Counting the World's Poor: Problems and Possible Solutions

Angus Deaton

As recent discussions have made clear, the apparent lack of poverty reduction in the face of historically high rates of economic growth—both in the world as a whole and in specific countries (most notably India)—provides fuel for the argument that economic growth does little to reduce poverty. How confident can we be that the data actually support these inferences? At the international level, the regular revision of purchasing power parity exchange rates plays havoc with the poverty estimates, changing them in ways that have little or nothing to do with the actual experience of the poor. At the domestic level, the problems in measuring poverty are important not only for the world count but also for tracking income poverty within individual countries. Yet, in many countries, there are large and growing discrepancies between the survey data—the source of poverty counts—and the national accounts—the source of the measure of economic growth. Thus economic growth, as measured, has at best a weak relationship with poverty, as measured.

The World Bank prepares and publishes estimates of the number of poor people in the world. Although these numbers should be taken with a pinch of salt, they are arguably important. In an institution where the reduction of poverty is the paramount objective, some overall yardstick of progress (or the lack of it) is required. The numbers are frequently quoted by politicians and by leaders of international organizations, including the World Bank itself, who believe the numbers are effective for advocacy. Indeed, there is a long history of studies of poverty mobilizing support among the nonpoor for antipoverty policies. So it is important to know whether the world and national poverty counts are sound enough to support these uses.

As recent discussions have made clear, the apparent lack of poverty reduction in the face of historically high rates of economic growth—both in the world as a whole and in specific countries (most notably India)—is providing fuel for the argument that economic growth does little to reduce poverty. How confident can we be that the data actually support these inferences? Are the changes in the poverty counts sufficiently well measured to support conclusions about growth and poverty reduction? Should

the World Bank stand ready to be judged by its success in reducing the current measures of world or even national poverty? If not, can better data collection or better methodologies improve the numbers?

A more fundamental issue is whether the number of people suffering consumption (or income) deprivation is the right yardstick. It is widely agreed that other dimensions of poverty are more important than income poverty, particularly deprivations in health, education, and democratic rights. There are difficult measurement issues associated with all of these other dimensions, especially if we try to combine them into a single measure, such as the Human Development Index. For many purposes, we need to monitor changes over periods of a few years, which is typically quite difficult for nonincome-based measures. For example, the average literacy rate of a population changes too slowly to be much use as a monitoring device in all but the longest of runs. Even so, and independently of measurement issues, the importance of these nonincome dimensions of poverty should be a warning not to overemphasize income poverty at the expense of other measures. It should also be noted that other dimensions of deprivation are just that, "other measures." They are not perfectly correlated with income or consumption poverty so that, just as we cannot subsume health or literacy measures into an income measure, we also should not try to measure income poverty by using health or education as a proxy. Difficulty in measuring income poverty is a poor argument for focusing on other dimensions of welfare and deprivation.

This article discusses a number of problems with the current $1/day poverty counts, makes some suggestions for improvement, and identifies issues that need further research. World poverty numbers are calculated in two stages. At the first or international stage, a world poverty line is set and used to derive comparable poverty lines for each country. At the second or domestic stage, the poverty lines are used to count the number of poor people in each country, and the totals are added up over countries. It is useful to think about each stage separately, and there is disquieting evidence about both. At the international level, the purchasing power parity (PPP) exchange rates that are used to turn the $1/day poverty line into national currencies are arguably inappropriate in theory. In practice, their regular revision (to different base years with different relative prices) plays havoc with the poverty estimates, changing them in ways that have little or nothing to do with the actual experience of the poor.

Within each country, at the domestic stage, the World Bank's poverty count comes from household surveys. The incorporation of these unit record data into its statistical base is surely one of the Bank's major achievements over the past 20 years. Yet in many countries—including India, where around a third of the world's poor live—there are large and growing discrepancies between the survey data and the national accounts. Because poverty counts come from the survey data and growth measures come from the national accounts, and because they are evidently measuring differ-

ent things, there is no consistent empirical basis for conclusions about the extent to which growth reduces poverty. That economic growth, *as measured*, has at best a weak relationship with poverty, *as measured*, means little more than would a finding that growth in China had failed to reduce poverty in India; they are simply different things.

International Issues in Counting the Poor

I discuss the international issues first, followed by the domestic ones. For those who are not persuaded that the world poverty counts are useful or believe them to be "essentially propagandist" (Srinivasan 2000), these international issues are irrelevant.

Purchasing Power Parities and the Base Poverty Line

The current concept for world poverty is the number of people who live in households whose daily consumption per head is less than the PPP equivalent of $1/day in constant 1985 PPP dollars. In the latest revision, this has been updated to $1.08 in constant 1993 PPP dollars, but is still conveniently referred to as the $1/day poverty line (Chen and Ravallion 2000). The virtues of this approach are many. It is simple, is easy to remember, and applies equally to all countries. It is denominated in a currency that is familiar to the relatively wealthy people who are the primary users of the measures and who are the primary target for rhetoric based on them. The $1/day was originally selected as being representative of poverty lines in use in low-income countries; the new $1.08 is representative of current lines and is thus anchored in actual practice. The $1.08/day is converted to the 1993 value of local currencies using internal World Bank PPP tables, and then updated to the target year using a local consumption deflator. The resulting poverty line is then used to count the number of people in each poor (technically World Bank part two) country using household survey data. When possible, the raw data are used; when this is not possible, values are interpolations based on published tables from such surveys.

Given world and U.S. inflation between 1985 and 1993, it is somewhat surprising that the international poverty line should have increased by only 8 percent. But the updating was carried out by going back to the country poverty lines and converting back to international dollars. Thus the modest increase comes not from a failure to allow for world inflation, but because the PPP international dollar has strengthened relative to the currencies of the poor countries whose poverty lines are incorporated into the international line.

The PPP numbers come from data collected from 110 countries in 1993 by the International Comparison Project. Earlier PPP numbers came from the Penn World Tables, which covered only 60 countries. The revision from the earlier PPPs to the new PPPs in the most recent revision of the poverty counts, Chen and Ravallion

(2000), resulted in some disconcertingly large changes in poverty counts even for the same country in the same year and using the same survey data. The worst cases are Sub-Saharan Africa, where the poverty rate in 1993 (that is, the same year) increased from 39.1 to 49.7 percent; Latin America, with the poverty rate falling from 23.5 to 15.3 percent; and the Middle East and North Africa, where the poverty rate fell from 4.1 to 1.9 percent. Changes of this size risk swamping real changes, and it seems impossible to make statements about changes in world poverty when the ground underneath one's feet is changing in this way. Note that Chen and Ravallion (2000) recalculate previous counts at the new PPPs, so the issue is *not* that we are comparing current poverty using one PPP conversion with earlier poverty using another conversion. The problem is that, if revisions can have such enormous effects, it is difficult to place much confidence in any set of poverty estimates, either of levels or of changes, even those based on the latest and best set of PPP exchange rates.

For countries whose poverty lines are near the middle of the distribution of consumption or income, poverty rates can be extremely sensitive to changes in the poverty line, including those induced by PPP (or other) errors. World Bank (1997) lists Thailand as having only 0.1 percent of its population living on less than $1/day at PPP. This virtual elimination of poverty was cited in the *New York Times* by then Chief Economist Joseph Stiglitz as one of the consequences of the Asian economic miracle (Stiglitz 1997), but it is much more likely a tribute to inappropriate PPP conversion. This anecdote suggests that, at a minimum, it would be wise to check the PPP poverty lines for plausibility in the countries to which they apply. Of course, if the lines are wildly implausible, we need some other method.

There are two conceptually different kinds of revisions to PPP exchange rates. First, there are the usual revisions that come with better information, elimination of previous errors, and so on. For example, it is claimed that for China, the new updated PPP rates are much more solidly based than were the old rates. In addition, the PPP exchange rates are now calculated, rather than imputed, for 50 more countries. Second, there are the revisions that come from the change in the base, given that PPP exchange rates depend on the relative prices of commodities in the base year. This second source of revisions ultimately traces back to the conceptual or "index number" problems associated with making international comparisons of well-being. Indeed, one might argue that the $1/day poverty line, by its very simplicity, is misleading just because it seems to sweep away the problems.

The calculation of the cost of subsistence for a poor Indian versus that for a poor Kenyan or Brazilian is fraught with conceptual difficulties. For example, we could calculate the relative cost of living of an Indian living in Calcutta compared with the (hard-to-imagine) cost of living if that person was moved to São Paulo. We would likely get a very different number from doing the same with a Brazilian transplanted to Calcutta. Suppose that, in country A, the staple food for poor people is beans and in country B it is rice. If we price beans in country B in B's currency and divide by the

bean price in country A in A's currency, we get an implicit "beans" PPP exchange rate. A similar calculation for rice gives a "rice" PPP exchange rate. These will give the same answer (or at least reciprocals of one another) only if both goods are freely traded between the two countries without tariffs, barriers, or transportation costs. These conditions are far from holding true in practice, even approximately.

Many (perhaps most) of the world's poor eat food staples that are grown only a short distance away and whose price is only loosely (if at all) connected to world commodity prices. And the staple food in one country (or even region) is often a relatively rare (and even expensive) luxury somewhere else (for example, sticky rice in Northern Thailand versus the United States). Therefore, the exchange rate obtained from the prices of the same commodity is not very useful for converting poverty lines from one country to another.

In practice, matters are not quite so stark. Actual PPPs price a representative bundle of goods in each country and compare the local cost of the bundle with the U.S. dollar cost of the same bundle. But note that the answer depends on the structure of *relative* prices at the time of the comparison, so that, when PPPs are revised with a new base year, the new exchange rates are not simply a new measure of the *old* concept but a new measure of a *new* concept. World prices of primary commodities are notoriously volatile, and for some countries, primary commodities make up a large share of gross domestic product (GDP). As a result, the PPP exchange rate for Nigeria (for example) relative to the United States will vary with the world price of oil in the base year. This is presumably why the African and Latin American poverty rates are so sensitive to revisions of the PPP base year. Although it might be argued that changes in the world prices of beans and rice ought to change the relative poverty lines of Brazilians and Indians, it is much harder to make the case for changes in the world price of oil. Even in theory, PPP exchange rates as currently defined are not designed to convert poverty bundles (see also Srinivasan 2000).

One way of improving the PPP poverty lines would be as follows. First best would be an international comparison program that focuses not on GDP and its components but on some more appropriate poverty bundle. In this case, the first best is perhaps not very good, if only because even a poverty-bundle PPP would not address the underlying index number issues. A second-best shortcut would be to take the current set of 1993 poverty lines, update them to the present, and check them in each country (or at least those countries that have significant numbers of poor people). In cases where the numbers are implausible (for example, showing no poverty in Thailand), the PPP exchange rate can be calculated using the first-best procedure. The calculation would price out in dollars and in local currency an appropriate poverty bundle, for example the average consumption bundle of the bottom quintile of the population. Although the choice of the bottom quintile is arbitrary (but sensible), the calculations are straightforward given a household expenditure survey, which often collects data on quantities as well as on expenditures. When the survey does not

contain quantity data, they can be calculated from the expenditure data using a set of local consumer price quotes, such as those that are routinely used for the construction of national consumer price indexes.

Armed with a set of PPP poverty lines that have been locally validated and adjusted as necessary, the crucial next step is to hold these fixed (in real terms) and not make further adjustments as PPP exchange rates are revised in response to changes in the base year and its associated relative prices. The benefits of this proposal are that African and Latin American poverty rates are not held hostage to fluctuations in the prices of primary commodities in the base year of PPP calculation, and that world poverty rates are consistently calculated against a fixed target. Of course, it could be argued that the line is no longer the constant PPP $1/day line, and that the World Bank is failing to update for new information. But this objection is more semantic than real. The lines originally started at $1/day and can retain the label. They are updated by local prices, just not for changes in world relative prices. The proposal effectively changes the *definition* of PPP exchange rates, whose shifting basis is an embarrassment, not an advantage.

Even if the lines, once set, are to some extent arbitrary, and even if it were the case that a reworking from scratch would lead to new lines, there is much to be said for holding them fixed. Indeed, it is hard not to be impressed by the durability of official poverty lines in countries as diverse as India and the United States, whose poverty lines have survived unchanged (except for inflation correction) for 30 and 40 years, respectively. Once an initial poverty line has been set, and provided it commands public and political support at the time it is set, it seems that it can generate useful estimates of poverty for many years. Moreover, and for both India and the United States, the original nutritional basis for the lines is often cited in their support. The same could be true for the $1/day lines.

International Comparability without Purchasing Power Parity

There are a number of alternatives to the $1/day PPP poverty line, although none is without problems. There is a tradition, followed in both the United States and India, of setting poverty lines with reference to what Ravallion (1998) refers to as "the nutritional requirements for good health." This is defined as the level of income (or total consumption) at which, on average, nutritional norms are met. Note that, because the standard is set in terms of income or total consumption, this is *not* the amount necessary to purchase the nutritional norms only. It is the amount spent on the nutritional norms and on other goods by an average household that just meets the norms. An internationally comparable set of poverty lines could be set in this way with reference to a universal calorie norm, say 2,000 calories per person per day. A national household expenditure survey for each country would be the basis of the

calculation of the level of total household expenditure per capita at which this norm is met in expectation.

Again, this method is not quite what it seems, and there are important practical qualifications. First, as with the recommendations for the amended PPP method, the procedure is used once and once only. The resulting poverty lines for each country are subsequently held fixed over time. This may seem contradictory; if the calorie method is correct, it should surely be applied consistently over time and space. But it is not correct in the sense of being a "scientifically sound" method of setting the line. Instead, it is a device for generating a socially acceptable poverty line that can then be held fixed over time. The food rhetoric helps build legitimacy for the line, but, judging from India and the United States, maintaining that legitimacy does not require updating, except possibly in the very long term. Successive commissions and expert groups have altered details of the Indian lines, for example, by introducing variation across states (Government of India 1993). In the United States, there have been frequent (sometimes detailed) discussions of reform, particularly National Research Council (1995), but none of the recommendations have been adopted.

Second, it is important *not* to set separate nutritional poverty lines for different regions or sectors within a country. At the same level of per capita total household expenditure, urban people spend less on food, buy more expensive calories, and consume fewer of them. Therefore, applying the method to urban and rural areas using the same calorie target will lead to higher poverty lines in urban than rural areas. Across regions with different income levels, the operation of Engel's Law has much the same effect, so that poverty lines can move more or less in proportion to average incomes, generating effectively relative, not absolute lines (Ravallion and Bidani 1994; Ravallion and Sen 1996). Although relative lines may make a good deal of sense in other contexts, they are not appropriate for world counts that are based on an explicitly absolute standard.

Domestic Issues in Counting the Poor

Even if there existed perfectly accurate and perfectly relevant international exchange rates, any problems with measuring domestic poverty will carry through to the world totals. There are several such problems.

National Accounts and Household Survey Data

A major source of controversy is the discrepancy between national accounts and survey estimates of consumption, especially when, as in India, the difference is increasing. India is the largest single contributor to the world poverty count, with more than 400 million poor people out of a world total of 1.2 billion. The ratio of National

Sample Survey (NSS) consumption to national accounts statistics (NAS) consumption has fallen from around unity in the 1950s, to 89 percent in 1968–69, to around 75 percent in the late 1970s, to a little more than 50 percent in recent years (see Bardhan 1974; Mukherjee and Chaterjee 1974; Srinivasan and others 1974 for early discussions; and Srinivasan 2000 for a recent update). Although different authors give somewhat different ratios, there is general agreement that the discrepancy is widening and that the difference in per capita growth rates is currently about 2 percent a year. As a result, the rapid postreform growth in GDP and consumption per head measured in the NAS does not show up in the NSS surveys. The latter are now the exclusive basis for the official poverty counts, and they show little or no reduction in official poverty during the reform period.

The Indian example is only the most notable of similar or related phenomena elsewhere, including some industrial countries, such as the United States. Triplett (1997) estimates for the United States that, even after correcting for differences in concept, between 1973 and 1994, per capita consumption in the national accounts has grown 0.4 percent a year faster than per capita consumption in the Consumer Expenditure Survey (CEX). Between 1984 and 1994, the difference in growth rates was a full 1 percent a year. (Note that the CEX data are not used to derive the national poverty counts in the United States.)

Ravallion (2000b) notes a similar drift in other countries and regions, particularly in China (the second-largest single contributor to world poverty) as well as in Latin America, where survey consumption growth is only 70 percent of NAS consumption growth. In Sub-Saharan Africa, consumption growth in the surveys is less well linked to the NAS but is also consistently slower. Although there is no drift in the transition economies of Eastern Europe and Central Asia, it is only because there appears to be no relationship whatever between consumption growth in the NAS and in the surveys.

Ignoring the transition economies, Ravallion cannot reject the hypothesis that the coefficient of the growth rate of NAS consumption on survey consumption is unity in a regression on a cross-section of countries. However, as is appropriate for his purpose, Ravallion's units are countries, not people, and most of the poor people in the world live in countries or regions (East Asia, India, and Sub-Saharan Africa) where survey growth rates of consumption are substantially less than national accounts growth rates. Even when the coefficient is insignificantly different from one, the point estimate is 0.836, indicating substantial slippage, and is lower still in East Asia (0.628), Sub-Saharan Africa (0.645), and South Asia (0.742).

That the drift is so common around the world and that it is generally in the same direction suggest a common cause; clearly, it would be an important research project to try to track it down. For example, does the positive skewness in consumption, together with the plutocratic basis of the NAS versus the democratic basis of the surveys, result in a tendency for average consumption in the surveys to lie beneath the

population mean? Is average consumption from the surveys even lower when inequality is increasing? Some limited calculations show this to be unlikely for reasonable parameters. There are other plausible links between increasing inequality and increasing bias. For example, the rich are more likely to be missed in household surveys—something that is widely suspected in India and has been well documented in the United States (Groves and Couper 1998).

Because growth around the world is almost universally assessed from the NAS, and because the poverty counts come entirely from the survey data, the discrepancies between the two drive a wedge between measured growth and measured poverty reduction. If we were to measure both growth and poverty from the household survey data, the discrepancy would vanish; there has been little or no measured growth, and little or no measured poverty reduction. Similarly, if we were to follow Bhalla (2000a, 2000b) and treat the national accounts estimate of mean consumption as the truth, using it to "correct" the survey data, the discrepancy would also vanish but in the other direction; there would be lots of growth and lots of poverty reduction. Meanwhile, claiming that growth has done little to reduce poverty is comparable to saying that, in spite of rapid growth in China, poverty in India remained the same. The NAS and the surveys evidently measure different things.

Measuring consumption. Which is correct? We don't know, although it seems safe to say that there are almost certainly errors in both the NAS and the surveys. There is a longstanding prejudice by many economists (although not in India) against surveys and in favor of national accounts that is probably without basis. However, the following points are worth noting.

First, the two concepts are different. In particular, there are items in NAS consumption, perhaps most important consumption of nonprofit organizations and the imputed rental of owner-occupied dwellings, which are not included in the surveys. In India, it is thought that as much as half of the current discrepancy between the two estimates can be accounted for by implicit rents. It is also plausible that these rents are growing more rapidly than other items of consumption, thus contributing to the differential rates of growth.

Second, although expenditure estimates from household surveys are sometimes used in constructing NAS consumption for some commodities, most items are derived as residuals, so that errors and omissions elsewhere in the accounts are automatically absorbed into consumption. A major problem is consumption by businesses, which has to be estimated and subtracted from the production totals to get the consumption of households. Consumption of coarse grains by animals presents a similar problem and is often estimated using a set of multipliers that are difficult to estimate with any precision and that are infrequently updated. It is not clear that the NAS estimates of consumption should be treated as the gold standard to which the survey estimates should ideally correspond. Indeed, Bardhan (1974:117) writes, "it

would be highly improper to judge the goodness of NSS estimates by pointing to their divergence (or lack of it) from such residual estimates."

Third, there are many well-known problems with data collection from household surveys. Best-practice consumption measures (as in India or Indonesia) use long lists of specific items. The evidence suggests that attempts to shortcut this process lead to underestimation of consumption (Deaton and Grosh 2000). Yet many surveys use such questionnaires, which tend to change slowly over time and may not keep up with the increase in consumption of goods and services that may have been unknown a decade or two ago.

Fourth, consumption surveys miss some households through refusals. Surveys in developing countries typically report low refusal rates (for example, 0 in the Indian NSS) and do not document the extent to which willing respondents have been substituted for unwilling ones. There is some evidence that wealthy households are most likely to refuse to cooperate; for example, it is often impossible for enumerators to gain access to gated communities. If so, it might be conjectured that real income growth, by increasing the fraction of such people, will lead to an increasing underestimation of consumption. Against this, it might be argued that the very poor, especially those without fixed abode, are also missed in household surveys.

Fifth, in most surveys, different kinds of households have different probabilities of being included, either by design or because of problems in the field, as when some types of household are more prone to refuse their cooperation. As a result, the survey results must be weighted to give an accurate representation of the population as a whole, and the calculation of suitable weights depends on the availability of accurate, up-to-date information about the population. Although statistical agencies have a number of tools for ensuring accuracy, it would not be surprising if the weights were to become progressively more inaccurate as the latest census becomes outdated. That said, the divergence between survey and NAS estimates in India shows no evidence of the saw-tooth pattern—growing between censuses and then falling immediately after each census—that would be predicted by such an account.

Sixth, it is tempting to try to use the behavior of other measures of poverty, such as literacy or infant mortality, to judge the relative plausibility of the NAS on one hand and the surveys on the other. There is undoubtedly good research to be done along these lines. Even so, definitive results should not be expected. Health and literacy measures should not be regarded as proxies for income, if only because they frequently change in ways that have little or nothing to do with income. We do not have a reliable link between income and nonincome measures of poverty that would allow us to infer the former from the latter.

Much the same can be said of attempts to look at indicators like the share of the budget devoted to food. Although it is true that the food share drops with increases in income, it is also influenced by other factors, including relative prices, demographic changes, and the distribution of income, to mention only the most obvious. As a re-

sult, factors (such as an increase in inequality) that are possible suspects in driving the discrepancy between the national accounts and the surveys are also likely to affect the relationship between food and total expenditure. Therefore it is difficult to know how much weight to attach to calculations such as that in Bhalla (2000a) that impute the growth in total expenditure from data on food availability. Perhaps more promising in practice, if not in theory, are the measures of living standards derived by Filmer and Pritchett (1998) from principal components analyses of the various asset ownership measures that are included in Demographic and Health Surveys.

What should be done? It is, of course, possible for poverty to remain constant or to increase in the face of positive growth in average consumption if people in the lower tail of the distribution are left behind, and if most or all of the growth accrues to people above the poverty line. That growth should disproportionately favor the better off is entirely plausible in some settings. For example, in India, the sources of growth have been outside of the agricultural sector where most poor people make their living. But differences of definition apart, the NAS and survey estimates of consumption cannot both be right because their means are different and are systematically becoming more different over time. Whatever the role of increasing inequality in moderating the poverty-reducing effects of growth, one of the reasons why measured growth has not reduced measured poverty is the growing statistical discrepancy between the two sets of data on which each is measured.

What to do? There is a school of thought in India among the pro-reformers, as in Bhalla (2000a), which argues that the NSS consumption figures should be scaled up so as to match the NAS, and that the rescaled survey estimates should be used to calculate poverty. (Indeed, this is how the official poverty counts used to be done in India, and no very convincing reason was ever given for the change following the recommendations in Government of India [1993].) Bhalla (2000b) argues that the "means from NAS, distribution from surveys" formula should be adopted for the world poverty accounts and calculates that growth has indeed reduced poverty. But there is a serious lack of evidence for at least some of the assumptions that would validate such a calculation. In particular, the following must be true: (a) the NAS estimates are correct; (b) the survey estimates of the mean are incorrect; and (c), in spite of b, the consumption levels of each household in the surveys are correct up to a multiplicative factor. The last condition is a real stretch.

Most of the plausible accounts of what might have gone wrong with the surveys, such as a progressively larger number of refusals, involve a systematic underestimation of inequality as well as of the mean. It seems particularly implausible that the degree of underestimation in the surveys should be the same for the urban and rural households (see Ravallion 2000a). Indeed, many people in India, particularly those skeptical of the benefits of the reforms, believe that the growth that has taken place in consumption has accrued to people well above the poverty line and to urban more

than rural people. They believe that the people who have most benefited are the most likely to be undersampled by the surveys. Moreover, the consumption of the poor—who do not own valuable housing—is unlikely to be understated by the omission of owner-occupied rents. If these arguments are correct, one might suppose that the NAS modestly overestimates the rate of growth of consumption, that the NSS greatly underestimates it by missing the wealthy, and that the NSS-based poverty figures are essentially correct.

A good statistical principle is that, if two sources of data disagree and we have no reason to favor one over the other, then we should combine them to make a better estimate. This principle would argue for a more modest version of the scaling proposal, in which the survey data are scaled up by some weighted average of the NAS and the survey means, at least after correcting for conceptual differences and coverage. Of course, the data on economic growth would also have to be scaled down accordingly. This proposal would make neither side happy, each of which has its reasons for believing that one set of accounts is biased and the other bias-free, but I can see no very convincing argument against it. Its limits are obvious; it takes no account of reasonable suspicions that the NSS errors are not neutral between rich and poor or between urban and rural. But in the absence of the sort of better information that one might hope would be produced in the longer term, it is surely better than either of the extreme alternatives that currently dominate the debate.

The urgent need is to start a serious program of reconciliation between the NAS and survey data in a few countries, of which India must surely be one. (The Bureau of Labor Statistics and the Bureau of Economic Analysis at Census are talking about setting up a similar program in the United States. Perhaps common cause could be made.) This program must work both with the national accounts and with the survey organizations and must not begin by assuming that one is right and the other wrong. It might be productive to start with a particular commodity or group of commodities, such as cereals, where the production, export, and import data are relatively strong and where there is a great deal of experience with household surveys. It would be possible to mount special supplementary household surveys for the task, which could focus on only a few goods. But it would also be a good idea to audit the performance of the survey agencies, just as the NAS agency is being audited. For example, the Indian NSS used to lead the world in household survey methodology and practice. But that was a long time ago; the organization is no longer as dynamic and open to new practice as it once was. There has been a great deal of progress in survey methodology in the last 20 years, with the result that those statistical offices that used to be the leaders, mostly in South and East Asia, are now the most likely to be behind.

The World Bank is the only international body that has the capacity and organization to run a research program designed to resolve the discrepancies between national accounts and survey data. To its credit, the World Bank has played a major

role, through the Living Standards Measurement Study (LSMS) surveys, in collecting household survey data on a nearly worldwide basis. The next step is to find out why these surveys give such different results from the national accounts. Until it does so, neither the Bank nor anyone else will have a firm platform on which to make statements about the central issue of how growth affects poverty.

Price Indexes

Local consumer price indexes are used to update a real poverty line for comparison with the nominal expenditures or incomes collected in the surveys. Consumer price indexes are of variable quality around the world and poverty counts can be very sensitive to any errors. Consumer price indexes are constructed from two components: a set of prices, collected on a regular schedule from retail shops and markets around the country, and a set of weights, which typically come from a household expenditure survey.

Problems can arise with both components. For example, some countries collect good data in urban centers, but find it much more difficult to collect adequate price data in the countryside, where many poor people live. Urban bias may result in prices that are more relevant for relatively affluent urban workers than for the poor whose poverty line we are trying to calculate. There are also perennial questions about the representativeness of the markets and whether the enumerators observe the prices that people actually pay. In India, the villages where prices are collected are updated infrequently, if at all.

Weights can also be troublesome, especially if not updated sufficiently often. When it was revised in late 1995, the Indian Consumer Price Index for Agricultural Laborers (CPIAL) had weights that were 35 years old. It is usually thought that Laspeyres indexes whose weights are held fixed for too long are increasingly likely to overstate the rate of inflation and, thus, in this context, to cause an underestimation of the rate of poverty reduction. (Note, however, that price indexes are *not* at the root of the discrepancy between NAS and survey data; the divergence is in nominal terms.) When updating its poverty lines, the Indian Planning Commission reweights the components of the CPIAL so as to match more closely the purchases of people near the poverty line, using weights from 1973–74, which are still elderly by international standards. (The World Bank, in its $1/day poverty calculations, uses a slightly modified version of the CPIAL.) Deaton and Tarozzi (2000) have made independent estimates of price indexes in India, which suggest that even the re-weighted CPIAL has been rising too quickly. Their alternative estimates show some decline in poverty in the 1990s, although much less than would be the case if the NSS data were scaled up to agree with the NAS data on consumption. I suspect that India is much better provided with good price indexes than are other countries. Indeed, much of the recent debate on the effects of the East Asian crisis on poverty

in Indonesia has foundered on the unavailability of good price indexes for rural areas (see Thomas and others 1999).

Local price indexes are used not only for updating over time but also for adjusting poverty lines for urban-rural price differences, as well as for regional differences in prices. In the World Bank's calculations of the number of poor in the world, separate urban and rural indexes are used only for India and China. In other countries, a single index does service for everyone, an expedient that must overstate rural relative to urban poverty.

For many years, India had urban and rural poverty lines of Rs 57 and Rs 49, respectively, at 1973–74 prices, a differential of 16 percent. Both lines were updated by the implicit price deflator of consumers' expenditure in the national accounts. Since 1993, the planning commission has adopted different urban and rural poverty lines for each state. The lines are designed to reflect different prices in different parts of the country, and each line is updated by a poverty-reweighted urban or rural price index specific to that state. For a number of reasons, the implicit urban-rural price difference in the current lines, a difference that was never explicitly measured, has risen to around 40 percent on average, with larger or smaller differentials in individual states.

This large urban-rural price difference, for which Deaton and Tarozzi (2000) find no evidence in prices actually paid, yields urban poverty rates that are larger than rural poverty rates in many states. This finding seems implausible to most observers and sits ill with the evidence on other poverty-related indicators. Deaton and Tarozzi recalculate the rates based on their price indexes, with differentials quite close to the once standard 16 percent, and find that urban poverty is reduced from 32.4 to 18.1 percent in 1993–94, which reduces the total number of Indians in poverty by 23 million people. Thus price differences within a country can have major effects on the poverty counts both nationally as well as in their contribution to the world totals.

Problems with Household Surveys

The World Bank is to be commended for building so much primary survey data into its poverty counts. This is a great advance over the situation 20 years ago, and the Bank has been the prime mover in bringing these data to bear on the measurement of poverty. But the surveys have a number of limitations for poverty work that need to be kept in mind, some of which might be ameliorated by more work or different approaches. Here I discuss three of the most important limitations: coverage, reference periods, and income versus consumption.

Coverage. Household income and expenditure surveys are good at collecting data on cash that passes through the household for both incomes and outlays. They are some-

what less good (but probably still good enough) at collecting data on own production and consumption. However, they do not attempt to include in the consumption or income estimates the benefits that people receive from publicly provided goods of one kind or another. Although it would be difficult to collect such information, some relevant data are collected in the surveys (for example, visits to clinics and school attendance) and perhaps more might be.

Problems develop if the poverty counts ignore this information. In particular, consider the debates about structural adjustment and the (widely credited) arguments by nongovernmental organizations and others that the World Bank has impoverished the poor by forcing cuts in public expenditure. It is not very useful, as the Bank sometimes does, to argue that there was no impoverishment because the poverty counts (which take no account of the provision of public goods) are going down. Any program that eliminated government services (closing clinics and schools) and shared the money among the population would reduce the poverty count as the World Bank currently makes it. The Bank has a good deal of experience in benefit incidence analysis, that is, the use of household surveys to allocate the benefits of public expenditures. It would be worth giving serious consideration to adding some of these estimated benefits to the consumption totals before making the poverty counts.

Reference periods. Different household surveys use different recall periods, even for the same types of goods. Even within the LSMS surveys, there is some variation. Given a true flow of expenditures, the rate of expenditure that is reported is sensitive to the reporting period; longer reporting periods lead to more forgetting, at least for common purchases like food. Again, India provides a good example. Based on experiments done long ago by Mahalanobis and Sen (1954), whose results were quite contrary to later findings in the literature, the Indian NSS has always used a 30-day recall period for food (see Deaton and Grosh 2000). Seven or 14 days is a more common standard in modern surveys, and Scott and Amenuvegbe (1990) have argued that estimates are biased downward after even one day. In the last several years, the NSS has experimented by using a 7-day reference period for half the households and a 30-day period for the other half. There were also changes for goods, such as durables and clothing, in which a 365-day reference period was used, again compared with a 30-day period in the standard survey.

Visaria (2000) shows that the shorter recall period causes more (food) expenditures to be reported. Using the official poverty lines and the new data, measured rural poverty in the first six months of 1998 falls from 42.6 to 23.6 percent, and urban poverty from 32.9 to 20 percent, removing about 175 million people from poverty. (The effect would be somewhat larger for the $1/day international line, which is higher than the Indian official poverty line.) Of course, Visaria's calculations offer no apparent help on the issue of differences in trend between the NAS and the NSS versions of consumption. (Although it is possible that the relatively underreported foods

are those that have been growing more rapidly, preliminary analysis of the data does not support the conjecture.) Somewhat surprisingly, the estimates based on the experimental questionnaire (7-day, 30-day, and 365-day recall, as opposed to 30 days for all goods) show *less* variance across households than do the estimates from the traditional questionnaire (Government of India 2000; Deaton 2000). Thus there is no evidence of people being incorrectly characterized as poor simply because they happened to make food purchases outside the shorter recall period.

One question about Visaria's calculations is whether it is legitimate to use the old poverty line with the new reporting periods and the greater food consumption that they reveal. The answer depends on where one thinks the poverty line comes from and the purpose it serves. Because the Indian line was originally set by the nutritional needs method, perhaps the obvious thing to do is to try to recalculate what the poverty line would have been had the original method been followed using data with the shorter reference period. This is not the only possible treatment. Instead, one might argue that the legitimacy and longevity of the line owes more to the general perception that Rs 57 per head in 1973–74 prices is a reasonable poverty line in reference to people's experiences of earning and spending and their notions of what constitutes poverty. If so, there is no reason to revise it in the light of the new survey data, and Visaria's calculations are correct as they stand. Alternatively, one might argue that the line is acceptable only because people accept the headcount ratio that it generates, in which case the survey data are essentially irrelevant.

If we follow the calorie approach, it turns out that the poverty lines should be revised, not upward as intuition might suggest, but *downward*, so that Visaria's calculations, if anything, *overstate* the rate of poverty. The argument is as follows. Pick a level of household per capita total expenditure and consider the average food expenditure over all households at that level, under the new and old survey methodologies. Because the new methodology recovers food expenditures that were previously forgotten, food expenditures rise on average. But total expenditure is both defined and measured as the sum of food and nonfood expenditures, so that total expenditure will rise by the same amount as food expenditure. (For the purpose of the argument, ignore the change in survey procedures for nonfood.)

Figure 1 illustrates. The old (30-day reporting period) Engel curve is the lower curve. For an average household (conditional on expenditure) at point A, food and total expenditure (or per capita) increase by equal amounts along the 45-degree line to B, say. Because the slope of the curve is less than 45 degrees—the marginal propensity to spend on food is less than one—the switch to the seven-day period moves the Engel curve outward to the higher curve. The constant level of nutritional needs, represented by point N, thus requires a shift downward in the poverty line from the old P to the new, lower P*. With the new, better reporting period, more food is reported, so that if we put ourselves back in the shoes of the planning

Figure 1. The Effects of a Shorter Reporting Period on a Nutritional Poverty Line
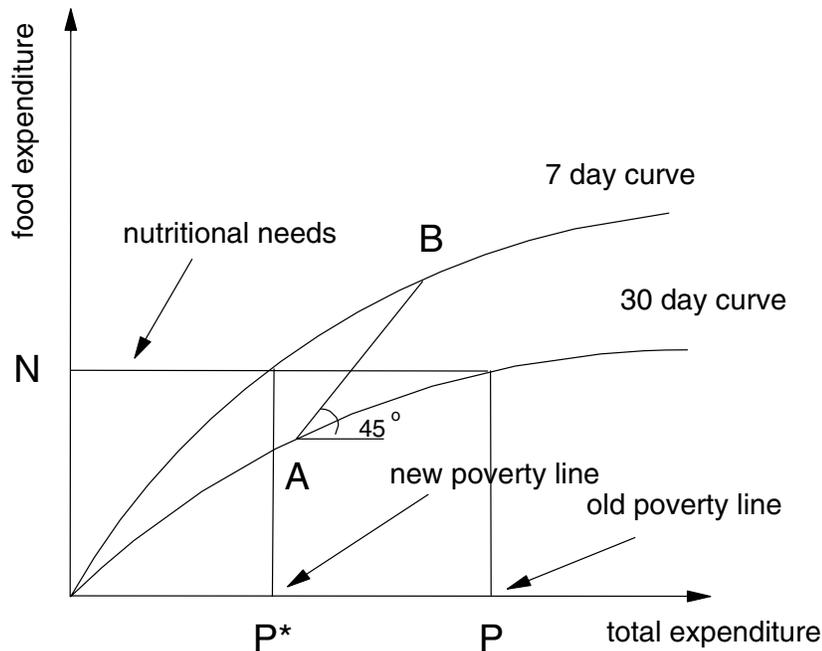
commission when the poverty line was first calculated, the original poverty line would have generated more calories than the norm, and would therefore have been revised down.

In Deaton (2000), I present evidence on the relevance of figure 1 using data from the 52nd Round of the NSS gathered during 1995–96. The Engel curves relating per capita expenditure on food to total household expenditure per capita are in the order predicted in figure 1, with the seven-day Engel curve lying on or above the 30-day Engel curve. However, the two curves essentially coincide below and up to the poverty line so that, unless the ratio of calories to food expenditure is different in the two questionnaires, there is no reason to revise the original poverty line to meet the new questionnaire design.

In summary, Visaria's reduced poverty counts have as much (or more, judged by international survey practice) legitimacy as the official counts, or indeed the $1/day calculations. To get some idea of the magnitudes involved, the change in the survey reporting period reduces measured Indian poverty by about as many people as the total number of poor in China. Clearly, there is a large margin of error in the world poverty estimate of 1.2 billion people. In the long run, household surveys should be

brought into reasonable conformity in their reporting periods. In the short run, the counts should not simply add together counts from surveys with different reporting periods without some sort of correction. This will involve sometimes admittedly arbitrary choices, and perhaps the removal altogether of some surveys. But that is better than adding together incommensurate numbers.

Income versus consumption. The surveys used for the world poverty count sometimes collect data on income and sometimes on consumption. The latter is the preferred measure (rightly, in my view), so that the surveys that collect income (38 out of the 91 in Chen and Ravallion [2000], including all the Latin American and Caribbean surveys) have to be "converted" to a consumption basis. This is done by scaling down the income data in the survey by the ratio of national consumption to national income (presumably personal disposable income) in the national accounts. This is plausible, but in my judgment, it is unlikely to lead to good results.

First, there is some inconsistency in using the national accounts data in this situation, but not when there is a full consumption survey, even when the NAS and survey data are mutually inconsistent. The argument is presumably that there is little choice. Even so, it would seem more logical to use some average of NAS and survey data to correct the expenditure survey when available or to make some adjustment to the NAS estimate of the consumption ratio in the case when it is not.

Second, in countries where household surveys collect data on both income and consumption, there is usually little or no evidence of positive saving among poorer households; indeed, the reverse is true. (China seems to be an exception to this generalization.) Even in industrial countries, survey saving rates are frequently negative until well up the income distribution. A frequent argument is that survey respondents understate income relative to consumption. If so, multiplying incomes by a number that is less than one is a "correction" in the wrong direction. More generally, whatever the accuracy of the consumption and income reports themselves, it remains true that, in most countries, multiplying survey income by the national consumption rate would do a very poor job of replicating survey consumption, certainly at the bottom of the income distribution.

Likewise, a proportional correction for saving is not likely to convert income data from a survey into something that looks like consumption data from a survey. In the usual situation, when the poverty line is below the mode of a unimodal distribution, the resulting underestimation of consumption by the poor (relative to a consumption survey) will tend to overstate poverty rates. Countries with income surveys probably tend to be assigned higher poverty rates relative to those with consumption surveys. This contention could be tested for a few countries where surveys collect both kinds of data, or by running a regression across the countries in the database and seeing whether those with income surveys have higher poverty, controlling for GDP and other variables.

Unfortunately, the degree of heterogeneity across countries is unlikely to permit any simple correction based on the results of such a regression.

Third, measured consumption is famously less variable than measured income. This is true in logarithms as well as in levels (the variance of the logarithm of consumption is less than the variance of the logarithm of income). Therefore, even if the proportional saving assumption were right on average, the corrected income data would still be too variable relative to measured consumption. Once again, and provided the poverty line is below the mode, consumption poverty will be overestimated by income poverty.

My guess is that it would not be too difficult to work out some pattern of corrections that, although hard to defend absolutely, would likely still be better than the current procedures. In the long term, the battle for expenditure surveys will be won. Even the holdouts in Latin America are moving in that direction.

Summary, Recommendations, and Directions for Research

The World Bank's worldwide count of the poor starts from a common international poverty line and counts the number of people in each country whose consumption lies below it. The international poverty line, at $1 or $2 a day, is converted into domestic currencies using PPP exchange rates. I argue that, although there is much to recommend the $1/day line, its dependence on PPP exchange rates has a number of unfortunate consequences.

An arguably better procedure for the future would be to hold fixed (in real terms) the current domestic poverty lines, and not to revise them along with changes in PPP exchange rates induced by updating base years. Such a program would work best if the current rates were subjected to detailed, local scrutiny and corrected in a way that would simultaneously give them local credence without major deviation from the $1/day standard. Such corrections will likely be difficult. They must not be interpreted as a license for individual countries to define their poverty lines with a view to their international political effects. Corrections will be easier if, as I suspect, most anomalies are the result of inappropriate PPP exchange rates, perhaps because interpolation from other countries gives a poor answer or because the bundle of goods being priced is inappropriate for the poor.

Other problems relate less to international comparability than to the calculation of poverty within each country. The proliferation of household surveys, for which the World Bank should take much credit, has greatly enriched the basis for good poverty calculations around the world. But the very availability of expenditure surveys has highlighted an old problem that seems to have recently become much worse, which is a divergence between estimates of average consumption based on surveys

and on national income accounts. The World Bank is probably the only organization that is capable of leading an international research initiative to address these discrepancies. Until it does so, policymakers and commentators will be invited to choose their evidence according to their political predilections. Furthermore, neither the World Bank nor anyone else will be able to make well-supported statements about what is surely a central issue in economic development: whether growth has reduced poverty in the past or is likely to do so in the future.

There is also much scope for the improvement of survey practice and the convergence of household expenditure surveys toward international best-practice standards. The experimentation with recall periods by the Indian National Sample Survey, although it has sown some confusion in the short run, is an excellent example of the sort of methodological work that is needed and which statistical offices are usually reluctant to perform. If nothing else, these experiments have highlighted the extreme sensitivity (175 million people less in poverty, the headcount ratio reduced by half) of poverty measurement to what many would previously have regarded as a technical, if not downright esoteric, issue of survey design.

Here, too, is an important research program. We need to understand much better than we do now how the measurement of consumption is affected by the design of consumer expenditure surveys—not only reference periods but such issues as the detail of the questionnaire, the need for bounding interviews and repeat visits, multiple versus single respondents, whether ownership information on durables is worth collecting, and so forth. Although the World Bank's LSMS has done much to normalize the use of enriched expenditure surveys in many countries, it has done little experimentation and has not greatly contributed to settling design questions (see Deaton and Grosh 2000).

Another important research issue is finding better ways to set the poverty line. Within the standard realm of income (consumption) poverty, there is much to be said for giving greater weight to people's own reports of consumption adequacy. Ravallion has done sterling work in this direction and there is widespread interest around the world in these methodologies (Ravallion 1998; Pradhan and Ravallion 2000). In the United States, for example, the Gallup poll questions people on how much it would take for a family like theirs to "just get by" in their community. The responses yield sensible poverty lines, at least once outliers have been removed. People can also be asked to self-assess their poverty, as in Ravallion's work. If such schemes could be put on a sound footing, they raise the tantalizing prospect of reducing our dependence on extremely expensive, time-consuming, and most likely inaccurate consumption surveys. It is a good deal cheaper to ask one or two questions than 200 or 300.

Recent work in South Africa shows strong correlation between measures of financial, physical, and mental health (Case and Wilson 2000). But there are real concerns about "adaptation." People do not perceive themselves to be better off, even in the face of large increases in real income, essentially because they adapt their expec-

tations to their circumstances. For example, Easterlin's (1995) famous finding shows that subjective levels of happiness did not increase among Japanese consumers between 1958 and 1987 in spite of a fivefold increase in real per capita income. The World Bank might find itself successfully promoting growth with a concomitant reduction in income poverty, but then being negatively assessed because people do not report themselves to be better off. The Washington Consensus derailed by the hedonic treadmill!

Those of us who have been exposed to field experience have been impressed by the prominence of health concerns in what people tell you about their poverty. Income, housing, and jobs tend to predominate when health is normal, but if someone gets sick, is hit by a car, or has a friend or relative who has been raped or murdered, income poverty recedes into the background in people's perceptions. Many millions of people around the world will die from AIDS, with untold misery and deprivation. It would be a terrible thing if the World Bank dealt with this crisis only in terms of its effects on income poverty, which it already shows some signs of doing (PovertyNet 2000).

My view is that the World Bank should back away from its current too-concentrated focus on income headcount numbers. It should emphasize a much wider range of other measures, focusing on deprivations that may be more important than deprivation of income.

Note

References

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Bardhan, Pranab K. 1974. "The Pattern of Income Distribution in India: A Review." In T. N. Srinivasan and P.K. Bardhan. eds., *Poverty and Income Distribution in India*. Calcutta: Statistical Publishing Society, 103–38.

Bhalla, Surjit S. 2000a. "Growth and Poverty in India: Myth and Reality." New Delhi (June). Processed.

———. 2000b. "Trends in World Poverty: Ideology and Research." New Delhi (June). Processed.

Case, Anne, and Francis Wilson. 2000. "Health and Wellbeing in South Africa: Evidence from the Langeberg Survey." Research Program in Development Studies, Princeton, N.J. (December). Processed.

Chen, Shaohua, and Martin Ravallion. 2000. "How Did the World's Poorest Fare in the 1990s?" World Bank, Washington, D.C. Processed.

Deaton, Angus. 2000. "Preliminary Notes on Reporting Periods in the Indian NSS, 52nd through 54th Rounds." Research Program in Development Studies, Princeton, N.J.

Deaton, Angus, and Margaret Grosh. 2000. "Consumption." In Margaret Grosh and Paul Glewwe, eds., *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 years of the Living Standards Measurement Study*, vol. 1. Washington, D.C.: World Bank.

Deaton, Angus, and Alessandro Tarozzi. 2000. "Prices and Poverty in India." Princeton University, Program in Development Studies, Princeton, N.J. (July). Available on the World Wide Web at www.wws.princeton.edu:80/~rpds/working.htm.

Easterlin, Richard A. 1995. "Will Raising the Incomes of All Increase the Happiness of All?" *Journal of Economic Behavior and Organization* 27:35–47.

Filmer, Deon, and Lant Pritchett. 1998. "Estimating Wealth Effects without Expenditure Data—Or Tears: With an Application to Educational Enrollments in States of India." World Bank Policy Research Working Paper 1994, Washington, D.C. (October). Processed.

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

———. 2000. *Choice of Reference Period for Consumption Data*. National Sample Survey Organization, Calcutta, India.

Groves, Robert M., and Mick P. Couper. 1998. *Nonresponse in Household Surveys*. New York: Wiley.

Mahalanobis, P. C., and S. B. Sen. 1954. "On Some Aspects of the Indian National Sample Survey." *Bulletin of the International Statistical Institute* 34.

Mukherjee, M., and G. S. Chaterjee. 1974. "On the Validity of NSS Estimates of Consumption Expenditure." In T. N. Srinivasan and P. K. Bardhan, eds., *Poverty and Income Distribution in India*. Calcutta: Statistical Publishing Society, 139–47.

National Research Council. 1995. *Measuring Poverty: A New Approach*. Washington, D.C.: National Academies Press.

PovertyNet. 2000. Newsletter #21 (July). Available on the World Wide Web at www.worldbank.org/poverty/newsl/jul00.htm#4.

Pradhan, Menno, and Martin Ravallion. 2000. "Measuring Poverty Using Qualitative Perceptions of Welfare." *Review of Economics and Statistics* 82(3):462–71.

Ravallion, Martin. 1998. "Poverty Lines in Theory and Practice." LSMS Working Paper No. 133, World Bank, Washington, D.C.

———. 2000a. "Should Poverty Measures Be Anchored to the National Accounts?" World Bank, Washington, D.C. (July). Processed.

———. 2000b. "Do National Accounts Provide Unbiased Estimates of Survey Based Estimates of Living Standards?" World Bank, Washington, D.C. (August). Processed.

Ravallion, Martin, and Benu Bidani. 1994. "How Robust Is a Poverty Profile?" *World Bank Economic Review* 8(1):75–102.

Ravallion, Martin, and Binayak Sen. 1996,. "Monitoring Poverty in Bangladesh." *Economic Development and Cultural Change* 44(4):761–92.

Scott, Christopher, and Ben Amenuvegbe. 1990. "Effect of Recall Duration on Reporting of Household Expenditures: An Experimental Study in Ghana." Social Dimensions of Adjustment in Sub-Saharan Africa Working Paper No. 6. World Bank, Washington, D.C.

Srinivasan, T. N. 2000. "Growth and Poverty Alleviation: Lessons from Development Experience." Department of Economics, Yale University, New Haven, Conn. (November). Processed.

Srinivasan, T. N., P. N. Radhakrishnan, and A. Vaidyanathan. 1974. "Data on Distribution of Consumption Expenditure in India: An Evaluation." In T. N. Srinivasan and P. K. Bardhan, eds., *Poverty and Income Distribution in India*. Calcutta: Statistical Publishing Society, 148–62.

Stiglitz, Joseph. 1997. "How to Fix the Asian Economies." *New York Times*, October 31, sec. A, p. 27, col. 2.

Thomas, D., K. Beegle, E. Frankenberg, and G. Teruel. 1999. "Household Budgets, Household Composition and the Crisis in Indonesia: Evidence from Longitudinal Household Survey Data." RAND Corporation, Santa Monica, Calif. Processed.

Triplett, Jack E. 1997. "Measuring Consumption: The Post-1973 Slowdown and the Research Issues." *Federal Reserve Bank of St Louis Review* May/June:9–42.

Visaria, Pravin. 2000. "Poverty in India during 1994–98: Alternative Estimates." Institute for Economic Growth, New Delhi (June). Processed.

World Bank. 1997. *World Development Report 1997*. Washington, DC: World Bank.