Chapter 2. Measuring Poverty

Summary

The first step in measuring poverty is defining an indicator of welfare such as income or consumption per capita. Information on welfare is derived from survey data. Good survey design is important. Although some surveys use simple random sampling, most use stratified random sampling. This requires the use of sampling weights in the subsequent analysis. Multistage cluster sampling is also standard; it is cost-effective and unbiased, but lowers the precision of the results, and this calls for some adjustments when analyzing the data.

The World Bank-inspired Living Standards Measurement Surveys (LSMS) feature multi-topic questionnaires and strict quality control. The flexible LSMS template is widely used.

Income, defined in principle as *consumption + change in net worth*, is generally used as a measure of welfare in developed countries, but tends to be seriously understated in less-developed countries. Consumption is less understated and comes closer to measuring "permanent income." However, it requires one to value durable goods (by assessing the implicit rental cost) and housing (by estimating what it would have cost to rent).

While consumption per capita is the most commonly-used measure of welfare, some analysts use consumption per adult equivalent, in order to capture differences in need by age, and economies of scale in consumption. The OECD scale ($= 1 + 0.7 \times (N_A - 1) + N_C$) is popular, but such scales are controversial and cannot be estimated satisfactorily.

Other popular measures of welfare include Calorie consumption per person per day; food consumption as a proportion of total expenditure; and nutritional status (as measured by stunting or wasting). But there is no ideal measure of well-being, and analysts need to be aware of the strengths and limitations of any measure they use.

Learning Objectives

After completing the module on *Measuring Poverty*, you should be able to:

6. Summarize the three steps required to measure poverty.
7. Recognize the strengths and limitations arising from the need to use survey data in poverty analysis, including the choice of sample frame, unit of observation, time period, and choice of welfare indicators.
8. Describe the main problems that arise with survey data, including
    a. survey design (sampling frame/coverage, response bias),
    b. stratification, and
    c. multistage cluster sampling.
9. Explain why weighting is needed when surveys use stratified random sampling.
10. Describe and evaluate the use of equivalence scales (including the OECD scale).
11. Define consumption and income as measures of welfare, and evaluate the desirability of each in the LDC context.
12. Summarize the problems that arise in measuring income and consumption, and explain how to value durable goods, and housing services.
13. Identify measures of household welfare other than consumption and income, including Calorie consumption per capita, nutritional status, health status, and food consumption as a proportion of total expenditure.
14. Argue the case that there is no ideal measure of welfare.

2.1 Steps in measuring poverty

The goal of this chapter is to set out a method for measuring poverty. There is an enormous literature on the subject, so we just set out the main practical issues, with some suggestions for further reading for those interested in pursuing the subject in more depth.

Three steps need to be taken in measuring poverty (for more discussion see Ravallion, 1998). These are:

- Defining an indicator of welfare;
- Establishing a minimum acceptable standard of that indicator to separate the poor from the non-poor (the poverty line), and;
- Generating a summary statistic to aggregate the information from the distribution of this welfare indicator relative to the poverty line.

This chapter defines an indicator of welfare, while chapter 3 discusses the issues involved in setting a poverty line and chapter 4 deals with measuring aggregate welfare and its distribution.

2.2 Household surveys

2.2.1 Key survey issues

All measures of poverty rely on household survey data. So it is important to recognize the strengths and limitations of such data, and to set up and interpret the data with care. The analyst should be aware of the following issues (see Ravallion (1999) for details):

i) *The sample frame*: The survey may represent a whole country's population, or some more narrowly defined sub-set, such as workers or residents of one region. The appropriateness of a survey's particular sample frame will depend on the inferences one wants to draw from it. Thus a survey of urban households would allow one to measure urban poverty, but not poverty in the country as a whole.

ii) *The unit of observation*: This is typically the household or (occasionally) the individuals within the household. A household is usually defined as a group of persons eating and living together.

iii) *The number of observations over time*: A single cross-section, based on one or two interviews, is the most common. *Longitudinal surveys*, in which the same households or individuals are re-surveyed over an extended period (also called panel data sets) are more difficult to do, but have been undertaken in a few countries (e.g. the Vietnam Living Standards Surveys of 1993 and 1998).

iv) *The principal living standard indicator collected*: The most common indicators used in practice are based on household consumption expenditure and household income. The most common survey used in poverty analysis is a single cross-section for a nationally representative sample, with the household as the unit of observation, and it includes data on consumption and/or income. This form of survey is cheaper per household surveyed than most alternatives, thereby allowing a larger sample than with a longitudinal or individual-based survey. A larger sample of household-level data gives greater accuracy in estimating certain population parameters, such as average consumption per capita, but can lose accuracy in estimating other variables, such as the number of under-nourished children in a population (which may require oversampling of the target group). It should not, however, be presumed that the large household consumption survey is more cost-effective for all purposes than alternatives, such as using smaller samples of individual data.

2.2.2 Common survey problems

One needs to be aware of a number of problems when interpreting household consumption or income data from a household survey.

2.2.2.1 Survey design

Even a very large sample may give biased estimates for poverty measurement if the survey is not random, or if the data extracted from it have not been corrected for possible biases, such as due to sample stratification. A random sample requires that each person in the population, or each sub-group in a stratified sample, have an equal chance of being selected.

However, the poor may not be properly represented in sample surveys; for example they may be harder to interview because they live in remote areas, or are itinerant, or live illegally in the cities and so do not appear on the rosters of the local authorities. Household surveys almost always miss one distinct sub-group of the poor: those who are homeless. Also, some of the surveys that have been used to measure poverty were not designed for this purpose, in that their sample frames were not intended to span the entire population.

Examples: This is true, for instance, of labor force surveys, which have been widely used for poverty assessments in Latin America; the sample frame is typically restricted to the "economically active population," which precludes certain sub-groups of the poor. Or to take another example, household surveys in South Korea have typically excluded one-person households from the sample frame, which makes the results unrepresentative.

Key questions to ask about the survey are:

a) Does the sample frame (the initial listing of the population from which the sample was drawn) span the entire population?

b) Is there likely to be a response bias? This may take one of two forms – *unit non-response*, which occurs when some households do not participate in the survey, and *item non-response*, which occurs when some households do not respond fully to all the questions in the survey.

It is sometimes cost-effective deliberately to oversample some small groups (e.g. minority households in remote areas) and to undersample large and homogeneous groups. Such *stratified random sampling* – whereby different sub-groups of the population have different (but known) chances of being selected but all have an equal chance in any given subgroup – can increase the precision in poverty measurement obtainable with a given number of interviews. When done, it is necessary to use weights when analyzing the data, as explained more fully below.

2.2.2.2 Sampling

Two important implications flow from the fact that measures of poverty and inequality are always based on survey data.

First, it means that actual measures of poverty and inequality are *sample statistics*, and so estimate the true population parameters with some error. Although it is standard practice to say that, for

instance, "the poverty rate is 15.2%," it would be more accurate to say something like "we are 99% confident that the true poverty rate is between 13.5% and 16.9%; our best point estimate is that it is 15.2%." Outside of academic publications, such caution is rare.

The second implication is that it is essential to know how the sampling was done, because the survey data may need to be weighted in order to get the right estimates of such measures as mean income, or poverty rates. In practice, most household surveys oversample some areas (such as low-density mountainous areas, or regions with small populations), in order to get adequately large samples to compute tolerably accurate statistics for those areas. Conversely, areas with dense, homogeneous populations tend to be undersampled. For instance, the Vietnam Living Standards Survey of 1998 (VLSS98) oversampled the sparsely-populated central highlands, and undersampled the dense and populous Red River Delta.

In cases such as this, *it is not legitimate to compute simple averages of the sample observations* (such as per capita income, for instance) in order to make inferences about the whole population. Instead, weights must be used, as the following example shows.

Example: Consider the case of a country with 10 million people, who have a mean annual per capita income of $1,200. Region A is mountainous and has 2 million people with average per capita incomes of $500; region B is lowland and fertile and has 8 million people with an average per capita income of $1,375.

Now suppose that a household survey samples 2,000 households, picked randomly from throughout the country. The mean income per capita of this sample is the best available estimator of the per capita income of the population, and so we may calculate this and other statistics using the simplest available formulae (which are generally the ones shown in this manual). The Vietnam Living Standards Survey of 1993 (VLSS93) essentially chose households using a simple random sample, using the census data from 1989 to determine where people lived; thus the data from the VLSS93 are easy to work with, because no special weighting procedure is required.

Further details are set out in Table 2.1. If 400 households are surveyed in Region A (one household per 5,000 people) and 1,600 in Region B (one household per 5,000 people), then each household surveyed effectively "represents" 5,000 people; a simple average of per capita income

($1,215.6), based on the survey data, would then generally serve as the best estimator of per capita income in the population at large, as shown in the "Case 1" panel in Table 2.1.

But now suppose that 1,000 households were surveyed in Region A (one per 2,000 people) and another 1,000 in Region B (one per 8,000 people). If weights were not used, the estimated income per capita would be $943.5 (see the "Case 2a" panel in Table 2.1), but this would be incorrect. Here, a weighted average of observed income per capita is needed in order to compute the national average. Intuitively, each household sampled in Region A should get a weight of 2,000 and each household in Region B should be given a weight of 8,000 (see Table 2.1). The mechanics are set out in the "Case 2b" panel in Table 2.1, and yield an estimated per capita income of $1,215.6.

Table 2.1.
Illustration of why weights are needed to compute statistics based on stratified samples

	Region A	Region B	Whole country
Population (m)	2.0	8.0	10.0
True income/capita ($ p.a.)	500	1,375	1,200
Case 1. Simple random sampling. Use simple average.			
Sample size (given initially)	400	1,600	2,000
Estimated total income, $	196,000 =400*490	2,235,200 =1,600 * 1,397	2,431,200 =196,000 + 2,235,200
Estimated income/capita, ($ p.a.)*	490	1,397	**1,215.6** =2,431,200/2000
Case 2. Stratified sampling.			
Sample size (given initially)	1,000	1,000	2,000
Estimated total income, $	490,000 =1,000*490	1,397,000 =1,000 * 1,397	1,887,000 =490,000 + 1,397,000
Case 2a. Stratified sample, using simple average. This is incorrect, so don't do this!			
Estimated income/capita ($ p.a.)	490	1,397	**943.5** =1,887,000/2000
Case 2b. Stratified sampling, using weighted average. This is the correct approach.			
Weight (Based on population)	0.2 = 2.0/10.0	0.8 =8.0/10.0	
Estimated income/capita ($ p.a.)	490	1,397	**1,215.6** = .2*490 + .8*1,397.

Note: * Estimated income per capita is likely to differ from true income per capita, due both to sampling error (only a moderate number of households were surveyed) and non-sampling error (e.g. under-reporting, poorly worded questions, etc.).

In picking a sample, most surveys use the most recent population census numbers as the sample frame. Typically, the country is divided into regions, and a sample picked from each region (referred to as a *stratum* in the sampling context). Within each region, subregional units (towns, counties, districts, communes, etc.) are usually chosen randomly, with the probability of being picked being in proportion to population size. Such multistage sampling may even break down the units further (e.g. to villages within a district).

At the basic level (the "primary sampling unit" such as a village, hamlet, or city ward) it is standard to sample households in clusters. Rather than picking individual households randomly throughout a whole district, the procedure is typically to pick a couple of villages and then randomly sample 15-20 households within each chosen village. The reason for doing cluster sampling, instead of simple random sampling, is that it is cheaper. But it has an important corollary: the information provided by sampling clusters is less reliable as a guide to conditions in the overall area than pure random sampling would be. To see this, compare Figure 2.1.a (simple random sampling) with Figure 2.1.b (cluster sampling). Although, on average, cluster sampling will give the correct results (for per capita income, for instance), it is less reliable because we might, by chance, have chosen two particularly poor clusters, or two rich ones. Thus cluster sampling produces larger standard errors for the estimates of population parameters. This needs to be taken into account when programming the statistical results of sample surveys. Not all statistical packages handle clustering; however, Stata deals with it well using the *svyset* commands (see Appendix for details).

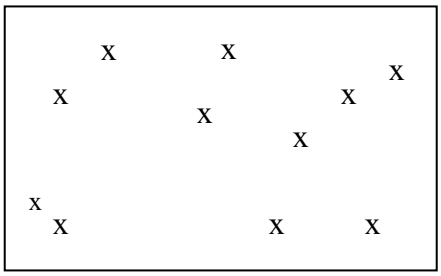


Figure 2.1.a Simple random sample

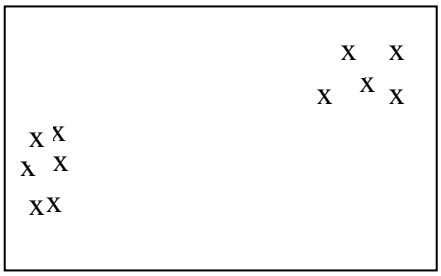


Figure 2.1.b Cluster sampling

Most living standards surveys sample households rather than individuals. If the variable of interest is household-based – for instance the value of land owned per household, or the educational level of the household head – then the statistics should be computed using household weights. But many measures relate to individuals (for instance, income per capita), in which case the results need to be computed using *individual weights*, which are usually computed as the household weights times the size of the household. Most, but not all, statistical packages handle this easily, but the analyst still has to provide the appropriate instructions.

2.2.2.3 Goods coverage and valuation

The coverage of goods and income sources in the survey should be comprehensive, including both food and non-food goods, and all income sources. Consumption should cover all monetary expenditures on goods and services consumed plus the estimated monetary value of all consumption from

income in kind, such as food produced on the family farm and the rental value of owner-occupied housing. Similarly, the income definition should include income in kind. Local market prices often provide a good guide for valuation of own-farm production or owner occupied housing.

However, whenever prices are unknown, or are an unreliable guide to reflect opportunity costs, serious valuation problems can arise. The valuation of access to public services is also difficult, and rarely done, though it is important. For transfers of in-kind goods, prevailing equivalent market prices are generally considered to be satisfactory for valuation. Non-market and durable goods present more serious problems, and there is no widely preferred method; we return to this problem in more detail below

2.2.2.4 Variability and the time period of measurement

Income and consumption vary from month to month, year to year, and over a lifetime. But income typically varies more significantly than consumption. This is because households try to smooth their consumption over time, for instance by managing their savings, or through risk-sharing arrangements (e.g. using remittances). In Less-Developed Countries, most (but not all) analysts prefer to use current consumption than current income as an indicator of living standards in poor countries, because:

i) in the short-run it reflects more accurately the resources that households control;
ii) over the long-term, it reveals information about incomes at other dates, in the past and future; and
iii) in poor countries, income is particularly difficult to measure accurately.

However, a number of factors can make current consumption a "noisy" welfare indicator. Even with ideal smoothing, consumption will still (as a rule) vary over a person's life-cycle, although this may be less of a problem in traditional societies where resource pooling within an extended family is still the norm. Another source of noise is that different households may face different constraints on their opportunities for consumption smoothing. It is generally thought that the poor are far more constrained in their ability to smooth consumption – mainly due to lack of borrowing options – than the non-poor.

2.2.2.5 Comparisons across households at similar consumption levels

Household size and demographic composition vary across households, as do the prices they face, including wage rates. As a result, it takes different resources to make ends meet for different households. In other words, at a given level of household expenditure, different households may achieve different levels of well-being: an annual income of $1,000 might suffice for a couple living in a rural area (where food and housing are cheap), but be utterly inadequate for a family of four in an urban setting.

There are a number of approaches, including equivalence scales, true cost-of-living indices, and equivalent income measures, which try to deal with this problem. The basic idea of these methods of welfare measurement is to use demand patterns to reveal consumer preferences over market goods. The consumer is assumed to maximize utility, and a utility metric is derived that is consistent with observed demand behavior, relating consumption to prices, incomes, household size, and demographic composition. The resulting measure of household utility will typically vary positively with total household expenditures, and negatively with household size and the prices faced.

The most widely-used formulation of this approach is the concept of "equivalent income", defined as the minimum total expenditure that would be required for a consumer to achieve his or her actual utility level but evaluated at pre-determined (and arbitrary) reference prices and demographics fixed over all households. This gives an exact monetary measure of utility (and, indeed, it is sometimes called "money-metric utility"). Quite generally, equivalent income can be thought of as money expenditures (including the value of own production) normalized by two deflators: a suitable price index (if prices vary over the domain of the poverty comparison) and an equivalence scale (since household size and composition varies).

One of the most serious problems that arises when using equivalent income as a measure of welfare is that it does not usually include a measure of the value of access to non-market goods (e.g. public services, community characteristics), yet this varies across households. Thus two households with the same income and demographic structure may not be equally one off if one of them has access to better roads and schools, and a nicer climate.

Unfortunately there is no satisfactory solution to this particular problem, although some studies do try to include a measure of the value of at least some publicly-provided services. Information on these often comes from a separate community survey (done at the same time as the interviews, and possibly by the same interviewers), which can provide useful supplementary data on the local prices of a range of goods and local public services.

2.2.3 Key features of Living Standards Measurement (LSMS) surveys

Motivated by the need to measure poverty more accurately, the World Bank has taken a lead in the development of relatively standard, reliable household surveys, under its Living Standards

Measurement (LSMS) project. The electronic version of the books edited by Grosh and Glewwe (2000) includes sample questionnaires and detailed chapters that deal with the design and implementation of such surveys. The LSMS surveys have two key features: multi-topic questionnaires, and considerable attention to quality control. Let's consider each in more detail.

2.2.3.1 Multi-topic questionnaires

The LSMS surveys ask about a wide variety of topics, and not just demographic characteristics or health experience or some other narrow issue.

- The most important single questionnaire is the *household questionnaire*, which often runs to 100 pages or more. Although there is an LSMS template, each country needs to adapt and test its own version. The questionnaire is designed to ask questions of the best-informed household member. The household questionnaire asks about household composition, consumption patterns including food and non-food, assets including housing, landholding and other durables, income and employment in agriculture/non-agriculture and wage/self-employment, socio-demographic variables including education, health, migration, fertility, and anthropometric information (especially the height and weight of each household member).
- There is also a *community questionnaire*, which asks community leaders (teachers, health workers, village officials) for information about the whole community, such as the number of health clinics, access to schools, tax collections, demographic data, and agricultural patterns. Sometimes there are separate community questionnaires for health and education.
- The third part is the *price questionnaire*, which collects information about a large number of commodity prices in each community where the survey is undertaken. This is useful because it allows analysts to correct for differences in price levels by region, and over time.

2.2.3.2 Quality control

The LSMS surveys are distinguished by their attention to quality control. Here are some of the key features:

- Most importantly, they devote a lot of attention to obtaining a representative national sample (or regional sample, in a few cases). Thus the results can usually be taken as nationally representative. It is surprising how many other surveys are undertaken with less attention to sampling, so one does not know how well they really represent conditions in the country.
- The surveys make extensive use of "screening questions" and associated skip patterns. For instance, a question might ask whether a family member is currently attending school; if yes, one

jumps to page x and asks for details; if no, then the interviewer jumps to page y and asks other questions. This cuts down on interviewer errors.

- Numbered response codes are printed on the questionnaire, so the interviewer can write a numerical answer directly on the questionnaire. This makes subsequent computer entry easier, more accurate, and faster.
- The questionnaires are designed to be easy to change (and to translate), which makes it straightforward to modify them in the light of field tests.
- The data are collected by decentralized teams. Typically each team has a supervisor, two interviewers, a driver/cook, an anthropometrist, and someone who does the data entry onto a laptop computer. The household questionnaire is so long that it requires two visits for collecting the data. After the first visit, the data are entered; if errors arise, they can be corrected on the second visit, which is typically two weeks after the first visit. In most cases the data are entered onto printed questionnaires, and then typed into a computer, but some surveys now enter the information directly into computers.
- The data entered are subject to a series of range checks. For instance, if an age variable is greater than 100, then it is likely that there is an error, which needs to be corrected.

This concern with quality has some important implications, notably:

- The LSMS data are usually of high quality, with accurate entries and few missing values.
- Since it is expensive to maintain high quality, the surveys are usually quite small; the median LSMS survey covers just 4,200 households. This is a large enough sample for accurate information at the national level, and at the level of half a dozen regions, but not at a lower level of disaggregation (e.g. province, department, county).
- The LSMS data have a fairly rapid turnaround time, with some leading to a statistical abstract (at least in draft form) within 2-6 months of the last interview.

2.3 Measuring poverty: choose an indicator of welfare

There are a number of conceptual approaches to the measurement of well-being. The most common approach is to measure economic welfare based on *household* consumption expenditure or *household* income. When divided by the number of household members, this gives a *per capita* measure of consumption expenditure or income. Of course, there are also non-monetary measures of individual welfare, which can include indicators such as infant mortality rates in the region, life expectancy, the

proportion of spending devoted to food, housing conditions, and child schooling. Well-being is a broader concept than economic welfare, which only measures a person's command over commodities.

If we choose to assess poverty based on household consumption or expenditure *per capita*, it is helpful to think in terms of an expenditure function, which shows the minimum expense required to meet a given level of utility u, which is derived from a vector of goods x, at prices p. It can be derived from an optimization problem in which the objective function (expenditure) is minimized subject to a set level of utility, in a framework where prices are fixed.

Let the consumption measure for the household i be denoted by y_i. Then an expenditure measure of welfare may be denoted by:

(2.1) $$y_i = p \cdot q = e(p, x, u)$$

where p is a vector of prices of goods and services, q is a vector of quantities of goods and services consumed, $e(.)$ is an expenditure function, x is a vector of household characteristics (e.g. number of adults, number of young children, etc.) and u is the level of "utility" or well-being achieved by the household. Put another way, given the prices (p) that it faces, and its demographic characteristics (x), y_i measures the spending that is needed to reach utility level u.

Typically, we compute the actual level of y_i from household survey data that include information on consumption. The details of this are discussed below. Once we have computed $y_{i,}$ we can construct *per capita* household consumption for every individual in the household, which implicitly assumes that consumption is shared equally among household members. For this approach to make sense, we must also assume that all individuals in the household have the same needs. This is a strong assumption, for in reality, different individuals have different needs based on their individual characteristics (age, gender, job, etc).

While estimating per capita consumption might seem straightforward, there are several factors that complicate its estimation. Table 2.2 reports estimates of both nominal and inflation-adjusted ("real") per capita consumption from three different household surveys in Cambodia. Using the 1997 Cambodia Socio-economic Survey (CSES), for example, nominal and real per capita consumption were 2,223 and 1,887 riels, respectively. However, across years the estimates in real terms for 1993/94 may not be directly comparable with the 1999 estimates *because the surveys did not have exactly the same set of questions regarding consumption*. For example, real consumption per capita was computed as 2,262 riels

for 1993/94, but was only 1,700 in 1999, despite economic growth during the interval; this may merely be an artifact of the different ways in which questions were asked.

Table 2.2: Summary of per capita consumption from Cambodian Surveys

Surveys	Nominal	Real (inflation adjusted)
SESC 1993/94	1,833	2,262
CSES 1997 (adjusted)	2,223	2,530
CSES 1997 (unadjusted)	1,887	2,153
CSES 1999 (Round 1)	2,037	1,630
CSES 1999 (Round 2)	2,432	1,964
CSES 1999 (both Rounds)	2,238	1,799

Note: All values are in Riels per person per day. Real values are estimated in 1993/94 Phnom Penh prices, as deflated by the value of the food poverty lines. Adjusted figures from 1997 incorporate corrections for possible underestimation of certain types of consumption (see Knowles 1998, and Gibson 1999 for details). Differences between Rounds 1 & 2 in 1999 are detailed in Gibson (1999). CSES: Cambodia Socio-Economic Survey. SESC: Socio-Economic Survey of Cambodia.
Source: Gibson (1999)

Traditionally, we use a monetary measure to value household welfare. The two most obvious candidates are income and expenditure.

2.3.1 Candidate 1: Income

It is tempting to measure household welfare by looking at household income. Practical problems arise immediately: what is income? and can it be measured accurately? The most generally accepted measure of income is the one formulated by Haig and Simons:

$$\text{Income} \equiv \text{consumption} + \text{change in net worth.}$$

Example: Suppose I had assets of $10,000 at the beginning of the year. During the year I spent $3,000 on consumption. And at the end of the year I had $11,000 in assets. Then my income was $4,000, of which $3,000 was spent, and the remaining $1,000 added to my assets.

The first problem with this definition is that it is not clear what time period is appropriate. Should we look at someone's income over a year? Five years? A lifetime? Many students are poor now, but have good lifetime prospects, and we may not want to consider them as being truly poor. On the other hand, if we wait until we have information about someone's lifetime income, it will be too late to help him or her in moments of poverty.

The second problem is measurement. It is easy enough to measure components of income such as wages and salaries. It may be possible to get adequate (if understated) information on interest, dividends, and income from some types of self-employment. But it is likely to be hard to get an accurate measure of farm income; or of the value of housing services; or of capital gains (e.g. the increase in the value of animals on a farm, or the change in the value of a house that one owns).

For instance, the Vietnam Living Standard Survey (VLSS; undertaken in 1993 and again in 1998) collected information on the value of farm animals at the time of the survey, but not the value a year before. Thus it was not possible to measure the change in the value of animal assets. Many farmers that reported negative cash income may in fact have been building up assets, and truly had positive income.

It is typically the case, particularly in societies with large agricultural or self-employed populations, that income is seriously understated. This certainly appears to be the case for Vietnam. Table 2.3 shows income per capita for households in 1993 for each of five expenditure quintiles: a quintile is a fifth of the sample, and quintile 1 contains the poorest fifth of individuals, etc. For every quintile, households on average reported less income than expenditure, which is simply not plausible. This would imply that households must be running down their assets, or taking on much more debt, which was unlikely in a boom year like 1993.

Table 2.3: Income and expenditure by per capita expenditure quintiles, Vietnam

(In thousands of dong per capita per year, 1992/93)

	Lowest	Lower-mid	Middle	Mid-upper	Highest	Overall
Income/capita	494	694	956	1,191	2,190	**1,105**
Expenditure/capita	518	756	984	1,338	2,540	**1,227**
Memo: food						
… spending/capita	378	526	643	807	1,382	**747**
…as % of expend.	73	70	65	60	54	**61**

Note: In 1993, exchange rate was about 10,000 dong/US$. *Source:* VLSS93

.

There are a number of reasons why income tends to be understated:

- People forget, particularly when asked in a single interview about items they may have sold, or money they may have received, up to a year before.
- People may be reluctant to disclose the full extent of their income, lest the tax collector, or neighbors, get wind of the details.
- People may be reluctant to report income earned illegally - for instance from smuggling, or corruption, or poppy cultivation, or prostitution.

- Some parts of income are difficult to observe - e.g. the extent to which the family buffalo has risen in value.

Research based on the 1969-70 socio-economic survey in Sri Lanka estimated that wages were understated by 30%, business income by 39%, and rent, interest and dividends by 78%. It is not clear how much these figures are applicable elsewhere, but they do give a sense of the potential magnitude of the understatement problem.

2.3.2 Candidate 2: Consumption expenditure

Note that consumption includes both goods and services that are purchased, and those that are provided from one's own production ("in-kind").

In developed countries, a strong case can be made that consumption is a better indicator of lifetime welfare than is income. Income typically rises and then falls in the course of one's lifetime, in addition to fluctuating somewhat from year to year, whereas consumption remains relatively stable. This smoothing of short-term fluctuations in income is predicted the permanent income hypothesis, under which transitory income is saved while long-term ("permanent") income is largely consumed.

The life cycle of income and consumption is captured graphically in figure 2.2. While the available evidence does not provide strong support for this *life-cycle hypothesis* in the context of less-developed countries, households there do appear to smooth out the very substantial seasonal fluctuations in income that they typically face during the year (see Alderman and Paxson 1994; Paxson 1993). Thus information on consumption over a relatively short period – a month for instance – as typically collected by a household survey is more likely to be representative of a household's general level of welfare than equivalent information on income (which is more volatile).

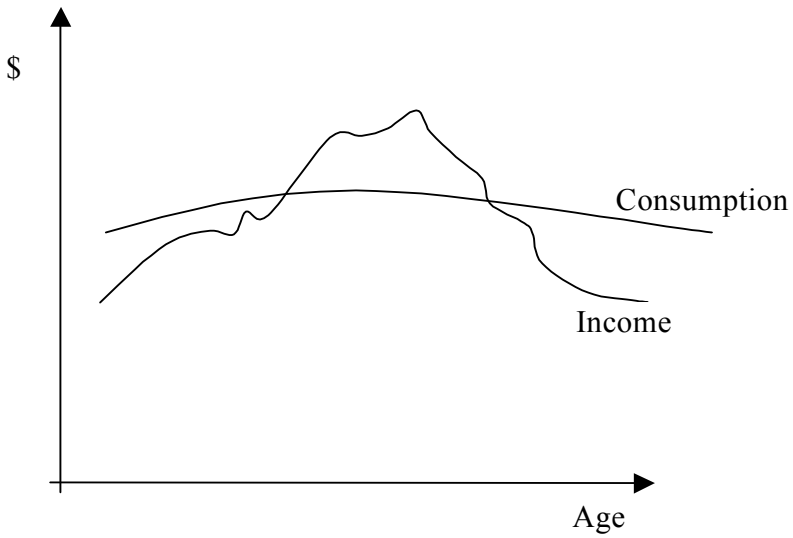


Figure 2.2 Life Cycle Hypothesis: Income and Consumption Profile over Time

A more practical case for using consumption, rather than income, is that households may be more able, or willing, to recall what they have spent rather than what they earned. Even so, consumption is likely to be systematically understated, because:

- Households tend to under-declare what they spend on luxuries (e.g. alcohol, cakes) or illicit items (drugs, prostitution). For instance, the amount that households said they spent on alcohol, according to the 1972-73 household budget survey in the US, was just half the amount that companies said they sold!
- Questions matter. According to VLSS93, Vietnamese households devoted 1.7% of their expenditure to tobacco; the VLSS98 figures showed that this had risen to 3%. An increase of this magnitude is simply not plausible, and not in line with sales reported by the cigarette and tobacco companies. A more plausible explanation is that VLSS98 had more detailed questions about tobacco use. When the questions are more detailed, respondents are likely to remember in more detail and to report higher spending.

2.3.2.1 Measuring durable goods

In measuring poverty it might be argued that only food, the ultimate basic need (which anyway constitutes three quarters of the spending of poor households), should be included. On the other hand, even households that cannot afford adequate quantities of food devote some expenditures to other items (clothing, shelter, etc.). It is reasonable to suppose that if these items are getting priority over food purchases, then they must represent very basic needs of the household, and so should be included in the poverty line. This argument also applies to durable goods (housing, pots and pans, etc.).

The problem here is that durable goods, such as bicycles and TVs, are bought at a point in time, and then consumed (i.e. eaten up and destroyed) over a period of several years. Consumption should only include the amount of a durable good that is eaten up during the year, which can be measured by the change in the value of the asset during the year, plus the cost of locking up one's money in the asset.

> **Example:** For instance, if my watch was worth $25 a year ago, and is worth $19 now, then I used $6 worth of watch during the year; I also tied up $25 worth of assets in the watch, money that could have earned me $2.50 in interest (assuming 10%) during the year. Thus the true cost of the watch during the year was $8.50.

A comparable calculation needs to be done for each durable good that the household owns. Clearly the margin of potential measurement error is large, since the price of each asset may not be known with much accuracy, and the interest rate used is somewhat arbitrary. The Vietnamese VLSS surveys asked for information about when each good was acquired, and at what price, and the estimated current value of the good. This suffices to compute the current consumption of the durable item, as the illustration in the following box shows.

One might wonder why attention needs to be paid to calculating the value of durable goods consumption when the focus is on poverty - in practice first and foremost the ability to acquire enough food. The answer is that when expenditure is used as a yardstick of welfare, it is important to achieve comparability across households. If the value of durable goods were not included, one might have the impression that a household that spends $100 on food and $5 on renting a bicycle is better off than a household that spends $100 on food and owns a bicycle (that it could rent out for $5), when in fact both households are equally well off (ceteris paribus).

Box: Calculating the value of durable goods consumption - an illustration.

A Vietnamese household is surveyed in April 1998, and says that it bought a TV two years earlier for 1.1m dong (about $100). The TV is now believed to be worth 1m dong. Overall prices rose by 10% over the past two years. How much of the TV was consumed over the year prior to the survey?

a. Recompute the values in today's prices. Thus the TV, purchased for 1.1m dong in 1996, would have cost 1.21m dong (=1.1m dong × (1+10%)) now.
b. Compute the depreciation. The TV lost 0.21m dong in value in two years, or 0.105m dong per year (i.e. about $7).
c. Compute the interest cost. At today's prices, the TV was worth 1.105m dong a year ago (i.e. 1.21m dong less this past year's depreciation of 0.105m dong), and this represents the value of funds locked up during the year prior to the survey. At a real (i.e. inflation-adjusted) interest rate of 3%, the cost of locking up these resources was 0.03315m dong over the course of the year.

Thus the total consumption cost of the TV was 0.138m dong (= 0.105 + 0.033), or about $10.

Note that this computation is only possible if the survey collects information on the past prices of all the durables used by the household. Where historical price data are not available, researchers in practice typically apply a depreciation+interest rate to the reported value of the goods; so if a TV is worth 1m dong now, is expected to depreciate by 10% per annum, and the real interest rate is 3%, then the imputed consumption of the durable good is measured as 1m × (10% + 3%) = 0.13m dong. Deaton and Zaidi (1998) recommend that one use average depreciation rates derived from the sample, rather than the rates reported by each individual household.

2.3.2.2 Measure the value of housing services

If you own your house (or apartment), it provides housing services, which should be considered as part of consumption. The most satisfactory way to measure the values of these services is to ask how much you would have to pay if, instead of owning your home, you had to rent it.

The standard procedure is to estimate, for those households that rent their dwellings, a function that relates the rental payment to such housing characteristics as the size of the house (in sq. ft. of floor space), the year in which it was built, the type of roof, whether there is running water, etc. This gives

Rent = f(area, running water, year built, type of roof, location, number of bathrooms, …)

This equation is then used to impute the value of rent for those households that own, rather than rent, their housing. For all households that own their housing, this imputed rental, along with the costs of

maintenance and minor repairs, represents the annual consumption of housing services.[1] In the case of households that pay interest on a mortgage, it is appropriate to count the imputed rental and costs of maintenance and minor repairs in measuring consumption, but not the mortgage interest payments as well, because this would represent double-counting.[2]

In the case of Vietnam there is a problem with this approach: almost nobody rents housing! And of those that do, most pay a nominal rent for a government apartment. Only 13 of the 5,999 households surveyed in VLSS98 paid private-sector rental rates. On the other hand the VLSS surveys did ask each household to put a (capital) value on their house (or apartment). In computing consumption expenditure, the rental value of housing was assumed to be 3 percent of the capital value of the housing. This is a somewhat arbitrary procedure, but the 3 percent is almost certainly too low.

2.3.2.3 Weddings and Funerals.

Families spend money on weddings. Such spending is often excluded when measuring household consumption expenditure. The logic is that the money spent on weddings mainly gives utility to the guests, not the spender. Of course if one were to be strictly correct, then expenditure should include the value of the food and drink that one enjoys as a guest at other people's weddings, although in practice this is rarely (if ever) included. Alternatively one might think of wedding expenditures as rare and exceptional events, which shed little light on the living standard of the household. Similar considerations apply to other large and irregular spending, on items such as funerals and dowries.

2.3.2.4 Accounting for household composition differences

Households differ in size and composition, and so a simple comparison of aggregate household consumption can be quite misleading about the well-being of individuals in a given household. Most researchers recognize this problem and use some form of normalization. The most straightforward method is to convert from household consumption to individual consumption by dividing household expenditures by the number of people in the household. Then, total household expenditure per capita is

[1] This assumes that renters are responsible for maintenance and repair costs, so that the rental paid does not include a provision for these items. In some countries the owner, rather than the renter, would bear these costs, in which case the imputed rental also includes the costs, and no further adjustment would be called for.

[2] However, if we want to measure income (rather than consumption), then we should use the imputed rental for households that own their property free and clear, and rental less mortgage interest payments for those who have borrowed against their housing.

the measure of welfare assigned to each member of the household. Although this is by far the most common procedure, it is not very satisfactory, for two reasons:

- First, different individuals have different needs. A young child typically needs less food than an adult, and a manual laborer requires more food than an office worker.
- Second, there are economies of scale in consumption (at least for such items as housing). It costs less to house a couple than to house two single individuals.

Example. For example, suppose we have a household with 2 members and monthly expenditure of $150 total. We would then assign each individual $75 as their monthly per capita expenditure. If we have another household with 3 members, it would appear that each member is worse off, with only $50 per capita per month. However, suppose we know that the 2-person household contains two adult males aged 35 whereas the second household contains 1 adult female and 2 young children. This added information may change our interpretation of the level of well-being in the second household, since we suppose that young children may have much lower costs (at least for food) than adults.

In principle, the solution to this problem is to apply a system of weights. For a household of any given size and demographic composition (such as one male adult, one female adult, and two children), an equivalence scale measures the number of adult males (typically) to which that household is deemed to be equivalent. So each member of the household counts as some fraction of an adult male. Effectively, household size is the sum of these fractions and is not measured in numbers of persons but in numbers of *adult equivalents*. Economies of scale can be allowed for by transforming the number of adult equivalents into "effective" adult equivalents.

In the abstract, the notion of equivalence scale is compelling. It is much less persuasive in practice, because of the problem of picking an appropriate scale. How these weights should be calculated and whether it makes sense to even try is still subject to debate, and there is no consensus on the matter. However, equivalence scales are not necessarily unimportant. For example, take the observation that in most household surveys, per capita consumption decreases with household size. It is probably more appropriate to interpret this as evidence that there are economies of scale to expenditure, and not necessarily as proof that large households have a lower standard of living.

There are two possible solutions to this problem: either pick a scale that seems reasonable on the grounds that even a bad equivalence scale is better than none at all, or try to estimate a scale typically based on observed consumption behavior from household surveys. Often the equivalence scales are based on the different calorie needs of individuals of different ages.

OECD scale

Commonly used is the "OECD scale," which may be written as

(2.2) $$AE = 1 + 0.7(N_{adults} - 1) + 0.5N_{children}$$

where *AE* refers to "adult equivalent." A one-adult household would have an adult equivalent of 1, a two-adult household would have an *AE* of 1.7, and a three-adult household would have an *AE* of 2.4. Thus the 0.7 reflects economies of scale; the smaller this parameter, the more important economies of scale are considered to be. In developing countries, where food constitutes a larger part of the budget, economies of scale are likely to be less pronounced than in rich countries. The 0.5 is the weight given to children, and presumably reflects the lower needs (for food, housing space, etc.) of children. Osberg and Xu (1999) use the OECD scale in their study of poverty in Canada. Despite the elegance of the formulation, there are real problems in obtaining satisfactory measures of the degree of economies of scale and even of the weight to attach to children.

Other scales.

Many other scales have been used. For instance, a number of researchers used the following scale in analyzing the results of the living standards measurement surveys that were undertaken in Ghana, Peru and the Côte d'Ivoire:

Age (years)	0-6	7-12	13-17	>17
Weight (i.e. adult equivalences)	0.2	0.3	0.5	1.0

An elegant formulation is as follows:

$$AE = (N_{adults} + \alpha N_{children})^{\theta}$$

Where α measures the cost of a child relative to an adult and $\theta \leq 1$ is a parameter that captures the effects of economies of scale. Consider a family with two parents and two children. For $\alpha = \theta = 1$, AE = 4 and our welfare measure becomes expenditure per capita. But if $\alpha = 0.7$ and $\theta = 0.8$, then AE = 2.67, and the measure of expenditure per adult equivalent will be considerably larger.

Estimate an equivalence scale.

It is also possible to estimate (econometrically) an equivalence scale, essentially by looking at how aggregate household consumption of various goods during some survey period tends to vary with household size and composition, although Deaton and Zaidi (1998) argue, "there are so far no satisfactory methods for estimating economies of scale."

A common method is to construct a demand model in which the budget share devoted to food consumption of each household is regressed on the total consumption per person. Deaton (1997) gives an example using Engel's method with household expenditure survey data from India and Pakistan. Specifically, household food share is regressed on per capita expenditure, household size, and household composition variables such as the ratio of adults and ratios of children at different ages. The equivalence scales – here the ratio of costs of a couple with a child to a couple without children – can then be calculated with the estimated coefficients. They are displayed in table 2.4:

Table 2.4: Equivalence scales using Engel's method

Age	Maharashtra, India	Pakistan
0-4	1.24	1.28
5-9	1.28	1.36
10-14	1.30	1.38
15-54	1.34	1.42

Note: Reproduced from Deaton (1997) table 4.6. Numbers show cost of 2 adults plus 1 person of the age shown, relative to a childless couple.

The numbers show the estimated costs of a family of two adults plus one additional person of various ages calculated relative to the costs of a childless couple. So, for example, a child between 0 and 4 years is equivalent to 0.24 of a couple, or 0.48 of an adult. As the age of the additional member rises, the extra costs associated with the child rise. We can compare these estimates with the last row, which shows the equivalence scale when an additional adult is added to the household. An additional adult costs 34% more for the couple, or incurs 68% of the cost of one member of the couple. So, by these calculations, these households experience economies of scale to additional adults, plus younger members are not equivalent to adults in terms of costs.

Unfortunately, there are a number of problems with this method (see Ravallion, 1994 and Deaton 1997 for details). Consider the following example from Ravallion (1994), where there are two hypothetical households as described in table 2.5.

Table 2.5: Consumption within two hypothetical households						
	Male adult	Female adult	First child	Second child	Per person	Per equivalent male adult
Household A	40	20	10	10	20	29.6
Household B	25	-	-	-	25	25
Source: Adapted from Ravallion (1994). Uses the "OECD scale": $AE = 1 + 0.7(N_{adults} - 1) + 0.5N_{children}$.						

In this example, four persons live in household A but just one in household B. The government can make a transfer to the household that is deemed to be the poorest, but it cannot observe the distribution of consumption within the households. All the government knows is the aggregate expenditure and the household composition. In this case, ***which of the two households should have priority for assistance?***

Household A has lower consumption per capita and so looks worse off. But using equivalence scales as calculated here, household B would have priority in receiving assistance. This example demonstrates two points. First, while observable consumption behavior is important information, assumptions about unobservables (e.g. how the aggregate is split within the household) will be required. Second, assumptions in computing consumption for individuals using household data can have considerable bearing on policy choices.

Most rich countries measure poverty using income, while most poor countries use expenditure. There is a logic to this; in rich countries, income is comparatively easy to measure (much of it comes from wages and salaries) while expenditure is complex and hard to quantify. On the other hand, in less-developed countries income is hard to measure (much of it comes from self employment), while expenditure is more straightforward and hence easier to estimate. The arguments for and against income and consumption as the appropriate welfare measures for poverty analysis are summarized in Table 2.6.

Table 2.6. ***Which indicator of welfare: income or consumption?***	
Income ("potential")	
Pro: • Easy to measure, given the limited number of sources of income. • Measures degree of household "command" over resources (which they could use if they so wish). • Costs only a fifth as much to collect as expenditure data, so sample can be larger.	***Con:*** • Likely to be under-reported. • May be affected by short-term fluctuations (e.g. the seasonal pattern of agriculture). • Some parts of income are hard to observe (e.g. informal sector income; home agricultural production, self employment income). • Link between income and welfare is not always clear.

	• Reporting period might not capture the "average" income of the household.
Consumption ("achievement")	
Pro: • Shows current actual standard of living. • Smoothes out irregularities, and so reflects long-term average well-being. • Less understated than income, because expenditure is easier to recall.	**Con:** • Households may not be able to smooth consumption (e.g. via borrowing, social networks). • Consumption choices made by households may be misleading (e.g. if a rich household chooses to live simply, that does mean it is poor). • Some expenses are not incurred regularly, so data may be noisy. • Difficult to measure some components of consumption, including durable goods.
Source: Based on Albert, 2004.	

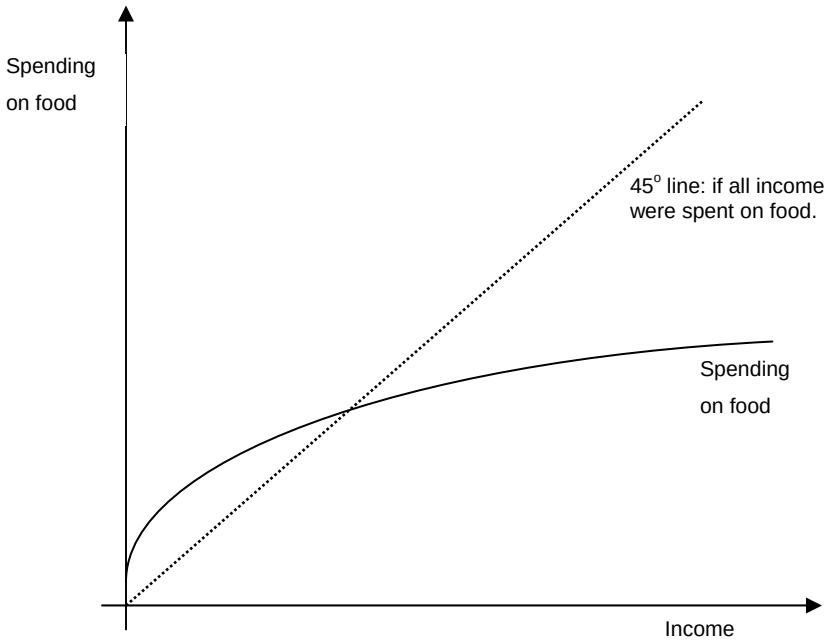
2.3.3 Candidate 3. Other measures of household welfare

Even if they were measured perfectly, neither income nor expenditure would be a perfect measure of household well-being. For instance, neither measure puts a value on the leisure time enjoyed by the household; neither measures the value of publicly-provided goods (such as education, or public health services); and neither values intangibles such as peace and security.

There are other possible measures of well-being. Among the more compelling are:

- *Calories consumed per person per day.* If one accepts the notion that adequate nutrition is a prerequisite for a decent level of well-being, then we could just look at the quantity of calories consumed per person. Anyone consuming less than a reasonable minimum - often set at 2,100 calories per person per day - would be considered poor. Superficially, this is an attractive idea, and we will return to it in chapter 3. However, at this point we just note that it is not always easy to measure calorie intake, particularly if one wants to distinguish between different members of a given household. Nor is it easy to establish the appropriate minimum amount of calories per person, as this will depend on the age, gender, and working activities of the individual.

Figure 2.3. Engel curve: food spending rises less quickly than income



- *Food consumption as a fraction of total expenditure.* Over a century ago Ernst Engel observed, in Germany, that as household income per capita rises, spending on food rises too, ***but less quickly***. This relationship is shown in figure 2.3. As a result, the proportion of expenditure devoted to food falls as per capita income rises. One could use this finding, which is quite robust to come up with a measure of well-being and hence a measure of poverty. For instance, households that devote more than (say) 60% of their expenditures to food might be considered to be poor. The main problem with this measure is that the share of spending going to food also depends on the proportion of young to old family members (more children indicates a higher proportion of spending on food), and on the relative price of food (if food is relatively expensive, the proportion of spending going to food will tend to be higher).

- *Measures of outcomes rather than inputs.* Food is an input, but nutritional status (being underweight, stunting or wasting) is an output. So one could measure poverty by looking at malnutrition. Of course, this requires establishing a baseline anthropometric standard against which to judge whether someone is malnourished. Anthropometric indicators have the advantage that they can reveal living

conditions within the household (rather than assigning the overall household consumption measure across all members of the household without really knowing how consumption expenditure is divided among household members). However, there is one further point about these measures: by some accounts, the use of child anthropometric measures to indicate nutritional need is questionable when broader concepts of well-being are invoked. For example, it has been found that seemingly satisfactory physical growth rates in children are sometimes maintained at low food-energy intake levels by not playing. That is clearly a serious food-related deprivation for any child.

- *Anthropological method.* Close observation at the household level over an extended period can provide useful supplementary information on living standards in small samples. However, this is unlikely to be a feasible method for national poverty measurement and comparisons. Lanjouw and Stern (1991) used subjective assessments of poverty in a north Indian village, based on classifying households into seven groups (very poor, poor, modest, secure, prosperous, rich and very rich) on the basis of observations and discussion with villages over that year.

  An issue of concern about this method is clearly its objectivity. The investigator may be working on the basis of an overly stylized characterization of poverty. For example, the poor in village India are widely assumed to be landless and underemployed. From the poverty profiles given by Lanjouw and Stern (1991) we find that being a landless agricultural laborer in their surveyed village is virtually a sufficient condition for being deemed poor. By their anthropological method, 99% of such households are deemed poor, though this is only so for 54% when their measurement of permanent income is used. It is clear that the perception of poverty is much more strongly linked to landlessness than income data suggest. But it is far from clear which data are telling us the most about the reality of poverty.

When one is looking at a community (e.g. province, region) rather than individual households, it might make sense to judge the poverty of the community by life expectancy, or the infant mortality rate, although these are not always measured very accurately. School enrollments (a measure of investing in the future generation) represent another outcome that might indicate the relative well-being of the population. Certainly, none of these other measures of well-being are replacements for consumption per capita–and nor does consumption per capita replace these measures. Rather, when taken together they allow us to get a more complete and multidimensional view of the well-being of a population, although this does not guarantee greater clarity. Consider the statistics in table 2.7, which refer to eleven different

countries. How countries are ranked in terms of living standards clearly depends on which measure or indicator is considered.

Table 2.7: Poverty and quality of life indicators

Countries	GNP per capita (1999 dollars)	% population below poverty line	Female life expectancy at birth, years (1998)	Prevalence of child malnutrition, % children <5 years (1992-1998)	Female adult illiteracy rate, % of people 15+years, (1998)
Algeria	1,550	22.6 (1995)	72	13	46
Bangladesh	370	35.6 (1995/96)	59	56	71
Cambodia	260	36.1 (1997)	55	na	80
Colombia	2,250	17.7 (1992)	73	8	9
Indonesia	580	20.3 (1998)	67	34	20
Jordan	1,500	11.7 (1997)	73	5	17
Morocco	1,200	19.0 (1998/99)	69	10	66
Nigeria	310	34.1 (1992/93)	55	39	48
Peru	2,390	49.0 (1997)	71	8	16
Sri Lanka	820	35.3 (1990/91)	76	38	12
Tunisia	2,100	14.1 (1990)	74	9	42
Source: World Bank (2000)					

In sum, there is no ideal measure of well-being. The implication is simple: all measures of poverty are imperfect. That is not an argument for avoiding measuring poverty, but rather for approaching all measures of poverty with a degree of caution, and for asking in some detail about how the measures are constructed.

Selected further reading:

Angus Deaton and Salman Zaidi. 1999. *Guidelines for Constructing Consumption Aggregates For Welfare Analysis*. Available at http://www.wws.princeton.edu/%7Erpds/downloads/deaton_zaidi_consumption.pdf [Accessed May 13, 2004]. Subsequently issued in 2002 as Living Standards Measurement Study Working Paper: 135. v. 104, pp. xi, Washington, D.C.: The World Bank.

A clear, sensible discussion of the practical issues that arise in measuring a consumption indicator of welfare. Includes a sample questionnaire and some useful Stata code.

Margaret Grosh and Paul Glewwe. 1998. "The World Bank's Living Standards Measurement Study Household Surveys," *Journal of Economic Perspectives*, 12(1): 187-196.

Margaret Grosh and Paul Glewwe (eds.). 2000. Designing Household Survey Questionnaires for Developing Countries: Lessons from Fifteen Years of Living Standard Measurement Study. Oxford University Press. Every statistics office should have a copy of this pair of volumes, or better still the CD-ROM version. This reference work includes sample questionnaires as well as detailed chapters on all aspects of designing, implementing and using living standard measurement surveys.

Margaret Grosh and Paul Glewwe. 1995. *A Guide to Living Standards Measurement Study Surveys and Their Data Sets*, Living Standards Measurement Study Working Paper No. 120, World Bank.

General Statistical Office (Vietnam). *Vietnam Living Standards Survey 1997-1998*, Statistical Publishing House, Hanoi 2000.

World Bank. *Vietnam: Attacking Poverty*, Hanoi, 1999.