CHAPTER III. POVERTY MEASURES

Jonathan Morduch[13]

Introduction

This chapter focuses on ways in which statisticians aggregate survey data to describe the condition of poverty. All governments make poverty reduction part of their policy agendas, but how exactly should poverty be measured? This chapter takes up that question with respect to money-based measures: those poverty statistics that measure the degree to which individuals and households fall below a poverty line.[14] Just as there is much diversity in how surveys are collected, the practice of calculating poverty statistics also varies widely. The past twenty years have seen a great deal of convergence in understandings, though, and this chapter draws on what has been learned.

This chapter focuses on ways in which statisticians aggregate survey data to answer questions such as:

- How many poor people are there in a region?
- How deep is their deprivation?, and

[13] I have benefited from the comments of the United Nations Expert Group on Poverty Statistics, and in particular from input from John Gibson, Christian Grootaert, and Branko Milanovic.

[14] The poverty line, as described in the previous chapter, is set to reflect the money needed to purchase those goods and services deemed necessary for living a life free of basic deprivation.

- Has poverty risen since the last survey?

Any discussion of how to form poverty measures must begin with recognition that statistics have multiple constituencies (e.g., government policy makers, NGOs, researchers, and the general public), and these scattered constituencies often have competing needs and agendas. Choosing which poverty measure is best depends in large part on the uses to which it will be put. Since no single statistic is likely to answer the needs of all, most statistical offices publish a range of statistics. Below, recommendations are made for ways to expand the data range to make comparisons easier. Even better, although not always easy, would be to also make the raw survey data available for others to analyze (after taking appropriate actions to protect the confidentiality of surveyed households).

This chapter begins by describing and comparing four commonly-used poverty measures. The discussion emphasizes interpretation, and a recently-introduced metric, “exit time,” which is then defined and illustrated with examples from Bangladesh and Papua New Guinea. The chapter then moves to a discussion of measurement error and comparisons based on stochastic dominance that allow researchers to analyze trends in poverty without using explicit measures. The discussion is framed in terms of income or consumption deprivation—since this is the form of poverty usually analyzed with the measures below. But the measures can, in principle, also be used for analyzing other forms of deprivation when the underlying quantities are measured on a cardinal scale, such as under-nutrition, stunting (height for age), and wasting (weight for height). The chapter concludes with a set of recommendations on how to improve the measurement of poverty.

3.1 Desirable features of poverty measures

Poverty measures are used first and foremost to monitor social and economic conditions and to provide benchmarks of progress or failure. Here, poverty measures are indicators by which policy results are judged and by which the impact of events (e.g., runaway inflation or the introduction of a government transfer program) can be weighed. Measures used for monitoring and targeting need to be trusted and require rigorous underpinning. The measures will function well as long as everyone agrees that when poverty numbers rise, conditions have indeed worsened (and conversely, when poverty measures fall, that progress has been made). The first question in judging measures is how well does each index reflect basic properties desirable on philosophical grounds.[15]

A second important use for poverty measures is descriptive. Poverty statistics play critical roles in summarizing complex social and economic conditions that inform conversations around economic and social priorities. For this purpose, effective measures need not completely capture all (or even most) morally relevant aspects of poverty. But the limits of measures need to be understood, and transparency and ease of

[15] Transparency of method is critical in helping to achieve a consensus, and interested parties should be given enough information to understand exactly how the numbers were constructed, beginning with data collection methods and ending with aggregation techniques.

interpretation are critical here. These two notions—the need for rigor balanced against a desire for ease of interpretation—run through the discussion below.

Economists have sharpened discussions by identifying a set of desirable normative characteristics of poverty measures (often stated mathematically as axioms) around which consensus can be built. The search focuses not on identifying descriptively useful measures in the sense above; instead, the focus is on moral relevance—even if the outcome is a set of measures that yield numbers with little intuitive meaning.

If we can agree that acceptable poverty measures must satisfy a given set of axioms or must have certain characteristics, it is possible to sharply narrow the number of potential candidates for poverty measures. In the most desirable case, a single, unique measure would emerge that would be fully "characterized"—that is, there would be only one possible candidate that satisfies all of the axioms on which we agree. So far, though, the search has left a long list of possible poverty measures still on the table, and the task for analysts remains to understand the basic properties of the chief contenders.

While not succeeding at singling out a particular, universally-acclaimed poverty measure, the axiomatic approach pushes discussions forward in useful ways, and the central ideas are worth reviewing. Building blocks include concepts such as "scale invariance." This is the idea that poverty measures should be unchanged if, for example, a population doubles in size while everything else is maintained in the same proportions.

A second building block focuses on the well-being of those below the poverty line—so that changes among better-off people do not affect measured poverty This "focus axiom" rules out measures based on relative notions of poverty (i.e., where poverty is not measured by absolute deprivations relative to a fixed poverty line but instead the poor are identified relative to a shifting statistic like the median income of the whole population). Our focus here is on "absolute poverty" as measured by a fixed poverty line.

A third attribute, the "monotonicity" axiom, states that, holding all else constant, when a poor person's income falls, poverty measures must rise (or at least should not fall).

The "transfer" axiom (sometimes referred to as the Pigou-Dalton principle, after those who employed it first in their analyses) has more analytical bite. It states that, holding all else constant, taking money from a poor person and giving it to a less poor person must increase the poverty measure. Conversely, poverty falls when the very poor gain through a transfer from those less poor.

"Transfer sensitivity," a related notion, goes further. It is best seen by example. Consider a population where the poverty line is set at $1,000. Next, imagine that $10 is taken from someone earning $600 and given to a neighbor earning $500. Any poverty measure that satisfies the transfer axiom will fall. Measured poverty should also fall (for such indices) when $10 is taken from someone earning $300 and given to someone earning $200. The transfer-sensitivity axiom says that the reduction in the second case (in which a very poor person is made better off relative to her neighbor) should be greater than the reduction in the first case (in which the recipient is less poor).[16]

An additional desirable characteristic is the ability to decompose poverty measures by sub-population. Sub-populations may include, for example, residents of different regions. The critical feature for decomposition is that the sub-groups are distinct from each other (so that there is no overlap in membership) and that together they encompass the entire population. All additive indexes are decomposable, and all of the measures described below share the feature.[17]

3.2 Four common measures

The simple headcount index is the most used poverty measure, but it violates several important axioms. Of the four measures described below, the one that satisfies all of the desirable axioms above--the Watts measure--turns out to be the least used. These two facts suggest an ongoing tension between the desire for simplicity and transparency pitched against the desire for rigor. The measures below will be compared in that light.

All of the measures will be described in terms of shortfalls of "income." The focus on income keeps discussion simple, but the measures may instead may be used to gauge shortfalls in consumption and spending—following the discussion in Chapter 1.

[16] For a broader discussion of axioms, see Sen (1976), Foster (1984), and Foster, Greer, and Thorbecke (1984).
[17] Decomposability is the focus of Foster, Greer, and Thorbecke (1984).

Also note that given that nearly all surveys are household-level surveys, income (or consumption) will either be put in per capita terms or per adult equivalent.

3.2.1 Headcount measure

The headcount is the simplest and best known poverty measure. It identifies the share of a population whose income is less than the poverty line. It is, not surprisingly, the most commonly calculated poverty measure. The measure literally counts heads, allowing policymakers and researchers to track the most immediate dimension of the human scale of poverty.

The headcount is calculated by comparing the income y_i of each household to the poverty line z. (The index $i = 1 \ldots M$, where M is the total number of households in the sample.) Concretely, an indicator variable is constructed for each household, taking the value 1 when income falls below the poverty line or 0 if income is greater:

$$I(y, z) = 1 \text{ if } y_i \leq z$$

$$I(y, z) = 0 \text{ if } y_i > z$$

The headcount index is simply the sample average of the variable $I(y, z)$, weighted by the number of people in each household n_i.[18] The measure is calculated by first counting the number of poor individuals, G:

$$G = \sum_{i=1}^{M} I(y, z) n_i .$$

[18] Note that total household size, n, is used even where income and the poverty lines are designated in terms of adult equivalents. This is not the case when calculating the poverty gap, as discussed below.

Total population of the sample can be calculated similarly as

$$N = \sum_{i=1}^{M} n_i ,$$

and the overall headcount is then the ratio of the two numbers:

$$H = G/N .$$

Where the sample is not representative of the underlying population (e.g., if the sampling strategy involved random stratification), population weights should also be included in the calculation (see chapter 5 for further discussion).

The headcount is an important descriptive tool. As a sole guide to allocating resources, though, the headcount can significantly mislead. There are two large tensions. First, the headcount registers no change when a very poor person becomes less poor. Nor does the headcount change when a poor person becomes even poorer. Most observers, though, following Watts (1968) and Sen (1976), argue that changes in the income distribution below the poverty line matter in a moral sense. This notion is captured by the transfer axiom above, but the headcount fails the test.

A second tension flows from the failure of the transfer axiom, combined with the focus on whether people are above or below the poverty line. If policymakers see their task as reducing poverty as measured by the headcount, their work will be made easier by focusing on improving the lot of individuals just below the poverty line. A little improvement at this level can raise the incomes of the "barely poor" above the poverty line and hence can reduce the poverty headcount fairly rapidly. Directing resources to very poor people, on the other hand, may be socially beneficial, but far larger income gains are required to take them over the poverty line and thus to make a dent in the poverty headcount. So if efforts are allocated specifically to reduce the headcount, priority will likely go to helping the least poor over helping the poorest.

The headcount remains a highly valuable measure, even if, when used on its own, it is a poor guide for resource allocation. One step to make the approach more useful is to calculate the headcount for "sub-poverty" lines at lower thresholds than the overall poverty line. These may capture, for example, the income required to purchase the food basket only, excluding non-food needs. Tracking the population under sub-poverty lines is a first, simple step—and often a powerful descriptive tool.

3.2.2 Poverty gap

This second widely-used measure has a problem similar to the headcount: it is descriptively very useful but, if used alone, would also serve as a poor guide to resource allocation. The poverty gap measures the amount of money by which each individual falls below the poverty line. It matters here whether income and the poverty line are measured on a per capita basis or whether they have been put into adult equivalent terms or adjusted for scale economies (Milanovic, 2002). The appropriate formulas are given below.

The starting point is to calculate the total shortfall in income for the poor population:

$$Shortfall = \sum_{i=1}^{M} (z - y_i)\, I(z, y_i)\; n_i \,,$$

where the poverty line is z, income is y, $I(z,y_i)$ is a 0/1 indicator of poverty for each household, household size is n_i, the total number of households in the sample is M, and individuals are indexed by i. The calculation gives the total sum of money that would be needed to make up for the gap between the existing incomes of the poor and the official poverty line.

The calculation above is correct only if income is in per capita terms. When income is made instead in adult equivalent terms (or adjusted for scale economies), the correct calculation is:

$$Shortfall = \sum_{i=1}^{M} (z - y_i)\, I(z, y_i)\; a_i \,,$$

where a_i gives the number of adult equivalent units in household i.

As a sum, the figure above may be helpful for budget planners, but it obscures the sense of individual deprivations. An alternative is to instead calculate the average shortfall for the population below the poverty line:

$$\frac{Shortfall}{G} \qquad (1)$$

When viewed together with the headcount, this version of the poverty gap measure shows the distance (on average) to be traveled in raising incomes. Because the figure is denominated in currency, conversion to a common international currency (e.g., euro or dollar) will aid global comparisons.

A different approach that can enhance comparability is to divide the index by the poverty line:

$$\frac{Shortfall}{Gz} \tag{2}$$

Normalization puts the average gap in terms of the percentage shortfall from the poverty threshold, freeing the measure from denomination in a particular currency. The measure is now easily comparable across countries and across time, a helpful improvement. Routinely publishing poverty lines alongside the normalized poverty gap and the headcount would allow observers to calculate for themselves all three of the poverty gap variations described above.

The three data points (headcount, poverty gap, and poverty line) can be combined to form another widely-used variant of the poverty gap:

$$\frac{Shortfall}{Nz} \tag{3}$$

Here, the resource shortfall is divided by the total population, rather than the population of the poor. The measure is often misinterpreted as giving the average income shortfall of the poor, but that is only the case for versions (1) or (2). Dividing by the total population sacrifices simple interpretations—the measure no longer gives a quick sense of deprivation of poor individuals since data on non-poor people are also included. The measure points up tensions between descriptively-useful measures and measures that can best serve as guides for monitoring and targeting.

The measure in (3) does, though, avoid a common problem with variants (1) or (2). Specifically, the poverty gap in (1) and (2) can rise—rather than fall—when individuals exit poverty. This occurs when the least poor are the ones who move above the poverty threshold (which is the typical pattern). Holding all else the same, those who exit poverty leave behind a population that is then smaller and, on average, poorer than before. Conditions would thus seem to worsen when someone exits poverty--if seen through the lens of the poverty gap as calculated following (1) or (2)--when in fact conditions have improved overall. The measure in (3) instead captures improvement due to the exit. The problem here occurs when the poverty gap in (1) or (2) is used as a *sole* indicator of progress; it performs poorly when people cross the poverty line.

Version (3) does not have that problem. But as noted, it lacks a simple interpretation and, along with all of the versions above, it fails to satisfy the transfer axiom. This failure follows from neutrality with regard to whose income goes up and down among the poor population. For the transfer axiom to be satisfied, it is not enough to be neutral: it must be that gains for the poorest are weighed more heavily than gains for the less poor. Even in version (3), no accommodation is made to weigh progress in reducing extreme poverty differently from progress in reducing moderate poverty.

In summary, version (2) can be useful as a descriptive tool, especially alongside other measures. Version (3) has some desirable properties from an analytical vantage. But as discussed below in sections 3.2.3 and 3.2.4, distributionally-sensitive measures do even better. Version (3) has gained favor by being a member of the Foster-Greer-

Thorbecke class of measures described below (specifically, the case in which $\alpha = 1$). But the attention seems misplaced. If a single poverty measure is required, version (3) of the poverty gap has less appeal than the measures described in sections 3.2.3 and 3.2.4--although it would be preferred over having, say, version (2) on its own.

All of this is rhetorical since it is seldom that only one poverty measure is calculated. When multiple poverty measures are produced and published simultaneously (say, the headcount and a few others), version (2) of the poverty gap stands as a useful part of a collection. Version (2) features a clear, simple interpretation that is relevant for policy discussions, an attribute lacking in version (3). A recommended set of basic poverty indicators would thus include the headcount, version (2) of the poverty gap, the median income of the poor, and the squared poverty gap described in section 3.2.4.

3.2.3 Watts index

A simple poverty measure that satisfies the transfer axiom was first put forward by Watts (1968), who argued for the following measure:

$$\frac{1}{N}\sum_{i=1}^{M}\left[\ln(z)-\ln(y_i)\right] I(z,y_i)\, n_i . \tag{4}$$

As above, when income is calculated in adult equivalent terms, the household size variable n_i should be replaced with the adult equivalent size a_i. The measure is "distributionally-sensitive" by virtue of its use of logarithms.[19] The way that the logarithm is used means that the Watts index is much more sensitive to changes in the lowest incomes than it is to changes for those with higher incomes. That is, transferring \$10 to a very poor person counts as a far larger contribution to poverty reduction than transferring \$10 to a richer (but still poor) neighbor.

Allocating anti-poverty resources to minimize the Watts index would thus tilt efforts toward the poorest—which is a feature that many analysts find appealing (and one also featured by the squared poverty gap and the influential (but seldom applied) measure of Sen, 1976). The Watts index also satisfies the transfer-sensitivity axiom described above, and it is decomposable into the population-weighted sum of the poverty indices of regions or groups. (The squared poverty gap of section 3.2.4 shares this feature too.) Being decomposable is useful when a population can be divided into a number of distinct groups or regions. Poverty measures can then be calculated for each group or region, and, if the poverty measure is decomposable, the individual poverty measures can be aggregated (using population shares as weights) to form the overall poverty measure for the entire population. Decomposing poverty measures in this way can help to pinpoint the groups or regions contributing most—and least--to overall poverty.

Given these appealing features, the Watts index can be a useful measure, and we return to it below in comparison to the "squared poverty gap" and the "average exit time" measures (see sections 3.2.4, 3.3, and 3.4). The comparisons in section 3.3 show that an

[19] If a given number is larger than another, then the logarithm of the first will also be larger than the logarithm of the second. But the logarithmic transformation is nonlinear, i.e., the ratio of the two numbers will be larger than the ratio of their logarithms.

important difficulty with the Watts index is that the weights on money in the hands of the "least poor" and the "moderately poor" are quite close, while the weights on the "most poor" and destitute are particularly heavy. As section 3.3 shows, the "squared poverty gap," in contrast, leads to greater differentiation across the entire distribution of incomes below the poverty line and may be preferable for that reason.

3.2.4 Squared poverty gap

One way to transform the poverty gap described above into a distributionally-sensitive measure is to raise the individual gaps to a power greater than 1. Foster, Greer, and Thorbecke (1984; henceforth, FGT) propose a class of measures built on this idea which have found their way into much of the poverty analysis published by the World Bank. With income expressed in per capita terms, the measures take the form:

$$\frac{1}{N}\sum_{i=1}^{M}\left(\frac{(z-y_i)}{z}\right)^{\alpha} I(z,y_i)\, n_i. \qquad (5)$$

When income is in adult-equivalent terms, the household size variable n_i should be replaced with the adult equivalent size a_i.

The parameter α determines the degree to which the measure is sensitive to the degree of deprivation for those below the poverty line. When α is zero, the measure collapses to the headcount measure described above, and when α is one, the measure is the normalized version of the poverty gap (equation 3 above). In neither case is the measure distributionally-sensitive. But for $\alpha > 1$, the measure is distributionally sensitive. And the particular case in which $\alpha = 2$ (often referred to as the squared poverty gap) is now the most widely-used distributionally-sensitive measure.

By squaring the poverty gap, improvements in the resources of the poorest individuals count most, since they are the ones for whom the initial resource gap is largest. The measure satisfies the transfer axiom but not the transfer-sensitivity axiom. To satisfy the latter, the poverty gap would have to be raised to a higher power—cubed rather than squared, say. Cubing adds "transfer sensitivity," a property that many find appealing. But it puts very heavy weight on the well-being of the poorest—perhaps weight that would be judged too great in a social calculus.[20] At levels of α between 1 and 2, not only is transfer sensitivity not satisfied but the reverse holds: holding all else the same, a regressive transfer among the very poor increases poverty *less* than a same-sized regressive transfer among the moderately poor—a clearly undesirable feature.

Distributional-sensitivity is achieved by weighing deprivations of the poor inversely to their base incomes. There are many ways to do this, and the weighting scheme in the squared poverty gap has the advantage of relative simplicity. The simplicity can help provide some intuition in understanding why the poverty measure moves over time. This is described below in the context of Figure 1.

3.3 Comparing the measures

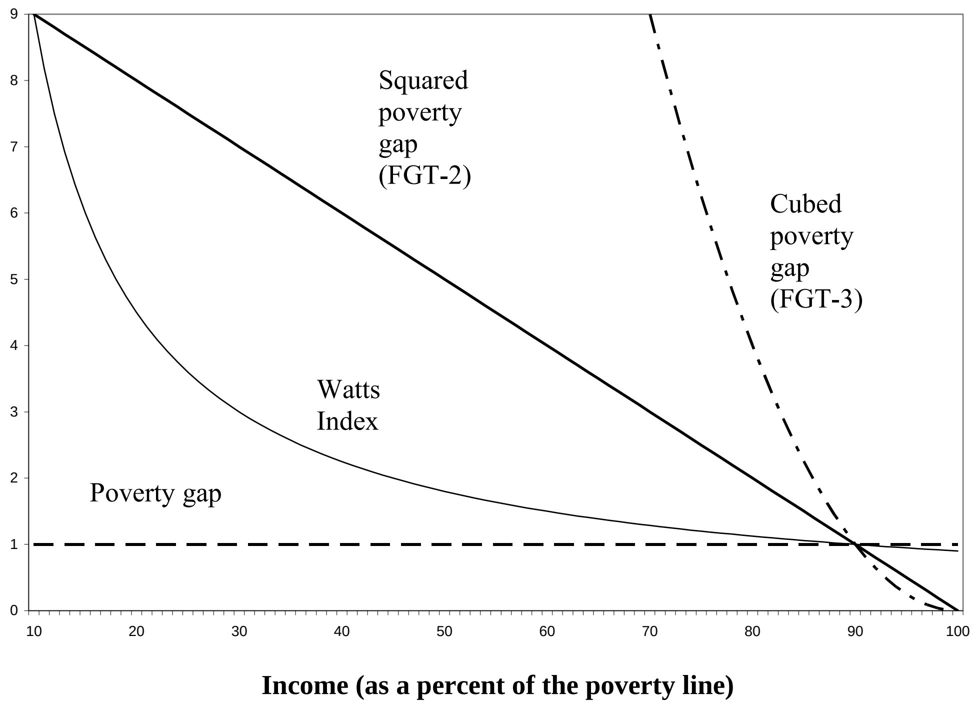
[20] Transfer sensitivity is obtained through setting α at any level larger than 2 (2.2, for example), but analyses have focused on integer values.

Neither the Watts index, nor the squared poverty gap, nor the cubed poverty gap yield particularly easy-to-interpret numbers. As noted above, we rely for interpretation on our knowledge that the indices satisfy (or do not satisfy) certain axioms, notably the transfer axiom and transfer-sensitivity axioms. Often that is enough, and at times each of the three indexes will rank income distributions identically. So the form of the poverty index may not matter when answering relevant policy questions. In other cases, though, results will depend on the choice of index, and the relatively opaque nature of the indexes hides the fact that even indexes with similar basic properties (e.g., Watts and the cubed gap) weigh income gains very differently for different people.

Figure 1 makes this explicit. The question asked in the figure is: how does giving \$1 to a person with income equivalent to, say, 50 percent of the poverty line compare to giving the \$1 to someone with income equivalent to 90 percent of the poverty line? Specifically, how does the transfer contribute to measured poverty?

The figure shows how the answer depends on the choice of poverty measure. It demonstrates the relative weight of a \$1 increase in income as implied by the four measures above—the poverty gap (equation 3 above), squared poverty gap, cubed poverty gap, and the Watts index. The horizontal axis gives a poor person's income relative to the poverty line, from 0 percent at the left to 100 percent at the right. Those on the far right are just at the poverty line (which was set at \$100 for this illustration). The \$1 increases are depicted relative to a \$1 increase for a person whose income is at 90 percent of the poverty line (i.e., an income of $90 when the poverty line is $100). By construction, all four curves thus meet at 90 on the horizontal axis.

Figure 1: Comparison of implicit weights in poverty measures.



The weight of an additional $1 of income for poor individuals at different levels, relative to the weight on an additional $1 of income for individuals with income equal to 90 percent of the poverty line.

The curve giving weights for the poverty gap is perfectly flat, showing that the measure is not sensitive to who gains the income--a dollar is a dollar no matter whether it accrues to the most poor or the least poor. The other three curves slope downward, though, indicating that a dollar earned by the most poor weighs more in these indexes.

The curve for the cubed-poverty gap arcs so sharply that it goes off the graph. Giving $1 to a person weighs 9 times more if the recipient has income of $70 than if the recipient starts with income of $90 (relative to a poverty line of $100). The weighting is thus heavily skewed to the very poor – and certainly far more skewed than is the case for the squared poverty gap.

For the squared poverty gap, the ratio of weights is proportional to base income. For example, a dollar accruing to someone with income of $70 (i.e., 30 percent below the line) is 3. At $80 (i.e., 20 percent below the line), the ratio is 2. And so forth. This property would be highlighted if the measure were renamed the "gap-weighted" index, say, since the income gap for each household (i.e., $z - y_i$) is weighted by the size of the income gap itself (which is, again, $z - y_i$).

The Watts index takes a very different form, weighing increases from very low incomes very heavily, but staying flatter for less poor individuals. The ratio in the figure only hits 2 when the added dollar accrues to individuals whose incomes are below $50. The relative weights, as seen by the ratio in the figure, are always below those of the squared gap, except for transfers made to people with incomes below $10. At that level, income is so low that (it is hoped) few if any people could survive at those levels. The weighting scheme corresponds with moral concerns that focus on the very, very poorest with extreme intensity, in contrast to the squared poverty gap which has a linear profile of relative weights. On the other hand, the Watts index exacerbates bias due to measurement error if the reported incomes at the very low end are largely the result of poor data collection and mis-reporting.

None of the weighting schemes can claim universal preference from a normative stance, but the square poverty gap is a middle ground, and this likely explains much of its continuing appeal. The Watts index, though should not be altogether dismissed. It has a useful feature through its association with the exit-time concept and value as a descriptive tool.

3.4 Exit time and the value of descriptive tools

For all of the theoretical appeal of the distributionally-sensitive measures described immediately above, the headcount remains—by far—the most common poverty measure in use. The Millennium Development Goals, for example, focus on reducing the headcount of poverty below $1/day, rather than minimizing a distributionally-sensitive measure.

One reason for the continuing use (and usefulness) of the headcount is its descriptive properties. It is a simple means for illustrating the scale of poverty. In this sense, it is an intrinsically meaningful measure.

The poverty gap is also intrinsically meaningful, taking us from counting people to counting shortfalls of income or consumption. It answers the question: how much would have to be spent to eliminate poverty through costless (and perfectly) targeted transfers. Its underlying assumptions are clearly unrealistic: in practice, transfers will never be administratively costless, nor will they ever be perfectly targeted. However, this hypothetical question still provides a helpful way to quickly gauge the scale of deprivation.

The "average exit time," introduced in Morduch (1998), is based on a similar sort of hypothetical question. The underlying assumption is as unrealistic as that under the poverty gap, but the measure nevertheless can fruitfully frame discussions of poverty. The measure is based around the number of years that it would take poor households to grow out of poverty given a hypothetical, steady growth of income. (An equivalent question could also be asked about consumption growth, rather than income.)

In practice, income growth will seldom be steady over long periods, nor will all poor households be able to grow at the same rate. But, as with the poverty gap, asking the hypothetical question provides a quick way to gauge one important aspect of the condition of poverty.

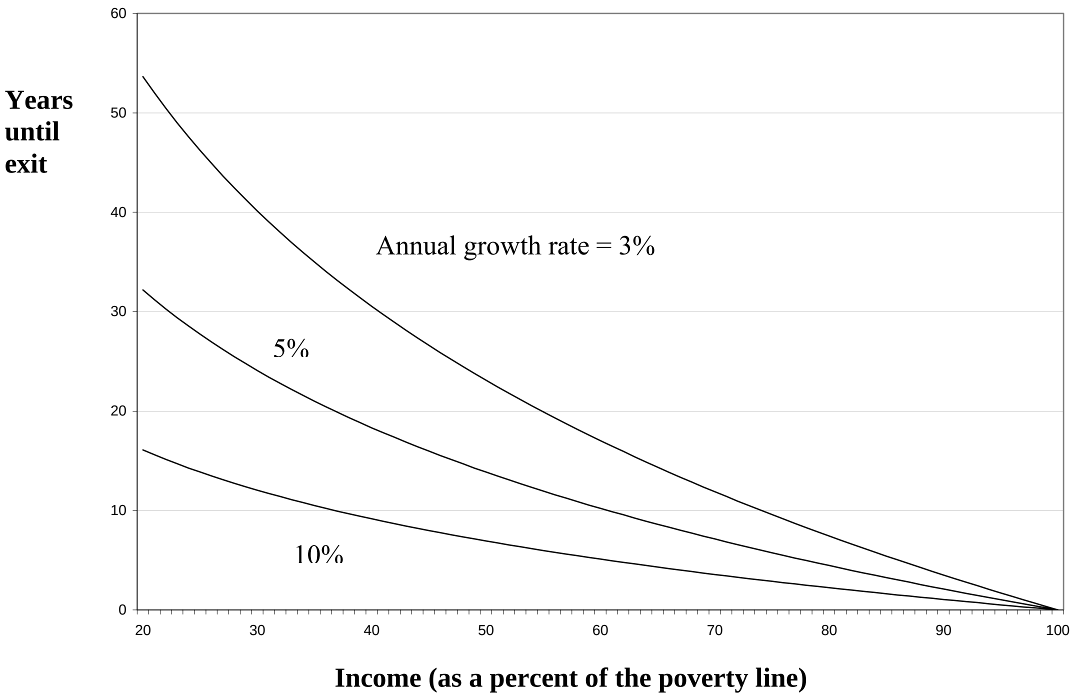
Hypothetical exit times are simple to calculate. If the assumed growth rate of income is g percent per year, an individual whose income starts at y_i will take T years to exit, where T solves this equation:

$$z = y_i (1+g)^T \tag{6}$$

The equation can be solved by taking logarithms, yielding that the number of periods of growth required before exit is $T = \ln(z/y_i) / g$. Of course, $T = 0$ for all households already above the poverty line. So, for example, if a person's income starts at 80 percent of the poverty line and his income grows at 5 percent per year (after adjusting for inflation), his exit time will be ln(100/80) / 0.05 = 4.5 years.

The calculation shows that within five years, consistent, broad-based income growth of 5 percent would be enough to push from poverty everyone whose income is 80 percent of the poverty line or higher. Exit times based on this calculation are shown in Figure 2, based on assumed growth rates of 3 percent, 5 percent, and 10 percent. Alternatively, a figure could be constructed that fixes the hypothetical growth rate and maps a profile of exit times for the entire poor population, with exit times shown according to the fraction of the population at each level of income below the poverty line.

Figure 2: Hypothetical exit times as a function of income below the poverty line



In a related calculation, Morduch (2000) focuses on the "median exit time" when illustrating the use of exit times with data from Bangladesh. The calculation is simple: $\ln(z/y_M)/g$. The only data required are the poverty line z, the assumed growth rate g, and the median income of the population below the poverty line y_M. The median poor rural household in the 1988-89 Household Expenditure Survey spent Taka 284 per month per capita relative to the poverty line of Taka 370. (In 1989, Taka 32.1 = $1.) So, if median expenditures grew steadily at 3 percent per year, it would take just under 9 years to reduce half of rural poverty through growth alone—ln(370/284)/.03.

If the hypothetical exit time , T_i, for each poor household is averaged over the population below the poverty line, the "average exit time" is:

$$A = \frac{1}{G}\sum_{i=1}^{M}\frac{\ln(z/y_i)}{g}\,I(z,y_i)\,n_i = \frac{1}{G}\sum_{i=1}^{M}T_i\,n_i \qquad (7)$$

This equation is analogous to the average poverty gap described in equation (2), and it shares similar weaknesses and strengths. Its chief strength is its simplicity and descriptive value. Its main weakness is that when a less poor household exits poverty and all else is unchanged, the average exit time, A, will fall. This makes A a poor candidate to be the sole measure of poverty. However, A can still be a very useful component of the analyst's toolkit.

One aspect that makes the average exit time potentially valuable is that it can be decomposed explicitly to show the impact of inequality below the poverty line. Start with the average income of households below the poverty income:

$$y^{avg} = \frac{1}{G}\sum_{i=1}^{M} y_i \; I(z, y_i) \; n_i .$$

If everyone below the poverty line had exactly this income (i.e., there was no inequality below the poverty line), then this hypothetical average exit time would be:

$$T^{avg} = \ln(z / y^{avg}) / g .$$

Using these relationships, the average exit time of the poor population (A in equation 7 above) can be rewritten simply as:

$$\text{A} = T^{avg} + L / g, \tag{8}$$

where L is the Theil-L measure of inequality--a commonly used inequality index. The decomposition shows the explicit contribution of income inequality to the average exit time.

In their survey of 1,144 households in Papua New Guinea, Gibson and Olivia (2002) found that given an assumed, hypothetical growth rate of 2 percent per year, the average exit time of the population would be 20.5 years. Their calculation helps to frame the potential importance of growth-based strategies—if growth is steady and broad. They decompose the result to show that the exit time of a person with income equal to the average income of the poor population (i.e., T^{avg}) is 17.8 years, so the explicit impact of inequality below the poverty line is an increase in the average exit time by 2.7 years.

The exit time has a useful relationship to another established measure. If exit times are calculated for the entire population of a country (with those above the poverty line having 0 exit times), the average turns out to be simply the Watts poverty index divided by *g*, the hypothetical annual growth rate.[21] This measure, the "population average exit time," naturally shares all properties of the Watts index, satisfying both the transfer axiom and the transfer-sensitivity axiom. But it has the addition of a new interpretation, akin to the interpretation of the poverty gap described by equation (3) above.

With economic growth very much a part of the poverty reduction policy agenda, tools like exit times provide ways to summarize data in a manner relevant to policy debates on growth-based poverty strategies. They complement the other measures described above, rather than substituting for measures whose appeal rests primarily on their axiomatic properties. By the same token, those theoretically-appealing measures cannot substitute for simpler tools that provide new ways of describing the data and identifying trends. And, importantly, it should be remembered that exit times describe possibilities based on simple assumptions—as used here, the exit times are not based on actual forecasts or careful predictions. These simple exit times, though, can be useful in identifying opportunities and constraints to guide policy.

[21] This measure satisfies the monotonicity transfer, and transfer-sensitivity axioms. As a result of its additively separable form, the measure is also decomposable into the population-weighted measures of sub-populations of the poor.

3.5 Broader concerns

3.5.1 Comparisons without poverty measures

For many purposes, the relevant policy question is simply: is poverty larger or smaller in one survey versus another? As noted above, the answer can depend on which poverty line and which measure is used in the analysis. But that is not always so. It may be that one survey finds a greater fraction of the population exists at *every* income level below the poverty line, relative to the fractions of the population reported by another survey.

When that is so, the first survey can be said to "stochastically dominate" the latter (Ravallion, 1994), and conclusions about poverty rates can be made irrespective of whether the poverty line is moved lower or whether the poverty measure chosen is the Watts index, headcount, or one of the poverty gap measures described in section 3.3. Formally, the condition for "first order" stochastic dominance is that with regard to two samples, A and B, if the cumulative distribution function of their income distributions is such that $D_A(y) < D_B(y)$ for all incomes y below the poverty line, then sample A stochastically dominates sample B.

When income distributions cross below the poverty line (i.e., when $D_A(y) < D_B(y)$ in some income range but $D_A(y) > D_B(y)$ in another), more restrictive notions of stochastic dominance can be applied which, nonetheless, allow broad statements that are robust to a wide range of choices of poverty lines or measures. These approaches—"second order" and "third order" stochastic dominance--are described further by Ravallion (1994). When these robust approaches can be employed, questions about the choice of method can be avoided. Their disadvantage is that they only answer a simple question—is poverty higher or lower in sample A than B? Richer descriptive tools (like the hypothetical exit times discussed above) are required to inform richer policy questions.

3.5.2 Measurement error

The chapters that follow in the handbook focus largely on methods for collecting surveys. No survey is perfect, but some collection methods are far more reliable than others. Particular problems arise when expenditures (or incomes) are either substantially over-counted or under-counted, and the biases can be exacerbated by the choice of poverty measure. Under-counting leads to exaggerations of poverty rates, and the distributionally-sensitive measures described here are particularly susceptible to the exaggeration of under-counting in the lower tail of the income distribution.

Figure 1 (above), which shows the implicit weights in several popular poverty measures, also helps to illustrate the potential problems of under-counting. The figure shows the weight given to an extra $1 of income for households at different levels of base income, relative to the weight of an extra $1 for an individual whose income is 10 percent below the poverty line. The figure also can be interpreted as giving the weights of $1 of mis-measured income for different individuals. Clearly, for distributionally-sensitive measures, every dollar that is mis-measured at the low end of the distribution has a far larger impact on poverty than a dollar mis-measured closer to the poverty line.

The pattern of weights can mean that mis-measurement in just a small fraction of observations can make a large difference to results if it takes the form of severe under-counting. Many household surveys, for example, include responses from some households about spending and income patterns that, for one reason or another, are implausibly close to zero. If those observations are taken at face value, they can translate into large movements in poverty measures.

When using the Watts index, for example, adding an extra dollar for someone with measured income that is extremely low—say, their total income is no more than 5 percent of the poverty line (a level so low that long-term survival is hard to imagine if income is measured correctly and savings are unavailable)--would lead to a change in the index which is 18 times the weight of the dollar for someone with income at 90 percent of the poverty line. (This result is so extreme that it cannot be seen in Figure 1.) When using the squared poverty gap, the relative weight would be 9.5 times. It is unclear from a moral standpoint which is the "correct" weight, but in choosing a poverty measure analysts are implicitly making a choice, and it should be borne in mind. In making that choice, it matters greatly whether one believes that very low incomes are most likely a function of measurement error or whether they reflect actual conditions. Fortunately, there are very few observations in these data ranges, but the weighting accentuates their importance.

The best solution is to maintain high-quality data and to be especially vigilant about potential measurement error at the low end. The following chapters, especially Chapter 5, provide guidance on this matter. But where the quality of data is uncertain, it is important to remain alert to prevent the choice of measure that could worsen quality problems. Robustness checks become all the more helpful.

3.6 Conclusions

There is now an extensive literature on poverty measures, and statistical offices have a wide range of numbers to analyze. Focusing on the most commonly used measures promotes comparability across countries. These include the headcount index, poverty gap, and squared poverty gap. This chapter describes how to calculate and interpret these measures (in addition to the Watts index), and identifies their respective strengths and weaknesses.

When multiple poverty measures are produced and published simultaneously (say, the headcount and a few others), a recommended set of basic poverty indicators includes the headcount in section 3.2.1, version (2) of the poverty gap in section 3.2.2, and the squared poverty gap described in section 3.2.4.

Statistical offices can go further, though, by also publishing simple statistics that provide a richer picture of conditions. These statistics are seldom very costly to compute and can substantially enrich analysis. The first is the median income of the poor population. The median gives the income level below which the bottom 50 percent of the poor population lives. This simple measure indicates whether the bulk of the poor population is close to or far away from the poverty line. In section 3.5 above, it was also shown how median income can be employed in the exit-time framework.

Going further, it would be helpful to also publish the income of households at the 25th percentile and the 75th percentile of the income distribution below the poverty line. Ultimately, publishing the entire Lorenz curve (the mapping of population shares to income shares) would be most revealing and would add little extra cost. The median, though, is the natural place to start, followed by incomes at other important focal points of the distribution.

The move toward publishing headcounts of the number of people below a "hardcore" poverty line has accelerated. This is a useful step from a descriptive vantage, although there is currently little consensus on how to define hardcore poverty.

The exit-time framework was introduced in section 3.4 as an example of a simple metric that can help policy makers debate ways to promote economic growth and alleviate poverty. It is a descriptive device based on an unrealistic, best-case scenario. It asks: how quickly would households exit poverty if their incomes grew at a given, fixed rate each year? While hypothetical, the answers integrate the notion of time into poverty programs. The exit time is put forward as a complement to the poverty-gap measure of section 3.2.2, which is built on a similarly unrealistic hypothetical policy scenario, but which nonetheless offers useful insights.

Over time, new measures and approaches will continue to emerge. One of the most valuable steps that statistical offices can take is to put in place ways to make the raw data for poverty analyses available to researchers. Steps would have to be taken to secure confidentiality to households in the survey, but fortunately methods to do so are now well-established. Broadening access to data will allow analysts to better compare conditions and to develop new tools that can ultimately benefit statistical offices, policy makers, and citizens.

References

Foster, James (1984), "On Economic Poverty: A Survey of Aggregate Measures," *Advances in Econometrics* 3, 215 - 251.

Foster, James, Joel Greer and Erik Thorbecke (1984), "A Class of Decomposable Poverty Measures." *Econometrica* 52, May: 761-766.

Gibson, John and Susan Olivia (2002). Attacking Poverty in Papua Guinea, But for How long?" *Pacific Economic Bulletin* 17 (2).

Kanbur, Ravi (1987), "Measurement and Alleviation of Poverty, with an Application to the Effects of Macroeconomic Adjustment," *IMF Staff Papers* 34(1), March, 60 -- 85.

Milanovic, Branko (2002). "Do We Tend to Overestimate Poverty Gaps? The Impact of Equivalency Scales on the Calculation of the Poverty Gap." *Applied Economics Letters* 2002 (9): 69-72.

Morduch, Jonathan (1998). "Growth, Poverty, and Average Exit Time," *Economics Letters* 58: 385 - 390.

Morduch, Jonathan (2000). "Reforming Poverty Alleviation Policies," in Anne Krueger, ed., *Economic Policy Reform: The Second Stage.* Chicago: University of Chicago Press, 2000.

Ravallion, Martin (1994). *Poverty Comparisons.* Chur, Switzerland: Harwood Academic Press.

Ravallion, Martin (2002). "How *Not* to Count the Poor: A reply to Reddy and Pogge." World Bank working paper.

Sen, Amartya (1976), "Poverty: An Ordinal Approach to Measurement," *Econometrica* 44: 219- 231.

Sen, Amartya K. (1987), *The Standard of Living*. The Tanner Lectures. Cambridge: Cambridge University Press.

Watts, Harold (1968), "An Economic Definition of Poverty," in *On Understanding Poverty*, ed. by Daniel Patrick Moynihan. New York: Basic Books.