

gued that poverty lines should move with the general standard of living, although perhaps not at the same rate. Some would argue that poverty is a purely relative phenomenon, defined by current social customs, and that the poor are simply those in the bottom percentiles of the distribution of welfare.

An intermediate view comes from Sen's (1985, 1992) view of welfare in terms of the capability to function in society. If economic growth means that food is sold with an increased amount of packaging and service built in, if city center stores relocate to suburban areas that cannot be reached on foot, and if urban growth increases the cost and time to travel to work, then a fixed absolute poverty line makes no sense. There is also some relevant empirical evidence that comes from asking people whether they are poor and what the poverty line ought to be (see Mangahas 1979, 1982, 1985, who makes good use of such surveys to assess poverty in the Philippines). In the United States, Gallup polls have regularly asked respondents how much money they would need "to get along," and more occasionally what they think would be an adequate poverty line. In the 1960s, the mean responses about the latter were close to the official (Orshansky) line, but have since increased in real terms, although not always as fast as has average real disposable income (see Rainwater 1974 and Vaughan 1992). Ravallion (1993) has also examined the cross-country relationship between real gross domestic product (GDP) and poverty lines, and found that the elasticity is close to unity. While many people—including this author—are uncomfortable with an entirely relative concept of poverty, it is surely right that there should be some movement of the line in response to changes in mean levels of living.

The conceptual and practical difficulties over the choice of a poverty line mean that all measures of poverty should be treated with skepticism. For policy evaluation, the social welfare function is all that is required to measure welfare, including an appropriate treatment of poverty. While it is possible—and in my view desirable—to give greater weight to the needs of the poorest, I see few advantages in trying to set a sharp line, below which people count and above which they do not. Poverty lines and poverty counts make good headlines, and are an inevitable part of the policy debate, but they should not be used in policy evaluation. Perhaps the best poverty line is an infinite one; everyone is poor, but some a good deal more so than others, and the poorer they are the greater weight they should get in measuring welfare and in policy evaluation.

The concept of a poverty line is deeply embedded in the poverty literature, and measures of poverty are typically based on it. Even so, a good deal of the recent literature on poverty has followed Atkinson (1987) in recognizing that the poverty line is unlikely to be very precisely measured, and trying to explore situations in which poverty measures are robust to this uncertainty. I shall return to this approach below once I have introduced some of the standard measures.

### Measures of poverty

There are a number of good reviews of alternative poverty measures and their properties, see in particular Foster (1984) and Ravallion (1993), so that I can

confine myself here to a brief discussion of the most important measures. The obvious starting point—and the measure most often quoted—is the *headcount ratio*, defined as the fraction of the population below the poverty line. If the line is denoted by  $z$ , and the welfare measure is  $x$ , then the headcount ratio is

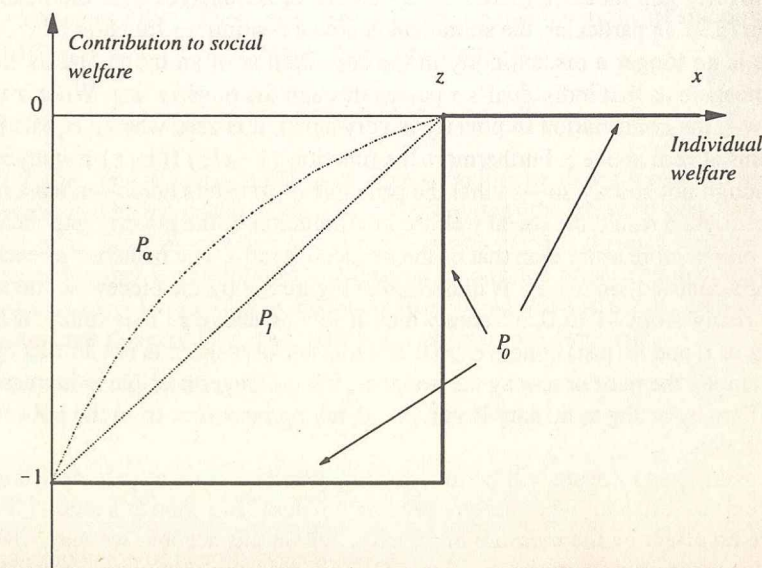
$$(3.9) \quad P_0 = \frac{1}{N} \sum_{i=1}^N 1(x_i \leq z)$$

where  $1(\cdot)$  is an indicator function that is 1 if its argument is true and 0 otherwise. The sum of the indicators on the right-hand side of (3.9) is the number of people in poverty, so that  $P_0$  is simply the fraction of people in poverty.

It is worth noting that with a change of sign, (3.9) could conceivably be regarded as a social welfare function. It is the average value of a rather strange valuation function in which  $x$  counts as  $-1$  when it is below the poverty line  $z$ , and as 0 when it is above  $z$ . This function is illustrated as the heavy line labeled  $P_0$  in Figure 3.2; it is nondecreasing in  $x$ , so it has some of the characteristics of a utility function, but its discontinuity at the poverty line means that it is not concave. It is this lack of concavity that violates the principle of transfers, and makes it possible to increase social welfare by taking money from the very poor to lift some better-off poor out of poverty.

Even if the poverty line were correctly set, and even if it were acceptable to view poverty as a discrete state, the headcount ratio would be at best a limited measure of poverty. In particular, it takes no account of the degree of poverty, and would, for example, be unaffected by a policy that made the poor even poorer. The

Figure 3.2. Alternative poverty measures and social welfare





headcount ratio gives the same measure of poverty whether all the poor are just below a generous poverty line, or whether they are just above an ungenerous level of subsistence. One way of doing better is to use the *poverty gap* measure

$$(3.10) \quad P_1 = \frac{1}{N} \sum_{i=1}^N \left(1 - \frac{x_i}{z}\right) 1(x_i \leq z).$$

According to (3.10), the contribution of individual  $i$  to aggregate poverty is larger the poorer is  $i$ .  $P_1$  can also be interpreted as a per capita measure of the total shortfall of individual welfare levels below the poverty line; it is the sum of all the shortfalls divided by the population and expressed as a ratio of the poverty line itself. Hence if, for example,  $P_1$  were 0.25, the total amount that the poor are below the poverty line is equal to the population multiplied by a quarter of the poverty line.

It is tempting to think of  $P_1$  (or at least  $P_1 z$ ) as a measure of the per capita "cost" of eliminating poverty, but this is far from being so except in the impractical case where lump-sum taxes and subsidies are possible. Even when tax and subsidy administration is efficient and is not corrupt, redistributive taxes have incentive effects that may render the elimination of poverty neither possible nor desirable given the actual range of feasible policies. This is clearly the case in an economy where everyone is poor, but applies much more widely. Once again, the appropriate way to think about tax systems for poverty alleviation is to go back to the social welfare function (3.1), to make sure that it incorporates the appropriate degree of weighting towards the poor, and to apply the general theory of tax design (see Newbery and Stern 1987 for a general discussion of such problems in the contexts of developing countries, and Chapter 5 below for some of the empirical issues).

The poverty gap measure (3.10) has a number of advantages over the headcount ratio (3.9). In particular, the summand is now a continuous function of  $x$ , so that there is no longer a discontinuity in the contribution of an individual to the poverty measure as that individual's  $x$  passes through the poverty line. When  $x$  is just below  $z$ , the contribution to poverty is very small, it is zero when  $x$  equals  $z$ , and remains at zero above  $z$ . Furthermore the function  $(1 - x/z) 1(x \leq z)$  is convex in  $x$ —although not strictly so—so that the principle of transfers holds—at least in a weak form. As a result, the social welfare interpretation of the poverty gap measure also makes more sense than that of the headcount ratio. The behavior of each individual's contribution to  $-P_1$  is illustrated in Figure 3.2 by the piecewise linear function rising from  $-1$  to  $0$ , a value which it retains above  $z$ . This function is increasing in  $x$ , and is (just) concave, so that while social welfare is not altered by transfers among the poor or among the nonpoor, it is no longer possible to increase social welfare by acting as an anti-Robin Hood, taking resources from the poor to give to the rich.

The poverty gap measure will be increased by transfers from poor to nonpoor, or from poor to less poor who thereby become nonpoor. But transfers among the poor have no effect on the measure of poverty, and on this account we may wish to consider other poverty measures. Sen's (1976b) measure of poverty remedies

the defect by incorporating the inequality among the poor. The definition is

$$(3.11) \quad P_S = P_0 \left(1 - (1 - \gamma^p) \frac{\mu^p}{z}\right)$$

where  $\mu^p$  is the mean of  $x$  among the poor, and  $\gamma^p$  is the Gini coefficient of inequality among the poor, calculated by treating the poor as the whole population. Note that when there is no inequality among the poor,  $P_S$  reduces to the poverty gap measure  $P_1$ . Conversely, when all but one of the poor has nothing,  $P_S = P_0$  and the Sen measure coincides with the headcount ratio. More generally, the Sen measure is the average of the headcount and poverty gap measures weighted by the Gini coefficient of the poor,

$$(3.12) \quad P_S = P_0 \gamma^p + P_1 (1 - \gamma^p).$$

Because Sen's measure depends on the Gini coefficient, it shares two of its inconveniences. First, the Gini—and thus the Sen index—is not differentiable. Although there is no economic reason to require differentiability, the inability to differentiate is sometimes a nuisance. More seriously, Sen's measure cannot be used to decompose poverty into contributions from different subgroups, something that is often informative when monitoring changes in poverty. If the aggregate poverty measure can be written as a weighted average of the poverty measures for the rural and urban sectors, or for households by age, or by occupation of the head, then changes over time can be similarly decomposed thus helping to identify groups that are particularly at risk, as well as sometimes pointing to the underlying mechanisms. While decomposability is hardly as fundamental a property as (say) the principle of transfers, it is extremely useful.

Our final poverty measure, or set of measures, comes from Foster, Greer, and Thorbecke (1984). Their measures are direct generalizations of the poverty gap (3.10) and are written, for some positive parameter  $\alpha$ ,

$$(3.13) \quad P_\alpha = N^{-1} \sum_{i=1}^N (1 - x_i/z)^\alpha 1(x_i \leq z)$$

so that  $P_0$  and  $P_1$  are special cases corresponding to values for  $\alpha$  of 0 and 1, respectively. The larger the value of  $\alpha$ , the more does the measure penalize the poverty gaps. Most frequently used is  $\alpha = 2$ , which yields a poverty measure like the Sen index that is sensitive to distribution among the poor. The decomposability property of (3.13) follows immediately from its additive structure. In particular, if sectors are denoted by  $s$ , and there are  $S$  of them, we can write

$$(3.14) \quad P_\alpha = N^{-1} \sum_{s=1}^S \sum_{j \in s} (1 - x_j/z)^\alpha 1(x_j \leq z) = \sum_s (n_s/N) P_\alpha^s$$

where  $n_s$  is the number of people in sector  $s$  and  $P_\alpha^s$  is the Foster, Greer, and Thorbecke index for poverty within the sector. Using (3.14), changes in aggregate poverty can be assigned to changes in sectoral poverty measures or to changes in the proportion of people in each sector.



Figure 3.2 shows how  $P_\alpha$  measures up as a social welfare function. The third line, shown as dots and dashes, is the contribution of individual welfare to social welfare defined as  $-P_\alpha$  in the case where  $\alpha$  is greater than unity; as with  $P_1$ , this line continues along the  $x$ -axis to the right of  $z$ . Because the function is strictly concave below  $z$ , it is sensitive to the degree of inequality among the poor. It is also continuously differentiable, even at the poverty line, so that the implicit marginal social utility declines continuously from its maximum among the poorest to zero at and above the poverty line.

### *The choice of the individual welfare measure*

Apart from a brief reference in the context of choosing a poverty line, I have so far avoided discussion of exactly how welfare is to be measured, and what practical concept should replace the  $x$ 's in the various poverty and inequality formulas. Ideally, we should like a survey based measure that approaches as closely as possible the individual welfare measures of economic theory. Particularly useful here is the concept of money metric utility—see Deaton and Muellbauer (1980a, ch. 7) for an overview—whereby the indifference curves of individual preference orderings are labeled by the amount of money needed to reach them at some fixed set of prices. In order to avoid the specification of a parametric utility function, money metric utility can be approximated by real income or real expenditure, the two leading candidates for practical welfare measures. However, there are other possibilities, indicators of nutritional status being perhaps the most important, and even if we settle on income or expenditures, there are many other questions that have to be settled before going on to compute the measures. In this subsection, I discuss a few of the most important: the choice between consumption and income or other concepts, the choice between individual and household measures, the choice of time period, as well as some data issues, particularly the effects of measurement error and reporting periods.

In the context of measuring welfare in developing countries, there is a very strong case in favor of using measures based on consumption not income. The standard argument—that by the permanent income hypothesis, consumption is a better measure of lifetime welfare than is current income—is much weaker than arguments based on practicality and data. It is unwise to condition a welfare measure on the validity of a hypothesis whose empirical support is at best mixed. In particular and as we shall see in Chapter 6, there is very little evidence from developing countries—or anywhere else—that lifetime consumption profiles are detached from lifetime income profiles as is required if consumption is to be superior to income as an indicator of lifetime welfare. Of course, there is no doubt that households smooth their consumption over shorter periods, certainly days, months, and seasons, and to some extent over runs of years. Income, especially agricultural income, can be extremely variable, and a farmer's income in any month is a poor indicator of living standards in that month. A better case can be made for annual income, but if farmers can even partially smooth out good years and bad, consumption will be the better measure.

At the practical level, and as discussed in Section 1.2, the difficulties of measuring income are much more severe than those of measuring consumption, especially for rural households whose income comes largely from self-employment in agriculture. Given also that annual income is required for a satisfactory estimate of living standards, an income-based measure requires multiple visits or the use of recall data, whereas a consumption measure can rely on consumption over the previous few weeks. Note that these arguments are likely reversed if we were dealing with, for example, the United States, where individual consumption surveys are much less developed than income surveys, where a much smaller fraction of the population is self-employed, where seasonality is much less of an issue, and where it is both feasible and economical to obtain accurate estimates of income for most people.

The conversion of nominal measures of consumption to real measures requires a price index. In most cases, there will exist an adequate consumer price index or cost-of-living estimate that can be used to compare data collected in different time periods. In countries with rapid inflation, this may even have to be done within each survey year, since different households are interviewed at different times. What is often more difficult is the comparison of living costs across regions at a given time, for example, when we are trying to compare living standards or poverty rates across different regions. In some surveys—but not typically in the Living Standards Surveys—households are asked to report both quantities and expenditures on a range of goods, particularly foods, and these data can be used to calculate unit values. Although unit values are not the same as prices—an issue that will be discussed in some detail in Chapter 5—accurate price indexes for each region can nevertheless be obtained from the unit values by averaging within regions and calculating a Laspeyres index for each, that is by pricing out a fixed bundle of goods at the average unit values for each region. The Living Standards Surveys have usually collected price data, not from households, but from observations on prices in markets used by the households in the survey, and these data can also be used to construct regional price indexes.

Although consumption and income are the standard measures of economic welfare, we will often want to supplement them with other measures of well-being, such as nutritional and health status, life expectancy, and education. While it is possible to consider methods for combining these indicators into a single measure, there is no adequate theory underlying such an aggregate so that weighting schemes are inevitably arbitrary, and it is more informative—as well as honest—to keep the different indicators separate. This is not to downplay the importance of these other indicators, nor to deny that public goods such as hospitals and schools contribute an important part of individual welfare. However, it is important not to confuse the components of economic welfare with their aggregate. We have already seen how the definition of a poverty line in terms of calories can give misleading results when relative prices differ. The same argument applies to attempts to shortcut welfare measurement using indicators such as housing, or the ownership of durable goods. Immigrants to big cities often live in very poor-quality housing in order to have access to employment. In such cases, their poor housing



reflects the high price of housing in urban areas, but may tell us little about their living standards.

Because surveys collect data at the level of the household, and not the individual, poverty and welfare measures must be based on consumption and income totals for the household, not for the individual. Although some surveys collect data on individual earnings, and even on individual income from assets, there is typically a component of household income—a large component in the case of family farms—that is not readily attributable to individual household members. For consumption, the position is even worse. Data on purchases are inevitably purchases for the household as a whole, and although some items—such as food—are conceptually assignable to particular individuals, the cost of observing who eats what is too large for all but specialist nutritional surveys. Even then there are questions about contamination of behavior by the presence of the enumerator during food preparation and family meals. There are also public goods in most households—goods and services the consumption of which by one member of the household does not exclude consumption by others. The consumption of these goods cannot be assigned to specific individuals.

As a result, we can either treat households as the units whose welfare is being measured, or we can use some rule to divide household total expenditure between its members, usually equally or in proportion to some measure of needs, and then treat each individual as the unit in the poverty and welfare calculations. Since it is hard to think of households as repositories for well-being, even in the best case where their membership does not change, an individual basis for measurement is conceptually clearer and is the recommendation carried throughout this book. One difficulty is that the assumption of equal division, or of each according to her or his needs, is bound to understate the true dispersion of consumption among individuals, and thus understate inequality and poverty. As pointed out by Haddad and Kanbur (1990), who have also investigated the magnitude of the biases, the assumed equal distribution within the household could be reached from the unknown true one by a system of equalizing transfers, so that any welfare measure that respects the principle of transfers will be overstated (or understated if a poverty measure) using household data.

It is also necessary to recognize that children do not have the same needs as adults. Assigning household PCE to each person gives too little to adults—especially those who do heavy manual work—and too much to children. If there are economies of scale, PCE will understate individual welfare levels, even if all household members are adults. Attempts to do better than PCE measures for individuals are discussed in Chapter 4, where I take up the question of allocation within the household, and the construction of “equivalence scales,” numbers that are the theoretically appropriate deflators to move from household to individual welfare. However, I should point out in advance that the equivalence scale literature is still very far from providing satisfactory answers to these questions, and that the use of household PCE assigned to individuals is still best practice. Even so, it is wise to remain skeptical of estimates that appear to be purely statistical but rely heavily on arbitrary assignments, such as the number of children in poverty, or the average

living standards of the elderly. The elderly rarely live by themselves in poor countries, and children do not do so anywhere, so that estimates of their welfare are determined as much by assumption as by measurement and should be treated as such. Measures of the fraction of children in poverty, or of women in poverty, are particularly fragile and international comparisons of such concepts cannot be treated seriously.

The choice of time period, like all of these issues, is partly one of theory and partly one of practicality. In theory, we need to decide the reference period for welfare measurement, whether someone is poor if they go without adequate consumption or income for a week, a month, or a year. The reference period can be shorter for consumption than for income, and if we use income, the choice of reference period will depend on what mechanisms—credit markets, familial support—are available to help people ride out fluctuations in income. In practice, long reference periods require either multiple visits or recall questions; the former are expensive and the latter risk measurement error. Note also that, if poverty and welfare measures are to be comparable across countries or over time, the reference periods must be the same. Because the dispersion of both consumption and income decrease the longer is the reference period, both the extent of inequality and poverty will be larger at short than at long reference periods.

One of the most difficult practical issues in estimating poverty and inequality is to separate genuine dispersion from measurement error. If we start from any distribution of welfare and add measurement errors that have zero mean and are uncorrelated with the true values, the new distribution is a spread-out version of the original one, so that if our measures respect the principle of transfers, measured inequality will be higher and social welfare lower. Poverty measures that satisfy the principle of transfers will also be higher. For the headcount, which does not satisfy the principle, matters are more complicated and measurement error can bias the count in either direction. If the country is wealthy enough for the poverty line to be below the mode, the addition of mean-zero measurement error will cause the measured headcount to overstate the number in poverty, and vice versa. Similarly if we try to assess the persistence of poverty using panel data by seeing who remains in poverty and who escapes it, measurement error will exaggerate the extent of mobility, and make poverty seem less persistent than is truly the case. In most cases, we have little idea of the magnitude of measurement error, or how much of the variance of consumption or income changes is noise as opposed to signal. However, and bearing in mind the problems of estimating consumption and income in surveys in developing countries, it is always wise to consider the robustness of conclusions to the presence of substantial measurement error.

### *Example 1. Inequality and poverty over time in Côte d'Ivoire*

This subsection applies some of the foregoing concepts to Living Standards data from Côte d'Ivoire for the four years 1985 through 1988, focussing on change over time, while the next subsection uses data from South Africa in 1993 to look at differences by race. The translation of the formulas into numbers is essentially