\frac{p_k^h}{p_k^0}\right) \tag{20}$$

which corresponds to equation (14) or, alternatively, corresponding to equation (16),

$$u_m^h \approx \frac{p^h \cdot q^h}{P_P^h} = \frac{x^h}{P_P^h} \tag{21}$$

The discussion of measuring price relatives for foods and non-foods, and of aggregation over households goes through as before, although when we average the Laspeyres indexes, only the price relatives are being averaged, not the weights—the principle of averaging price indexes over households, however, remains unchanged.

The welfare ratio approach requires comparison of actual indifference curves with a baseline indifference curve, here taken to be the poverty-line indifference curve, and the theory requires that the weights for the Laspeyres index used for deflation be calculated at that indifference curve. In practice, it may not be obvious how to do this. There are usually many households near the poverty line, although rarely many (or even any) *exactly* at it, so we lack the data for the quantity or budget share weights in equations (20) and (21). A useful solution to this problem is to calculate weights by averaging over the expenditure patterns of households near the poverty line, with those closer to it given more weight than those further away. Weights with this property are conveniently provided by a 'kernel' function, here denoted $K_h(\cdot)$ and the weights in equations (14), (20) or (21) are calculated from

$$\tilde{w}_k^{z^0} = \sum_{h=1}^{K} K_\tau(x^h - z)w_k^h \tag{22}$$

This sum is a weighted average over *all* households in the sample of the budget shares w_k^h using the kernel weights. There are a number of suitable choices for the kernel function which must be positive, must sum to one over all households, and which must be smaller the larger is the absolute difference between x^h and the poverty line z. One convenient choice is the 'bi-square' function

$$K_\tau(x-z) = \frac{1}{\tau}\left(1-\left(\frac{x-z}{\tau}\right)^2\right)^2 \text{ for } \left|\frac{x-z}{\tau}\right| \leqslant 1 \tag{23}$$

and

$$K_\tau(x-z) = 0, \text{ otherwise} \tag{24}$$

The number τ is a 'bandwidth' that controls how many households are included in the sample. The larger is τ, the more households are used, which makes the average more precise, but can cause bias by including households a long way from the poverty line. In practice, setting τ so as to include a few hundred households around the poverty line will usually be satisfactory. These equations are also likely to work better if x^h and z in equations (22) to (24) are replaced by their logarithms, so that distances from the poverty line are measured proportionately, not absolutely.

4. Adjusting for Household Composition

Sections 2 and 3 have discussed how to use survey data to construct a nominal measure of total household consumption and of how to adjust it to take into account cost-of-living differences. However, as argued by Pyatt in his paper, we are ultimately interested in *individual* welfare, not the welfare of a household, something that is hard to define in any very useful way. If it were possible to gather data on consumption by individual family members, we could move directly from the data to individual welfare, but except for a few goods, such data are not available, even conceptually—think of public goods that are shared by all household members. As it is, the best that can be done is to adjust total household expenditure by some measure of the number of people in the household, and to assign the resulting welfare measure to each household member as an individual. Other determinants of individual welfare at the intra-household level, not elaborated upon here, are age and gender (see Deaton, 1997, pp. 223–241) for further discussion on these points. Equivalence scales are the deflators that are used to convert household real expenditures into money metric utility measures of individual welfare. If a household consists entirely of adults, and if they share nothing, each consuming individually, then the obvious equivalence scale would be household size, which is the number of people over which household expenditures are spread. Even when households consist of adults and children, welfare is often assessed by dividing expenditures by household size, as a rough-and-ready concession to differences in family size. However, such a correction does not allow for the fact that children typically consume less than adults, so that deflating by household size will understate the welfare of people who live in households with a high fraction of children.

Moreover, simply deflating household expenditures by total household size also means implicitly ignoring any economies of scale in consumption within the household. Some goods and services consumed by the household have a 'public goods' aspect to them, whereby consumption by any one member of the household does not necessarily reduce the amount available for consumption by another person within the same household. Housing is an important household public goods, at least up to some limit, as are durable items like televisions, or even bicycles or cars, which can be shared by several household members at different times. Because people can share some goods and services, the cost of being equally well-off does not rise in proportion to the number of the people in the household. Per capita measures of expenditure thus understate the welfare of big households relative to the living standards of small households.

In this section, I discuss equivalence scales in general and outline some of the main approaches to their calculation. Before doing so, however, it is worth emphasizing the importance of per capita expenditure. Twenty years ago, per capita expenditure was itself something of an innovation and many studies worked with total household expenditure or income without correction for household size. In the years since, deflation to a per capita basis has become the standard procedure, and although its deficiencies are widely understood, none of the alternatives discussed have been able to command universal assent.

4.1. Equivalence Scales

The costs of children relative to adults and the extent of economies of scale are of the first-order of importance for poverty and welfare calculations. Indeed, the direction of policy can sometimes depend on exactly how equivalence scales are defined. Larger households typically have lower per capita expenditure levels than small households but until we know the extent of economies of scale, we do not know which group is better off, or whether anti-poverty programs should be targeted to one or the other. Rural households are often larger than urban households, and we are sometimes unable to compare rural with urban poverty without an accurate estimate of the extent of economies of scale. Another frequent comparison is between children and the elderly, and both groups have claims for public attention on grounds of poverty. Children tend to live in larger households than do the elderly, and (obviously) live in households with a higher fraction of children. As a result, comparisons of welfare levels between the two groups are often sensitive to what is assumed about both child costs and about economies of scale, see the calculations in Section 4 below. Issues involving comparison between children and the elderly have acquired a new salience in work on the transition economies of Eastern Europe which, compared with developing countries of Africa or Asia, have relatively large elderly populations that receive state support through pensions and health subsidies. As a result, the two groups are in competition for welfare support, and an accurate assessment of their relative poverty has become an important issue.

Unfortunately, there are no generally accepted methods for calculating equivalence scales, either for the relative costs of children, or for economies of scale. There are three main approaches to deriving equivalence scales: (i) one relying on behavioural analysis to estimate equivalence scales; (ii) one using direct questions to obtain subjective estimates; and (iii) one that simply sets scales in some reasonable, but essentially arbitrary, way. Each of these is discussed in turn in the sections that follow.

4.2. Behavioural Approach

The behavioural approach has generated a large literature, much of which is reviewed in Deaton (1997). While there are methods for calculating the costs of children that are relatively soundly based—although not all would agree even with this—there are so far no satisfactory methods for estimating economies of scale. Many of the standard methods, such as Engel's procedures for calculating both child costs and economies of scale, are readily dismissed, see again Deaton (1997) and Deaton & Paxson (1998a). One idea that seems correct, and that can sometimes give a useful, if informal, notion of the extent of economies of scale, is that shared goods within the household, or household public goods, are the root cause of economies of scale. In the simplest case, there are two sorts of goods in the household, private goods, which are consumed by one person and one person only and where consumption by one person precludes consumption by another, and public goods, where there is an unlimited amount of sharing, and where consumption by one member of the household places no limitation on consumption by others. In this case, Drèze & Srinivasan (1997) have shown that, in a household with only adults, the elasticity of the cost-of-living with respect to household size is the share of private goods in total household consumption. If all goods are private, costs rise in proportion to the number of people in the household, while if all goods are public, costs are unaffected by the number of people. This sort of argument supports the intuitive notion that, in very poor economies with a high share of the budget devoted to food—which is almost entirely private—the scope for economies of scale is likely to be small. In other settings, where housing—which has a large public component—is important, economies of scale are likely to be larger. Unfortunately, attempts to extend this sensible approach to a more formal estimation of the extent of economies of scale have not been successful (Deaton & Paxson, 1998a).

4.3. Subjective Approach

The subjective approach to setting equivalence scales has attracted increased attention in recent years. One widely used technique is the 'Leyden' method pioneered by van Praag and his associates, see van Praag & Warnaar (1997) for a recent review. In the household survey, each household is asked to provide estimates of the amount of income it would need so that their circumstances could be described as 'very bad', 'bad', 'insufficient', 'sufficient', 'good', and 'very good'. Suppose that the answer to the 'good' question by household h is c^h. From the cross-section of results, c^h is regressed on household income and family size (or numbers of adults and children) in the logarithmic form

$$\ln c^h = \alpha + \beta \ln n^h + \gamma \ln y^h \tag{25}$$

This equation is used to calculate the level of income y^h that this household would have to have in order to name its actual income as 'good'. Evidently, this is given by

$$\ln \tilde{y}^h = \frac{\alpha}{1-\gamma} + \frac{\beta}{1-\gamma} \ln n^h \tag{26}$$

If $\tilde{y}^h$ is interpreted as a measure of needs in that it would be regarded by a household receiving it as 'good', then the quantity $\beta/(1-\gamma)$ can be interpreted as the elasticity of needs to household size, and thus (a negative) measure of economies of scale. Van Praag and Warnaar report an estimate of $\beta/(1-\gamma)$ for the Netherlands of 0.17, 0.50 for Poland, Greece and Portugal, and 0.33 for the US. Taken literally, these numbers indicate very large, not to say incredible, economies of scale.

Even if we accept the general methodology, it is hard to take these estimates seriously. In particular, if the costs of children, or more generally the costs of living together, vary from household to household, the estimation of equation (25) will lead to downward biased estimates of β. To see this, rewrite equation (25) including the error term as

$$\ln c^h = \alpha + \beta \ln n^h + \gamma \ln y^h + u^h \tag{27}$$

The term u^h varies from one household to the next, and represents the idiosyncratic costs of living for that household, the amount that household needs above the average for a household with its income and size. The trouble with this regression is that households choose their size n^h, partly through fertility, but more importantly by adults (and some children) moving in and out. People who like living with lots of other people will live in large households (high n^h) and will report that they need relatively little money to live in a large household (low u^h). As a result, the error term u^h will be negatively correlated with household size n^h and estimates of β will be biased downward, which is consistent with what van Praag and Warnaar report.

4.4. Arbitrary Approach

Given the current unreliability of either the behavioural or the subjective approach, there is much to be said for making relatively ad hoc corrections that are likely to do better than deflating by household size. One useful approach, detailed in National Research Council (1995), is to define the number of adult equivalents by the formula

$$AE = (A + \alpha K)^{\theta} \tag{28}$$

where A is the number of adults in the household, and K is the number of children. The parameter α is the cost of a child relative to that of an adult, and lies somewhere between 0 and 1. The other parameter, θ, which also lies between 0 and 1, controls the extent of economies of scale; since the elasticity of adult equivalents with respect to 'effective' size, $A + \alpha K$ is θ, $(1 - \theta)$ is a measure of economies of scale. When both α and θ are unity—the most extreme case with no discount for children or for size—the number of adult equivalents is simply household size, and deflation by household size is equivalent to deflating to a per capita basis. An alternative version of equation (28) is frequently used in Europe, whereby the first adult counts as one, and subsequent adults are discounted, so that the A in equation (28) is replaced by $1 + \beta(A - 1)$ for some β less than unity. This is really an alternative treatment of economies of scale so that, if this scheme is used, the parameter θ would normally be set to unity.

A case can be made for the proposition that current best practice is to use equation (28) for the number of adult equivalents, simply setting α and θ at sensible values. Most of the literature—as well as common sense—suggests that children are relatively more expensive in industrialized countries (school fees, entertainment, clothes, etc) and relatively cheap in poorer agricultural economies. Following this, α could be set near to unity for the US and western Europe, and perhaps as low as 0.3 for the poorest economies, numbers that are consistent with estimates based on Rothbarth's procedure for measuring child costs (Deaton & Muellbauer, 1986; Deaton, 1997). If we think of economies of scale as coming from the existence of shared public goods in the household, then θ will be high when most goods are private and low when a substantial fraction of household expenditure is on shared goods, see Section 4.3 above. Since households in the poorest economies spend as much as three-quarters of their budget on food, and since food is an essentially private good, economies of scale must be very limited, and θ should be set at or close to 1. In richer economies, θ would be lower, perhaps in the region of 0.75.

In Section 5 below, I argue that it is important to assess the robustness of poverty comparisons using stochastic dominance techniques, and I sketch out a simple methodology for doing so. When the results are not robust, for example when the comparison of poverty rates between children and the elderly is sensitive to the choice of α and θ within the sensible range for that country, there is probably not much alternative to facing failure squarely. Certainly, the behavioural approach is unlikely to provide estimates that would be sufficiently precise and sufficiently credible to support such fine distinctions. In such situations, it might be better to turn to other indications of well being, such as mortality or morbidity.

5. Methods of Sensitivity Analysis

Although the general procedures for calculating money metric utility are well defined in theory, in practice, compromises have to be made, and difficult choices have to be made between imperfect alternatives. Is it better to add in a poorly measured component of consumption—such as imputed rent, or a component that is lumpy and transitory—such as health expenditures—and sacrifice accuracy for an attempt at completeness? Decisions about equivalence scales are almost always controversial, and even if we use the formulas (28), how do we know that the results are robust to the choice of parameters that control child costs and economies of scale? Even with perfect estimates of money metric utility, poverty analysis is subject to its own inherent uncertainty associated with the difficulty of choosing a poverty line. Although there is much to be said for making the best decisions one can, picking a sensible poverty line, and pressing ahead, it is often informative to examine the sensitivity of key results to alternatives. In recent years, much use has been made of *stochastic dominance* analysis to examine the sensitivity of poverty measures to different poverty lines, and this work has led to a much closer integration between poverty measurement and welfare analysis more generally. Stochastic dominance techniques can also be useful in examining the sensitivity of poverty analyses to the way in which money metric utility is constructed, including the construction of equivalence scales. In this section, we explore some of these issues.

5.1. Stochastic Dominance

Suppose that we have a money metric utility measure which, for the moment and to reduce notational clutter, we denote by x. Suppose too that we are interested in the headcount ratio (HCR), the proportion of people whose money metric utility is below the poverty line z. If $F(.)$ is the cumulative density function of x in the population, $F(z)$ is the fraction below z, and thus is the HCR. The sensitivity of the HCR to changes in z, can be assessed simply by plotting the HCR as a function of z, i.e. by plotting the cdf $F(z)$ as a function of z. Suppose then that we have two measures of money metric utility, x_0 and x_1, corresponding to two different decisions about construction. Suppose that these decisions are such that it makes sense to use the same poverty line for both—this will be the case if both are unbiased for the true money metric utility, and neither is more precise than the other. We discuss what happens when this is not the case in the subsections below, although it is sometimes obvious how to adapt the poverty line in moving from one situation to the other. Then if the two cdfs are $F_1(.)$ and $F_2(.)$, the two HCRs are $F_1(z)$ and $F_2(z)$. Plotting both of these functions against z on a single graph shows which one gives the higher HCR, and how the difference in HCRs varies with the choice of the poverty line z.

Figure 2 illustrates the lower part of the cumulative distribution functions for two (imaginary) measures of welfare. If the horizontal axis is thought of as the poverty line, each line tells us the fraction of people in poverty corresponding to that poverty line. Putting the two graphs on the same figure tells us how robust

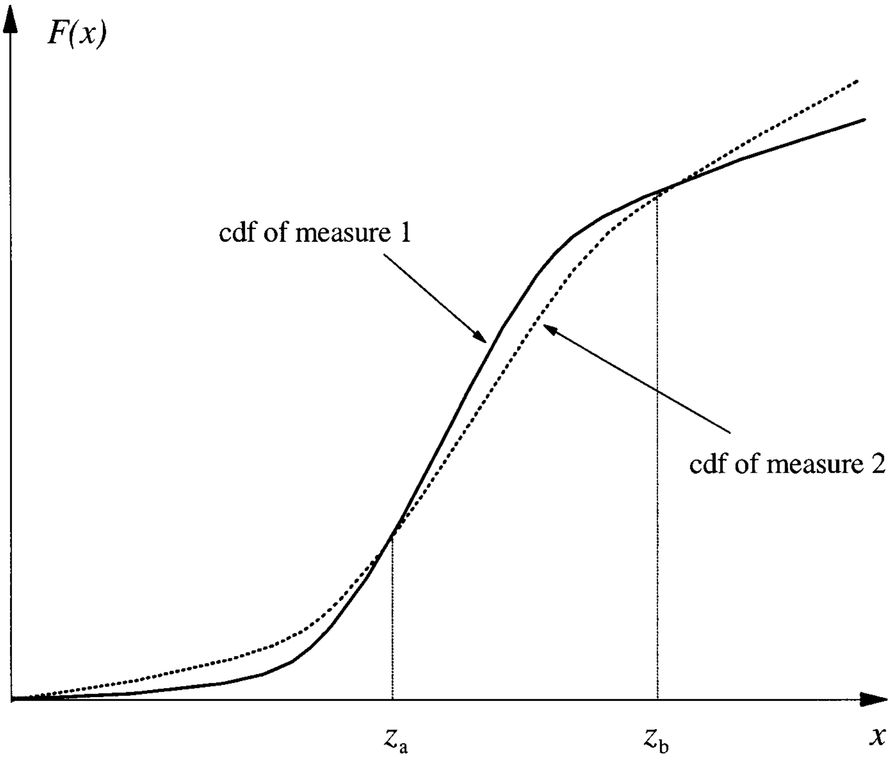

Figure 2. Cumulative distribution functions of two measures of welfare.

the head count ratio will be to the choice of measure at different poverty lines. For any low enough poverty line below z_a, the headcount ratio will be higher for measure 2. Between choice of poverty line between z_a, and z_b, measure 1 gives the higher poverty count, reversing again above z_b. Given some idea of the relevant poverty line, such figures tell us how the choice of measure affects the headcount.

This rather mechanical exercise becomes more interesting when we come to construct poverty profiles, for example for different groups, such as children and the elderly, or households in different regions. Suppose that we have two groups G and H, and that the conditional cdfs of the two measures are now $F_1(.|G)$ and $F_2(.|G)$ for G, with similar expressions for H. What we are typically concerned about is whether the relative poverty rates of G and H are sensitive to the choice between the two measures, and to what extent the conclusion depends on the choice of the poverty line. For poverty line z, and measure i, for i equal to 1 or 2, the difference in poverty rates between the two groups is

$$\Delta_i(z) = F_i(z|G) - F_i(z|H) \tag{29}$$

Plotting $\Delta_i(z)$ against z for a given i, and seeing whether it ever cuts the horizontal axis, tells us whether the poverty ranking of the two groups is sensitive to the choice of poverty line. Plotting the two Δ functions on the same graph tells us whether, at any given poverty line, the ranking is sensitive to the construction of the utility measure, and whether that sensitivity (or lack of it) depends on the choice of poverty line. An example of this kind of analysis is given in Section 5.3 below.

Sensitivity calculations for the headcount ratio involve the comparison of the cdfs of two distributions. Similar calculations are possible for other poverty measures; for example, the sensitivity of the poverty gap measure to the poverty line can be examined by plotting the areas under the cdfs, see Deaton (1997) for a review of the literature and for examples. These higher-order stochastic dominance comparisons can be used in the same way as above to examine the effects of construction on higher-order poverty measures.

5.3. Using Subsets of Consumption and the Effects of Measurement Error

It is often clear from the data collection exercise or from the subsequent analysis of the data that some components of consumers' expenditure are much better measured than others. Food is sometimes thought to be easier to measure than non-food, if only because, in households that eat from a common pot, there is a single well-informed individual who can act as respondent. Imputations are often quite suspect—for example, those for imputed rent for owner-occupiers in an economy where house tenancy is very rare. As a result, most analysts who have had to work through an LSMS survey, writing code to make the imputations, tend to be rather unwilling to make much use of the subsequent numbers. Whether it is better to use a subset of well-measured expenditures to assess poverty is an important question that has been raised by Lanjouw & Lanjouw (1997). As we have already seen, essentially the same issues arise in deciding whether or not to include an expenditure item where there are large, occasional expenditures. Transitory expenditure around a longer run mean is effectively the same as measurement error. In the rest of this subsection, I sketch out some results that are useful in thinking about measurement error and transitory expenditure. While I follow the lead of Lanjouw and Lanjouw, there are some differences in the analysis, both in methods and in results.

Before going on, it is worth noting that instrumental variable techniques for measurement error that are standard for making imputations, or for correcting regression analysis, are of more limited use when we are concerned with measuring poverty or inequality. The essential problem is that poverty and inequality depend on dispersion, not means, or even conditional means. If we are trying to estimate the mean expenditure of the population on some item, and some households have missing or implausible values, it is standard practice to impute an estimate, often from the mean of similar households, or more generally, from a regression using instruments, variables that are thought to be correlated with the missing information. But because such regressions only capture a fraction of the variation in the true variable, the fitted values will be less variable than the actuals, and imputation will tend to reduce inequality and poverty (if the poverty line is low enough.) Of course, for transitory expenditures and for measurement error, variance reduction is exactly what we want. But imputations are likely to eliminate not only the measurement error, but also the genuine variation across households, something that we need to preserve.

Start by assuming that there is a subset of total expenditure, such as food, expenditure on which is denoted by e, and that, conditional on total expenditure, x, we have

$$E(e|x) = m(x);\ V(e|x) = \sigma^2 \tag{30}$$

The regression function $m(x)$ can be thought of as an Engel curve, or as the true value of x when x is measured with error, or the long-run value of x when x has a large transitory component. The poverty line in terms of x is, as before, z, and the cdf of x is $F(.)$, so that the head count ratio is $F(z)$. Suppose that, instead of defining the poor in terms of low x, we define them in terms of low e; to do so, we must select an appropriate poverty line for e, and one obvious choice is to take the level of e on the Engel curve where total expenditure is equal to the poverty line, i.e. $m(z)$. The headcount ratio using e is then given by

$$P_e = F_e(m(z)) \tag{31}$$

where $F_e(.)$ is the cdf of e. If we assume that $m(x)$ is monotone, and therefore invertible, it can be shown that is P_e related to the 'true' headcount ratio P_x by the approximation

$$P_e \approx P_x + \frac{\sigma^2 f(z)}{[m'(z)]^2}\left(\frac{f'(z)}{f(z)} - \frac{m''(z)}{m'(z)}\right) \tag{32}$$

where $f(x)$ is the pdf of x. (This result is closely related to the results derived in a somewhat different context by Ravallion, 1988b.)

Note first that when the Engel curve fits perfectly (or there is no measurement error, or no transitory expenditure), so that $\sigma = 0$, the two poverty counts coincide, a result that is exact. Otherwise, the two poverty counts will diverge in a way that depends on the slope of the density of x at the poverty line, and on the convexity or concavity of the Engel curve. When the Engel curve is linear or when we are dealing with transitory expenditures or measurement error, the second term in brackets is zero, so that 'food' poverty will overstate 'true' poverty if $f'(z) > 0$ which will occur if the density of x is unimodal and the poverty line is below the mode. If this condition holds, the overstatement will be exacerbated if the Engel curve is concave, and moderated if it is convex.

These results are a useful starting point, but are not directly practical. If we knew both x and its component e, there would be no need to use the latter. Nevertheless, there are two immediate corollaries that are more useful. The first is the case where $m(x) = x$, so that e is just an error-ridden measure of x, so that equation (32) becomes

$$P_e \approx P_x + \sigma^2 f'(z) \tag{33}$$

which gives us a guide about how measurement error inflates (or deflates) the poverty measure. This formula is particularly useful when we have some idea of the variance of the measurement error which, for example, could be estimated from two error-ridden but independent measures of x. Note also that equation (33) is the basis for the result that, for unimodal distributions, where $f'(x)$ is first positive and then negative, adding measurement error increases the headcount ratio if the poverty line is below the mode, so that $f'(z) > 0$, and decreases it when the poverty line is above the mode, where $f'(z) < 0$. Except in the very poorest areas, we would expect the poverty line to be below the mode.

The approximation formula is also useful when considering whether or not to include a poorly measured component in the total. To simplify, suppose that e is the non-controversial component of the total x, so that adding in the controversial component would, in principle, take us to the total x. Suppose that the Engel curve for e is linear, so that the derivative $m'(x)$ is constant, equal to β say. To avoid confusion, rewrite its variance around the regression line as σ_e^2, where the subscript

e identifies the non-controversial component. From equation (32), the poverty count using the comprehensive, but noisy measure is

$$P_{\hat{c}} \approx P_{\tilde{c}} + \sigma_x^2 f'(z) \tag{34}$$

where σ_x^2 is the measurement error in the comprehensive (but noisy) total; c is for comprehensive. From equation (32), the poverty count using the non-controversial component alone is

$$P_e \approx P_x + \frac{\sigma_e^2 f'(z)}{\beta^2} \tag{35}$$

Since it is normally the case that the poverty line is below the mode, we can assume that $f'(z)$ is positive, in which case the poverty count based on the comprehensive but noisy measure will be closer to the truth if

$$\beta < \frac{\sigma_e}{\sigma_c} \tag{36}$$

Note that β is the share of the marginal rupee devoted to the non-controversial good, and that $1-\beta$ is the share going to the controversial good, so that the case for inclusion of the controversial item is strong if, at the margin, a large share of total expenditure is devoted to it, while the case is weaker the larger is the ratio of variance in the comprehensive measure to the non-controversial measure. This result is perhaps not surprising. A strong link to total expenditure is a case for inclusion, while making the total noisier is a case against inclusion. Note finally that equation (36) can be written in terms of the total-expenditure elasticity of the non-controversial component, ε_e and the *relative* measurement errors as:

$$\varepsilon_e < \frac{\sigma_e/e}{\sigma_c/x} \tag{37}$$

Since the (weighted) sum of the controversial and non-controversial elasticities is unity, equation (37) is a prescription for including controversial items if their total expenditure elasticities are large, provided they do not add too much measurement error. Of course, neither σ_e nor σ_c can actually be observed in practice, but the formulas (36) and (37) tell us what to look for and what to think about when making the decision to trade off comprehensiveness versus precision.

5.4. *Sensitivity Analysis with Equivalence Scales*

Suppose that we are working with the formula (28) that links adult equivalents to the number of adults A and the number of children K according to

$$AE = (A + \alpha K)^\theta \tag{38}$$

and that we do not know α or θ, although we may be prepared to commit to a range of values for each. Given values for the two parameters, we can compute money metric utility values for everyone so that, armed with a poverty line, we can calculate poverty rates for any groups. In this context, groups that we are particularly likely to be interested in are children, adults and the elderly, as well as other groups where households have different sizes and compositions, such as rural versus urban households. Sensitivity analysis to different values of α, θ and z, proceeds in very much the same way as discussed in Section 5.1.

However, as in Section 5.2 but in contrast to Section 5.1, we cannot simply change the parameters and leave the poverty rate unchanged. For example, suppose that α is set at 1, and θ is reduced from 1 to 0.5. As a result, EA would be reduced for all households except those with only a single person, so that, if the poverty line were held constant, poverty would be decreased. But this is not what we want changes in the parameters of the equivalence scale to do. Instead, we want to alter the relative standings of large households relative to small households, or households with large numbers of children relative to those with none. A straightforward way to do this is to select a particular household type as 'pivot', and to choose the equivalence scale in such a way that the money metric utility of people in such households are unaffected by changes in the parameters. Denote the number of adults and children in the reference or pivot household by (A_0, K_0); in practice, this should be chosen as the modal type, for example, a two-adult and three-child household. We then define money metric utility, not as x divided by AE, but as

$$x^* = \frac{x}{(A + \alpha K)^\theta} \frac{(A_0 + \alpha K_0)^\theta}{A_0 + K_0} \tag{39}$$

At any given value of α and θ, x^* is just a scaled version of x/AE; but for the reference household, x^* is always equal to per capita expenditure, and is unaffected by changes in α and θ.

An alternative procedure, not pursued here but equally useful in practice, is to alter the poverty line for use with equivalent expenditure so as to hold constant the measure of interest, for example the headcount ratio. This is most simply done by trial and error. Calculate per equivalent expenditures for each household by dividing total expenditure by equivalent adults calculated using the chosen values of α and θ. For a trial poverty line, calculate the head count ratio, and continue adjusting until the head count ratio returns to its value using per capita expenditure. Equivalently, the ratio of the new to the old poverty lines can be used to deflate expenditure per equivalent, at which point the original poverty line can be used.

Figures 3–5, reproduced from Deaton & Paxson (1998b), show what happens to the relative poverty of children, non-elderly adults, and the elderly in South Africa using the 1993 South African LSMS. These calculations are done on an individual basis, whereby when money metric utility is assigned to a household, it is assigned to each person in that household, or more precisely to every person in the number of people in the population calculated by multiplying the number of people in the household by its survey inflation factor. Figure 3 shows the cdfs for the three groups, for a range of possible poverty lines, and for nine combinations of values for α and θ. Irrespective of the values chosen, and irrespective of the poverty line, non-elderly adults always have a lower headcount ratio than do children or the elderly. The poverty profile of the elderly versus that of children depends on the values of the parameters. In the top right of the figure, where children are cheap, and economies of scale are large, children do better than the elderly, who benefit relatively little from either economies of scale or inexpensive children. At the bottom left of the picture, where there are no discounts for children or for large size, so that money metric utility is expenditure per capita, the children are more likely to be poor than the elderly at all poverty lines.

Figures 4 and 5 show plots of the difference between the cdf for the elderly and the cdf for children for the same range of the poverty line, but with plots for different values of α and θ on the same graph. By discarding the automatic increase in the cdf with the level of the poverty line, and looking only at differences, these

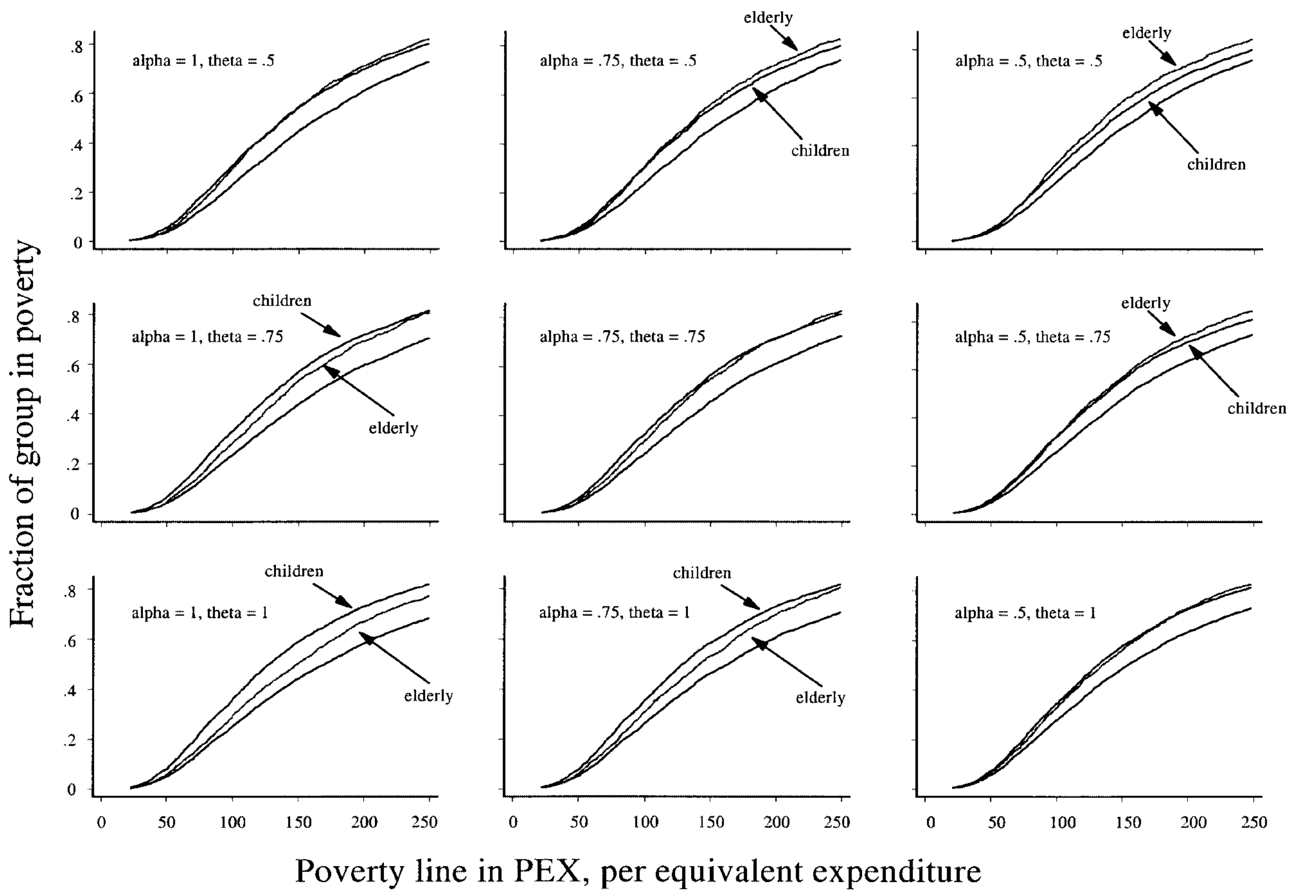

Figure 3. South Africa, poverty headcount ratios at various poverty lines and for various child costs and economies of scale.

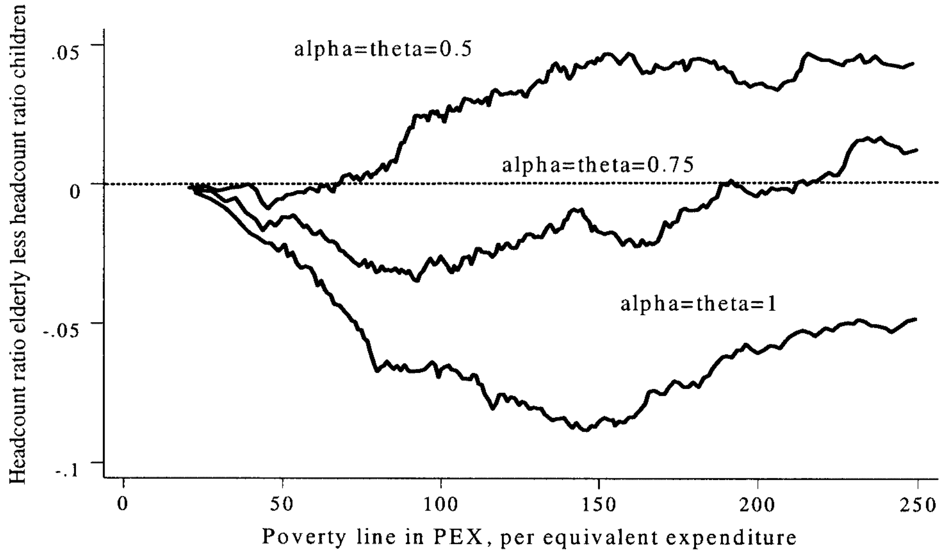

Figure 4. South Africa: poverty rates of the elderly and children.

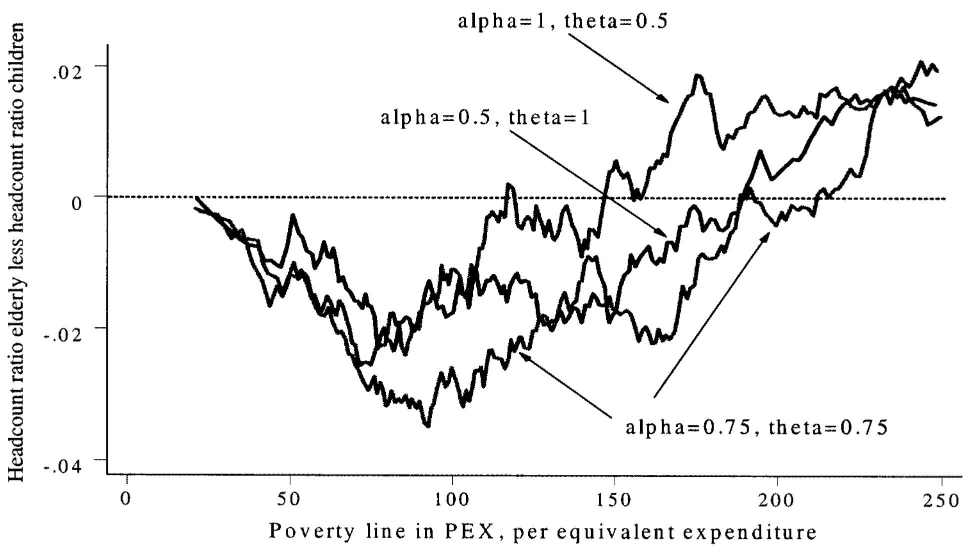

Figure 5. South Africa: poverty rates of the elderly and children.

graphs permit greater focus on the differences of interest, here the elderly versus children. The upper graph in Figure 4 shows the movement on Figure 3 from top right to bottom left, and shows how children become relatively poorer, and that, in the middle configuration, with $\alpha = \theta = 0.75$, the relative poverty rates depend on the value of the poverty line. The lower graph in Figure 4 shows the progress through Figure 3 from top left to bottom right, and shows a more muddied picture. All three graphs show that the relative poverty rates of the two groups depend on the poverty line, with children tending to be less poor at higher values.

What should we conclude from sensitivity analyses like these? Much of the time, the desired result from a sensitivity analysis is to find that the results are robust, so that clear conclusions can be drawn. This will sometimes be the case, but rarely for the analysis of equivalence scales, where we know from a large body of work that some important issues are not robust. Indeed, Deaton and Paxson show similar sensitivities between the relative poverty rates of children and the elderly, not only for South Africa, but also for Ghana, Pakistan, Taiwan and Thailand, but not Ukraine. In the absence of a breakthrough in behavioural and or subjective methods of measuring equivalence scales, it may simply be necessary for policy to be conducted in ignorance of the relative poverty of some groups.

6. Conclusions

In this paper I have reviewed four aspects of survey based poverty analysis where views have changed over the 20 years since the LSMS project was began. All of these are important, but lack of space ruled out the treatment of other topics that might well have been included. Perhaps the most important among these is the issue of questionnaire design. In spite of half a century of work, we still do not know what is the best reference period for asking questions about consumption. Most LSMS surveys use a two-week recall period, but other countries use periods as short as a few days, and some as long as a month, even for high frequency items such as food. In recent years, the Indian National Sample Survey Organization has experimented with 7-day and 30-day recall periods for a number of items, and discovered that the 7-day recall cuts Indian poverty rates by half, removing some 200 million people from dollar-a-day poverty. This is an issue that has not been much investigated by LSMS surveys, but has recently moved into the forefront of research. Clearly, the programme that Graham Pyatt began in 1979 still has many important topics to research.

References

Blackorby, C. & Donaldson, D. (1987) Welfare ratios and distributionally sensitive cost-benefit analysis, *Journal of Public Economics*, 34, pp. 265–290.

Blackorby, C. & Donaldson, D. (1988) Money metric utility: a harmless normalization? *Journal of Economic Theory*, 46, pp. 120–129.

Deaton, A. S. (1980) The measurement of welfare: theory and practical guidelines, *LSMS Working Paper*, No. 7 (Washington, DC, The World Bank).

Deaton, A. S. (1997) *The Analysis of Household Surveys: Microeconometric analysis for development policy* (Baltimore, MD, Johns Hopkins University Press).

Deaton, A. S. & Grosh, M. (2000) Consumption, in: M. Grosh & P. Glewwe (eds) *Designing Household Survey Questionnaires for Developing Countries: Lessons from Ten Years of LSMS Experience*, Vol. 1 (Washington, DC, The World Bank), pp. 91–133.

Deaton, A. & Muellbauer, J. (1980) *Economics and Consumer Behavior* (New York, Cambridge University Press).

Deaton, A. S. & Muellbauer, J. (1986) On measuring child costs: with applications to poor countries, *Journal of Political Economy*, 94, pp. 720–744.

Deaton, A. S. & Paxson, C. H. (1998a) Economies of scale, household size, and the demand for food, *Journal of Political Economy*, 106, pp. 897–930.

Deaton, A .S. & Paxson, C. H. (1998b) Poverty among children and the elderly in developing countries. Research Program in Development Studies, Princeton University, processed.

Deaton, A. S. & Tarozzi, A. (2000) Prices and poverty in India, Research Program in Development Studies (Princeton NJ, Princeton University). Available at http://www.wws.princeton.edu/~deaton

Drèze, J. & Srinivasan, P. V. (1997) Widowhood and poverty in rural India: some inferences from household survey data, *Journal of Development Economics*, 54, pp. 217–234.

Lanjouw, J. O. & Lanjouw, P. (1997) Poverty comparisons with non-compatible data: theory and illustrations, *Policy Research Working Paper* (Washington DC, The World Bank).

National Research Council (1995) *Measuring Poverty: a New Approach* (Washington, DC, National Academy Press).

Ravallion, M. (1988a) Expected poverty under risk-induced welfare variability, *Economic Journal*, 98, pp. 1171–1182.

Ravallion, M. (1998b) Poverty lines in theory and practice, *LSMS Working Paper* 133 (Washington, DC, The World Bank).

Samuelson, P. A. (1974) Complementarity—An essay on the 40th anniversary of the Hicks-Allen revolution in demand theory, *Journal of Economic Literature*, 15, pp. 24–55.

van Praag, B. M. S. & Warnaar, M. F. (1997) The cost of children and the use of demographic variables in consumer demand, in: M. Rosenzweig & O. Stark (eds) *Handbook of Population and Family Economics*, vol. 1A (Amsterdam, North-Holland), pp. 241–273.