

NBER WORKING PAPER SERIES

WHY MEASURE INEQUALITY?

Louis Kaplow

Working Paper 9342

<http://www.nber.org/papers/w9342>

NATIONAL BUREAU OF ECONOMIC RESEARCH

1050 Massachusetts Avenue

Cambridge, MA 02138

November 2002

I am grateful to Steven Shavell for comments, J. Raza Lawrence and David Schleicher for research assistance, and the John M. Olin Center for Law, Economics, and Business at Harvard University for financial support. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

© 2002 by Louis Kaplow. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Why Measure Inequality?  
Louis Kaplow  
NBER Working Paper No. 9342  
November 2002  
JEL No. D31, D63, H23, I32

**ABSTRACT**

A large body of literature is devoted to the measurement of income inequality, yet little attention is given to the question, Why measure inequality? However, the reasons for measurement bear importantly on whether and how measurement should be done. Upon examination, normative measures are found to be of questionable value. Descriptive measures, by contrast, may be useful, but the appropriate measure depends on the field of application rather than on general, a priori principles of the sort that are emphasized in the existing measurement literature. Measures of poverty are also considered, and similar conclusions are reached.

Louis Kaplow  
Harvard Law School  
Hauser 322  
Cambridge, MA 02138  
and NBER

## 1. Introduction

A substantial and growing literature develops various measures or indexes of economic inequality. Some use the Gini coefficient or other measures or relationships drawn from Lorenz curves; some prefer different indicators of dispersion, such as an entropy index; some offer axiomatic derivations of inequality indexes; and still others advocate the use of normative measures derived from social welfare functions. The extent of this work is indicated by the recent publication of two handbooks, the *Handbook of Income Distribution* (Atkinson and Bourguignon 2000), much of which addresses measurement issues, and the *Handbook on Income Inequality Measurement* (Silber 1999), devoted entirely to the subject.<sup>1</sup>

Despite the extent of writing on inequality measurement, little has been said about why inequality should be measured. Yet the proper index of inequality for a given use should presumptively depend on the reasons for measuring it; hence, the question of why we should measure inequality ought to be addressed explicitly.

This essay examines this question and advances two claims.<sup>2</sup> First, regarding normative measures of inequality — those designed to have direct policy relevance, as aspects of overall social welfare — it is suggested that there is little need to measure inequality per se. Although inequality is an aspect of overall social welfare, it is best to measure welfare directly; measuring inequality requires additional effort and yields no return. Second, regarding descriptive measures of inequality — such as may be used in regressions relating inequality and growth — measurement often is necessary. But what counts as a good measure depends on the economic theory and empirical facts in particular contexts, not (necessarily) on the properties and axioms that have generally been proposed for measures of inequality.

These two claims are elaborated in sections 2 and 3, respectively. Then, in section 4, similar points are made with respect to the measurement of poverty, a subject that has also received substantial attention.<sup>3</sup> (And in the notes, brief remarks applying the analysis to the measurement of progressivity and of redistribution are offered.<sup>4</sup>)

## 2. Normative Measures of Inequality

### 2.1. Motivation

A principle motivation for inequality measurement is normative, to guide policy. Thus, if it can be shown that inequality has been increasing, many would argue that there is a *prima facie*

---

<sup>1</sup>See also, for example, Cowell (1995, 2000) and Lambert (2001).

<sup>2</sup>Throughout this essay, problems regarding whether income is a good proxy for welfare, units of measurement (individual versus family versus household), the choice of a present or lifetime perspective, and the like will be set aside in order to focus on more basic questions. Obviously, the reasons for measuring inequality (and poverty, progressivity, or redistribution) will also bear on how these other matters should be addressed.

<sup>3</sup>See, for example, Atkinson (1987), Clark, Hemming, and Ulph (1981), Kakwani (1999), Lambert (2001), Ravallion (1994a, 1994b), Ruggles (1990), Seidl (1988), Sen (1976, 1983), and Watts (1968).

<sup>4</sup>See notes [10](#) and [16](#). On such measurement, see, for example, Blackorby and Donaldson (1984), Fellman (1976), Jakobsson (1976), Kakwani (1977), Lambert (2001), Musgrave and Thin (1948), and Suits (1977).

case for doing something about it, notably, increasing the extent of redistribution in the tax and transfer system. Or, in considering a particular policy proposal — for a new highway, for research subsidies, for raising cigarette taxes — that the policy may reduce or increase the extent of inequality is taken to be an argument for or against the policy. Accordingly, it seems appropriate to attempt to measure inequality as an input to policymaking.<sup>5</sup>

It has now become familiar that sometimes, though not often, it is easy to say (it is wholly unambiguous) whether inequality is higher or lower. For a given total income, if for any given percent of the poorest in society, they have more income in one situation than in another — that is, if there is Lorenz dominance — then inequality is lower in the former situation. Equivalently, if one distribution can be obtained from another by a series of transfers that all flow from richer to poorer individuals, that one distribution exhibits a lower level of inequality. These statements can usefully be presented in another manner: For any of a standard class of social welfare functions (SWF's), inequality is said to be unambiguously lower (holding total income constant) if and only if social welfare for that distribution is higher regardless of the degree of aversion to inequality.<sup>6</sup>

These results concerning inequality rankings, however, are limited in two important respects. First, what results is only an ordering. And an ordering alone is of limited use in policymaking for there is rarely available a redistributive free lunch, that is, a reform that holds total income constant and unambiguously reduces inequality. Usually, there are tradeoffs between equality and total income, which requires giving quantitative meaning to differences in inequality, and this in turn requires specifying a particular SWF (on which more below).

Second — and what has been the main motivation for further work on inequality measurement — even if one only seeks an ordering with respect to inequality, most comparisons of distributions of income do not involve such a simple relationship as involved with Lorenz dominance. For example, “flattening” the income tax may result in the middle class paying more with both the rich and the poor paying less.<sup>7</sup> Other reforms or changes over time may have more complex redistributive patterns. Furthermore, for a given income class (say, the poor), some may become better off and others worse off.

---

<sup>5</sup>A normative purpose also seems to underlie some axiomatic treatments of inequality measurement because the basis of appeal for many proposed axioms is related to the manner in which inequality may plausibly affect social welfare. This is especially apparent when stronger axioms are introduced, as is necessary to narrow the wide range of otherwise admissible indexes. To illustrate, consider the axiom that holds that a given difference in income should be weighted more heavily the poorer are the individuals involved.

<sup>6</sup>See Atkinson (1970), Dasgupta, Sen, and Starrett (1973), Kolm (1969), Rothschild and Stiglitz (1973), and Sen (1973). See also Shorrocks (1983) for an extension to distributions with different means, which involves the use of what are referred to as generalized Lorenz curves (the ordinary Lorenz curve multiplied by mean income), and Davies and Hoy (1995) for an extension involving some cases in which Lorenz curves intersect. The literature on equivalences between properties of the Lorenz curve, social welfare assessments, and certain other criteria draws on the close analytical relationship between assessment of inequality aversion and of risk aversion. This relationship is not surprising when social welfare is taken to be a concave function of each possible level of income, weighted by the density of individuals at that level of income, and utility is taken to be a concave function of each possible level of income, weighted by a probability density function. Likewise, the equally distributed equivalent level of income,  $y_e$ , described below, corresponds to the certainty equivalent. See also note 9.

<sup>7</sup>See, for example, Davies and Hoy (2002).

For the most part, there is nothing simple and unambiguous that can be said about whether inequality rises or falls in such cases, although there are distributive effects that are of consequence to overall social welfare. Another way to express the point is to observe that, holding total income constant, whether social welfare rises or falls due to many distributive changes will depend on the SWF and, in particular, on the degree of inequality aversion. For example, if the income of one individual at the median of the distribution falls by \$2 and the income of each of two individuals, one at 75% of the median income and one at twice the median income, rises by \$1, welfare will be higher (and inequality thus might be said to be lower) if the degree of inequality aversion is above some critical point, but welfare would be lower (and inequality deemed higher) if the degree of inequality aversion is below that critical point.

Because it is recognized that, in general, little can be said about inequality that is normatively relevant without reference to a particular SWF — and, relatedly, that social welfare may rise or fall for a given change in a descriptive measure like the Gini coefficient (even holding mean income constant) — it has been proposed since Dalton (1920) to develop and employ measures of inequality that are based explicitly on SWF's.

## 2.2. Measures

The most widely explored normative measure of inequality is due to Atkinson (1970).<sup>8</sup> His index of inequality is a number between zero and one that indicates the percentage of actual income that could be discarded while leaving total social welfare unchanged if the remaining income were distributed equally. Thus, if the existing distribution is one that is fully equal, the index will be 0. The more unequal is the distribution, and the more averse society is to inequality, the greater will be the value of the index.

This index is constructed in a four-step process. A description of each of these steps is essential for present purposes because once the four steps are understood, it will be straightforward to explain why such normative measures are unlikely to be very useful.

*Step 1.* Choose an SWF. Atkinson, and many following him (as well as Mirrlees (1971) in his exploration of the optimal income tax problem), often use an SWF of the following form:

$$(1) \quad W = \int \frac{y^{1-\epsilon}}{1-\epsilon} f(y) dy, \quad \epsilon \neq 1$$

In this expression,  $y$  refers to an individual's income,  $f(y)$  is the density of the income distribution, and  $\epsilon$  is an inequality aversion parameter. A choice of  $\epsilon$  equal to zero corresponds to no aversion to inequality; social welfare,  $W$ , equals mean income. The greater is  $\epsilon$ , the greater

---

<sup>8</sup>A similar approach, which at the time immediately following its publication was not as widely known, is offered by Kolm (1969). See also Sen (1973) for further discussion, and Jorgenson and Slesnick (1984) for an empirical implementation of the normative approach to inequality measurement.

is society's aversion to inequality.<sup>9</sup> In the special case in which  $\epsilon = 1$ , the SWF instead takes the following form:

$$(2) \quad W = \int \ln y f(y) dy, \quad \epsilon = 1$$

Choosing an SWF thus entails the choice of a functional form and, in this instance, choosing an inequality aversion parameter  $\epsilon$ .

*Step 2.* Using information on the density of the distribution of income  $f(y)$ , calculate the level of social welfare using the SWF chosen in step 1.

*Step 3.* Find the level of income,  $y_e$  — referred to as the equally distributed equivalent level of income — such that, if every individual had income of  $y_e$  instead of his actual income, social welfare would be unchanged. Thus, if  $\epsilon \leq 1$ , we can use the first formulation for  $W$  in expression (1):

$$(3) \quad \frac{y_e^{1-\epsilon}}{1-\epsilon} = W = \int \frac{y^{1-\epsilon}}{1-\epsilon} f(y) dy, \quad \text{or}$$

$$y_e = \left[ \int y^{1-\epsilon} f(y) dy \right]^{\frac{1}{1-\epsilon}}$$

Likewise, one can define  $y_e$  for the case in which  $\epsilon = 1$ , using expression (2).

*Step 4.* Let  $y_\mu$  refer to mean income. Construct the index of inequality,  $I$ , as follows:

$$(4) \quad I = 1 - \frac{y_e}{y_\mu}$$

The second term on the right side of expression (4), the ratio of equally distributed equivalent income to mean income, indicates what fraction of actual income would be needed to produce the existing level of social welfare if that income were instead equally distributed. Therefore, the index of inequality,  $I$ , indicates the fraction of income that is “wasted” in the sense that it

---

<sup>9</sup>Because  $W$  is expressed as a function of individuals' incomes rather than of their utilities,  $\epsilon$  has two interpretations: as individuals' (common) risk aversion coefficient in a constant-relative-risk-aversion utility function or as society's degree of aversion to inequality, independent of individuals' degree of risk aversion — or as a combination (individuals may be risk averse, and society may also be averse to inequality in utility levels).

could be destroyed while leaving social welfare unchanged if only the remaining income were equally distributed.

*Remarks.* Using  $I$ , it is clear that any two income distributions can be unambiguously ranked in terms of inequality. This is true regardless of their shapes (including whether, where, or how many times their Lorenz curves intersect). Likewise, rankings can be obtained for distributions with different means, whereas the possibility of doing this even using so-called generalized Lorenz curves is limited. The distributions being compared could be two distributions for a given society at two different points in time, distributions for different societies, or distributions before and after implementation of a proposed reform.

Two further observations are in order. First, it is clear from this four-step procedure that one must choose a specific SWF in order to define the inequality index  $I$ . This is because, in defining equally distributed equivalent income,  $y_e$ , in step 3, one needs to know the actual level of social welfare as well as the form of the SWF (to implicitly determine  $y_e$ ).

Second, the inequality index  $I$  does not allow one to rank social welfare in societies or to choose policies. It does provide a complete ordering, but the ordering is over inequality. Social welfare obviously depends on more than inequality; in particular, it also depends on the mean income. In the present context, knowing  $I$  and  $y_e$  is sufficient to determine a social welfare ranking (for a specified SWF), as should be obvious from the foregoing: Given  $I$  and  $y_e$ , one can determine  $y_e$  (reversing step 4); given  $y_e$ , one can determine the level of social welfare  $W$  (reversing step 3).

### 2.3. Usefulness

In light of the manner in which normative indexes of inequality, like  $I$ , are derived, it is hard to see how such measures can be useful. Recall why it is suggested that one would want a normative measure of inequality in the first place: to rank societies or to assess policies. But an inequality index gives an analyst or a policymaker only part of the information needed to accomplish this. As just explained, the information must be combined with other information — in the case of  $I$ , mean income is required — to provide the necessary assessment.

Moreover, it is unclear how a normative inequality index can help to provide even a partial assessment. Often, partial information provides partial illumination, and in a world where complete information is hard to come by, partial information should be welcomed. In addition, we often find that different types of information must be derived from different sources and then combined to provide an ultimate assessment. But it turns out that this is not true of normative inequality indexes for a simple reason: One must first undertake a complete, bottom-line analysis of social welfare as a *prerequisite* to measuring inequality.

This is apparent from the four-step process outlined above. Only at the conclusion of step 4 does one have a measure of inequality. But at the end of step 2, one had a complete social welfare assessment. Steps 3 and 4 involved further manipulations to, in a sense, remove part of the information contained in that assessment.

In essence, steps 3 and 4 decompose a measure of social welfare into two components, mean income and inequality. A measure of inequality might seem to be useful if analysts, policymakers, or the public disagreed about the proper SWF. The problem is that normative inequality indexes do nothing to resolve, or to avoid the need to resolve, such disagreement. Rather, step 1 requires that one begin by choosing a particular SWF; any normative inequality index, like  $I$ , is defined by reference to this SWF.

One solution would be for an analyst to present inequality measures based on some consensus SWF (perhaps for expression (1) with  $\alpha$  equal to a value near 2?), if such a consensus could be reached. Or, by analogy to what is often done in the optimal income tax literature, analysts could display inequality measures for a range of values of  $\alpha$  (say,  $\alpha$  equal to 1, 2, 3, and in the limit as  $\alpha$  approaches 4). Yet it remains the case that knowing  $I$  is insufficient to make social welfare assessments. One also needs to know mean income. Now, this too could be displayed. But further calculation is still required, so this approach seems inferior to just reporting values of social welfare for different SWF's, that is, in the present context, for different values of  $\alpha$ . And since social welfare requires less effort to calculate than  $I$ , and can be reported in less space, it is hard to see why computing and reporting  $I$  instead makes sense.

*Remark on grading.* Economists seem compelled — perhaps because of demand from others — to grade economies' performance with respect to inequality. Thus, OECD countries or some other group of nations might be ranked by the degree of inequality. Or one might examine how inequality changes over time to see how well a country is doing or whether policies in a given time period were supportive or perverse. Implicit in such exercises, a common connotation even if not a denotation, is the idea that countries with low grades should be doing something different, probably increasing the extent of redistribution or otherwise imitating the policies of those with higher scores. Or, for those more inclined to believe that modern societies go overboard with redistribution, high scores would be seen as indicating excess, suggesting a need to trim the welfare state.

Of course, none of these conclusions follow. Because inequality is not equivalent to social welfare, inequality rankings do not really indicate which countries really are doing better or worse. For example, a country with low inequality may have implemented effective policies aimed at the poor or may have destroyed the incentives and wealth of the upper classes, to the detriment of the poor. If one reported social welfare measures instead, one would know more. Focusing on inequality rather than welfare obscures the situation.

Furthermore, even with welfare rankings, the implied policies may still not be sensible. Countries differ in enough ways (natural resources, distribution of talent, infrastructure, quality of government institutions) that they cannot simply adopt the policies of others and expect to fare as well. A proposed redistributive policy, from a given policy status quo, could raise social welfare in one country and reduce it in another. Every polity instead needs to determine which policies, in its circumstances, would be most effective in redistributing income and then ask the further question whether implementing them would raise or lower social welfare. Proposals to reduce redistribution must similarly be evaluated. These assessments need the full social welfare information generated in step 2 and cannot be effectively guided by the partial and potentially misleading information in normative inequality indexes generated in step 4.



In sum, as attractive as issuing inequality “report cards” might be, whether to economists or to politicians looking for effective rhetoric to support their policy precommitments, it is difficult to justify such grading procedures.<sup>10</sup> Perhaps inequality scores are more persuasive than social welfare measures because the latter *seem* more controversial. That is, inequality scores appear to be objective indicators whereas social welfare measures obviously incorporate value judgments. However, it this impression is obviously mistaken given the origin of normative inequality indexes in SWF’s. (At best, reporting inequality scores might be seen as a corrective in light of the common practice of presenting comparisons of GDP per capita or measures of the inefficiency of redistributive policies, both of which ignore distributive effects that may be important in assessing social welfare. But this use is a rather limited one, itself inferior to reporting social welfare measures.)

### 3. Descriptive Measures of Inequality

Many measures of inequality, including some of the earliest and most widely used ones like the Gini coefficient, are understood as primarily descriptive.<sup>11</sup> Generically, one can identify at least two types of potential uses for descriptive measures of inequality, as explanatory variables and as dependent variables in understanding economic relationships. A priori, one would expect a combination of the pertinent theory and existing empirical evidence to determine what sort of inequality index, if any, is best for a particular application. Consider some examples.

Suppose that one is attempting to explain how a given distribution of talent combined with a particular technology translates into a distribution of labor income. For this purpose, it will be necessary to describe the income distribution. For most analysis, however, it is not clear that a single summary measure of the degree of inequality in the resulting income distribution is necessary. Rather, most models and empirical analysis will involve explaining the distribution in its entirety.<sup>12</sup>

Next, consider the relationship between income inequality and growth. Much of the literature on this subject measures inequality using the Gini coefficient, probably due to custom and convenience.<sup>13</sup> Once such a relationship is discovered, or hypothesized as a result of

---

<sup>10</sup>This conclusion as well as many other points raised in this section seem largely applicable to measures of progressivity and of redistribution. See the literature in note 4. Formally, this is suggested by close connection between results regarding Lorenz dominance for inequality measurement and those for the definitions and measurement of progressivity and redistribution. Also, when the literature addresses normative measurement explicitly, it uses an SWF approach and often draws on Atkinson (1970).

<sup>11</sup>For most descriptive measures, including the Gini coefficient, it is possible to state an SWF that underlies it. The point is that such an SWF is not generally seen as providing the justification for the measure, and often the implicit SWF is not one that is normatively compelling. On the Gini coefficient, see, for example, Atkinson (1970), Dasgupta, Sen, and Starrett (1973), Lambert (2001), Newbery (1970), Rothschild and Stiglitz (1973), Sheshinski (1972) and Yitzhaki (1979).

<sup>12</sup>See, for example, the survey by Neal and Rosen (2000) and the collection of essays in Welch (2001).

<sup>13</sup>See, for example, Barro (1999) and Forbes (2000). Recent work on this question often draws on the improved data set produced by Deininger and Squire (1996), which itself reports the Gini coefficient. They also report income shares by quintile wherever possible, and their analysis reveals that Gini coefficients often obscure important differences in distributions (for example, in countries with very different Gini coefficients, the income shares of the poorest quintile

theoretical analysis, however, further exploration is in order. In working with the data, different inequality measures can be tried to see which provides the best fit. In theoretical work, one would attempt to identify the source of the relationship, if any, between inequality and growth. Perhaps it derives from individuals being employed in different sectors of the economy (as suggested by Kuznets); maybe it is due to different marginal propensities to consume at different income levels, or differential access to capital for entrepreneurship, or different political forces determining macroeconomic policies. Different models would suggest different inequality measures. Or sometimes none at all. (For example, if the relationship involves movement between sectors of the economy, measures of the size of the sectors, differences in wages between the sectors, and rates of movement between the sectors would seem more pertinent than a summary measure of income inequality.)

Some work in political economy takes account of income inequality. Suppose, for example, that a median voter model is to be tested. Different distributions of income may imply that the median voter is a different distance from the mean level of income, which could translate into different policy preferences regarding, say, the shape of the income tax schedule. Greater inequality according to most measures would ordinarily be associated with a lower median (for a given mean). Nevertheless, there seems to be little reason to compute an inequality index that is related to median income rather than to use median income directly.<sup>14</sup>

Inequality is also thought to be related to crime. Depending on the channel hypothesized to explain this relationship, different inequality measures may be appropriate. If it is the lack of more attractive economic opportunities that fuel crime, then information about the bottom end of the distribution will be highly relevant and the rest may be largely irrelevant. (Also, mean income would seem quite important to know; hence, the common suggestion that measures of inequality should be invariant to multiplying the distribution of income by a constant seems inapt in this setting.) By contrast, if envy of the rich or having very attractive targets is a major contributor to crime, then concentration of income at the top end of the distribution may be highly relevant.<sup>15</sup>

Although the foregoing discussions are very brief and unavoidably superficial, they seem to make a strong *prima facie* case that descriptive inequality indexes should be chosen from the ground up. The question of what is the most appropriate measure of inequality for a particular purpose needs to be answered by a combination of theory and empirical testing in the specific field. Thus, there can be no presumption that there exists a common, ideal descriptive measure of inequality; quite to the contrary. No inequality measure can be said, in a vacuum, to be best for descriptive purposes. In addition, whether a measure satisfies one or another set of axioms

---

might be similar). Indeed, as indicated by Bénabou's (1996) survey, it is most common to use either Gini coefficients or some quintile information, possibly transformed.

<sup>14</sup>This assumes that data on the entire distribution of income or on the median is available. For example, in examining how distributive politics may affect growth, Alesina and Rodrik (1994) develop a model in which the situation of the median voter is most relevant but their empirical test has to rely on reported Gini coefficients from the work of other researchers. Some others use quintile information for the third quintile. See Bénabou (1996).

<sup>15</sup>Fajnzylber, Lederman, and Loayza (2002), in examining the link between inequality and violent crime across countries, measure inequality using the Gini coefficient, the ratio of income shares of the first to the fifth quintile, an index of polarization, and also a measure of inequality in educational attainment. Each measure yielded similar results.

that seem aesthetically or normatively appealing is not likely to be particularly important.<sup>16</sup>

Nevertheless, there is some value in having common measures. They facilitate discussion and the comparison of results among researchers, including results over different time periods and between fields. However, which common measures would be best still must be determined inductively, through experience.<sup>17</sup> Moreover, to the extent that some common agreement emerges, those investigating specific subjects should continually assess whether such measures are most suitable for their own purposes, and some (perhaps many) will find that they need to deviate from convention.

#### 4. Poverty Measures

The discussion in sections 2 and 3 regarding normative and descriptive measures of inequality are largely applicable to normative and descriptive measures of poverty and hence will not be repeated. This application is rather straightforward given the familiar point that poverty measures are generally a sort of inequality measure that confines attention to a specified bottom slice of the income distribution.

Indeed, compared to inequality measures, poverty measures have the added problem of being arbitrary.<sup>18</sup> They ignore most of the income distribution and often give substantial weight to an individual being at or just below the poverty line whereas no weight is given to those slightly above the poverty line. The commonly used headcount measure — the number of individuals below the poverty line — seems especially thin in terms of information provided by comparison to inequality measures. Furthermore, however the measures are constructed, the choice of the poverty line is itself problematic for both normative and descriptive uses. And, following the above analysis, even when a poverty index is helpful, one would not expect the same form of poverty index — or the same poverty line — to be correct for every application.

It should also be observed that — even more so than with inequality measures — normative implications are often assumed even if the analysis purports to be purely descriptive. To say that a program reduces poverty or that a regime has been responsible for steadily increasing poverty is often understood as praise or criticism.<sup>19</sup>

---

<sup>16</sup>Similar conclusions seem applicable to the literature cited in note 4 on the measurement of progressivity and of redistribution. For example, progressivity or redistribution might be relevant in a political economy model, but it might be more relevant to measure the taxes paid or received by particular groups (such as voters with income near the median) rather than some aggregate. The parallel is reinforced by the fact that certain aggregate measures in the progressivity and redistribution measurement literatures are closely related to inequality indexes. For example, the extent of redistribution from taxation is sometimes measured as the difference between the pre-tax and post-tax Gini coefficient.

<sup>17</sup>Cowell (2000, p. 150) suggests that the Gini coefficient is overwhelmingly prominent, despite its apparent shortcomings, because of cultural inertia and the ability to visualize its interpretation from the Lorenz curve.

<sup>18</sup>Many of the problems mentioned here that are specific to poverty measurement are familiar in the literature, whereas the questions of why poverty should be measured in the first place and how the answers relate to how measurement should be undertaken have received little attention.

<sup>19</sup>Ruggles (1990, pp. 1-3) is one of the few to address directly the purposes of poverty measurement, and her reasons are normative. Nevertheless, her book focuses almost entirely on the information needed to set a poverty line, not addressing further matters pertinent to normative poverty indexes, notably, the importance of falling various distances below the poverty line. Some of the analysis seems motivated by the desire to assess the actual situations of those with

A further point about normative uses is that, if the chosen SWF is sufficiently concave in income, social welfare assessments will often depend substantially on how policies affect poverty.<sup>20</sup> Hence the seemingly arbitrary focus on the bottom end of the income distribution implicit in normative poverty indexes is not entirely without justification. Nevertheless, such indexes may be rather sensitive to the choice of the poverty line; policies that have significant effects on poverty will often have significant effects on individuals just above the poverty line. And it cannot be taken for granted that effects on others will be negligible. Hence, there is a strong case for considering social welfare as a whole. Moreover, with normative poverty indexes, like with normative inequality indexes, one indeed must first specify an SWF, but then one could determine overall social welfare in a straightforward manner, without ever having to choose a poverty line and compute a poverty index.

There is much argument about whether measures of poverty should be relative, absolute, or some combination of the two. But it seems unnecessary to make this choice for normative purposes when it is more straightforward, and also more complete and less arbitrary, to measure social welfare directly. Related, it is unclear why as a normative matter it is important to distinguish the problem of poverty from that of inequality at all.<sup>21</sup>

It is sometimes suggested that a poverty line should have direct significance for program design, such as when it is argued that everyone in poverty should be exempt from the income tax or eligible for various welfare programs. But the optimal design of the income tax and transfer programs needs to be done directly. For program eligibility, it usually is not optimal to have a simple cutoff, but rather to phase program coverage out as income increases, and this is often done. The best phase-out range may well vary from one program to another and, for any given program, the beginning or the end of that range will correspond with any postulated poverty line only by chance. (One could define a poverty line by reference to the income level where, say, individuals just begin paying tax under the optimally designed income tax, but then the poverty line is merely another term for the breakeven point; it has no independent normative significance.)

Descriptive poverty measures are subject to the analysis of descriptive inequality indexes in section 3. For some theories of growth, voting behavior, crime, and so forth, it may be that a measure of poverty would be useful, which is to say that the number and income level of those at the bottom of the income distribution might matter a great deal, whereas others (including those

---

low incomes, a descriptive exercise that has clear normative relevance (but not one that requires adopting any particular poverty index).

<sup>20</sup>For example, if  $\alpha = 2$ , the social weight on an additional dollar of income to a person with \$10,000 is one hundred times higher than the weight on an additional dollar to a person with \$100,000. Ravallion (1994b) finds that, for developing countries, there is a very high correlation between poverty indexes and an Atkinson-type normative inequality index when the inequality-aversion parameter in the latter is 2 or higher.

<sup>21</sup>Hence, Yitzhaki (2002) entitles his article, "Do we need a separate poverty measurement?" Lewis and Ulph (1988) argue that indeed there is no point in distinguishing poverty from inequality unless there are indivisibilities or regions of increasing returns in individuals' utility functions, and they develop poverty measures on this basis. It seems unlikely, however, that these phenomena would be significant for most individuals at a given income level, especially when one allows for heterogeneity and individuals' ability to spread purchases or payments for durable goods over time. But if there was a range of incomes where there was significant increasing returns, the implication offered by Lewis and Ulph is that it may be optimal to favor those in such ranges over those with lower incomes.

with modestly higher incomes) might matter relatively little. It is not clear how often this would be the case, but in any event answers to the questions of when a poverty measure is useful and of what poverty measure to use are best determined through particular applications, not by a priori analysis.

Perhaps more with poverty measures than with inequality measures, there seems to be a strong demand for them by the public and by policymakers. In part, this may reflect habit, for such numbers have long been reported and publicized, whereas inequality indexes are understood primarily by economists and some other policy analysts, and social welfare measures may seem even more obscure. Some indicators — even noisy, incomplete, and somewhat arbitrary ones — are better than none. And, given the foregoing point that the effects of some policies on social welfare will largely be determined by their effects on the poor, poverty measures may have some value, though more for communication with the public than for research or policy analysis.

Two observations regarding this point are in order. First, for this sort of use, consistency may be the most important consideration since it is changes over time that are given the greatest attention. Thus, it may matter less whether the current poverty line is “correct” in some sense than whether it is maintained over time. (Accuracy in cost-of-living adjustments, no easy task, would thus be one of the more important considerations.) Second, this use does appear to be normative; poverty levels are invoked to assess specific policies and to evaluate government regimes. Hence, if the current highly crude headcount method is to be changed, moves in the direction of a measure based on an SWF would seem appropriate. Of course, this requires some agreement on the correct SWF. Also, normative poverty measures may be more opaque than simpler descriptive measures. However, it seems more appropriate to optimize the tradeoff between good normative measurement and public comprehensibility than, say, to choose the measure that is implied by what seem to mathematically trained analysts to be the most compelling axioms.

## **5. Conclusion**

This essay has asked questions that receive little attention in the measurement literature: Why measure inequality? Or poverty? It offers some preliminary thoughts regarding the answers. For normative purposes, direct measures of social welfare are superior in addressing the relevant questions and also are as easy or easier to derive than normative measures of inequality or poverty. For descriptive purposes, inequality and poverty indexes will sometimes be useful, but when this is so and which measures are most appropriate will depend on the context and can only be determined by examining the theoretical analysis and empirical results in each particular field of application. At a minimum, literature on the measurement of inequality and poverty should devote more attention to the purposes of measurement since whether and how measurement should be undertaken is likely to depend on the reasons for doing so.

## References

- Alesina, A., Rodrik, D. 1994. Distributive politics and economic growth. *Quarterly Journal of Economics* 109, 465-490.
- Atkinson, A.B. 1970. On the measurement of inequality. *Journal of Economic Theory* 2, 244-263.
- Atkinson, A.B. 1987. On the measurement of poverty. *Econometrica* 55, 749-764.
- Atkinson, A.B., Bourguignon, F., eds. 2000. *Handbook of Income Distribution*, vol. 1. Elsevier Science, B.V., Amsterdam.
- Barro, R. 1999. Inequality, growth, and investment. NBER Working Paper 7038.
- Bénabou, R. 1996. Inequality and growth. In: Bernanke, B.S., Rotemberg, J.J. (eds.), *NBER Macroeconomics Annual 1996*. MIT Press, Cambridge, 11-74.
- Blackorby, C., Donaldson, D. 1984. Ethical social index numbers and the measurement of effective tax/benefit progressivity. *Canadian Journal of Economics* 17, 683-694.
- Clark, S., Hemming, R., Ulph, D. 1981. On Indices for the Measurement of Poverty. *Economic Journal* 91, 515-526.
- Cowell, F.A. 1995. *Measuring Inequality*. 2nd ed. Harvester Wheatsheaf, Hemel Hempstead.
- Cowell, F.A. 2000. Measurement of inequality. In: Atkinson, Bourguignon (eds.), *Handbook of Income Distribution*, vol. 1. Elsevier Science, B.V., Amsterdam, 87-166.
- Dalton, H. 1920. The measurement of the inequality of incomes. *Economic Journal* 30, 348-361.
- Dasgupta, P.S., Sen, A.K., Starrett, D.A. 1973. Notes on the measurement of inequality. *Journal of Economic Theory* 6, 180-187.
- Davies, J.B., Hoy, M. 1995. Making inequality comparisons when Lorenz curves intersect. *American Economic Review* 85, 980-986.
- Davies, J.B., Hoy, M. 2002. Flat rate taxes and inequality measurement. *Journal of Public Economics* 84, 33-46.
- Deininger, K., Squire, L. 1996. A new data set measuring income inequality. *World Bank Economic Review* 10, 565-591.
- Fajnzylber, P., Lederman, D., Loayza, N. 2002. Inequality and violent crime. *Journal of Law and Economics* 45, 1-40.
- Fellman, J. 1976. The effect of transformations of Lorenz curves. *Econometrica* 44, 823-824.
- Forbes, K.J. 2000. A reassessment of the relationship between inequality and growth. *American Economic Review* 90, 869-887.
- Jakobsson, U. 1976. On the measurement of the degree of progression. *Journal of Public Economics* 5, 161-168.
- Jorgenson, D.W., Slesnick, D.T. 1984. Inequality and the distribution of individual welfare. In: Basmann, R.L., Rhodes, G. (eds.), *Advances in Econometrics*, vol. 3. JAI Press, Greenwich, CT, 67-130.
- Kakwani, N. 1977. Applications of Lorenz curves in economic analysis. *Econometrica* 45, 719-727.
- Kakwani, N. 1999. Inequality, welfare and poverty: three interrelated phenomena. In: Silber, Jacques (ed.), *Handbook of Income Inequality Measurement*. Kluwer Academic Publishers, Boston, 599-628.

- Kolm, S.-C. 1969. The optimal production of social justice. In: Margolis, J., Guitton, H. (eds.), *Public Economics: An Analysis of Public Production and Consumption and their Relations to the Private Sectors*. Macmillan, London, 145-200.
- Lambert, P.J. 2001. *The Distribution and Redistribution of Income*. 3rd ed. Manchester University Press, Manchester.
- Lewis, G.W., Ulph, D.T. 1988. Poverty, inequality and welfare. *Economic Journal* 98 (supp.), 117-131.
- Mirrlees, J.A. 1971. An exploration in the theory of optimal income taxation. *Review of Economic Studies* 38, 175-208.
- Musgrave, R.A., Thin, T. 1948. Income tax progression, 1929-40. *Journal of Political Economy* 56, 498-514.
- Neal, D., Rosen, S. 2000. Theories of the distribution of earnings. In: Atkinson, A.B., Bourguignon, F. (eds.), *Handbook of Income Distribution*, vol. 1. Elsevier Science, B.V., Amsterdam, 379-427.
- Newbery, D. 1970. A theorem on the measurement of inequality. *Journal of Economic Theory* 2, 264-265.
- Ravallion, M. 1994a. *Poverty Comparisons*. Harwood Academic Publishers, Chur.
- Ravallion, M. 1994b. Measuring social welfare with and without poverty lines. *American Economic Association Papers and Proceedings* 84, 359-364.
- Rothschild, M., Stiglitz, J.E. 1973. Some further results on the measurement of inequality. *Journal of Economic Theory* 6, 188-204.
- Ruggles, P. 1990. *Drawing the Line: Alternative Poverty Measures and Their Implications for Public Policy*. Urban Institute Press, Washington, D.C.
- Seidl, C. 1988. Poverty measurement: a survey. In: Bös, D., Rose, M., Seidl, C. (eds.), *Welfare and Efficiency in Public Economics*. Springer-Verlag, Heidelberg, 71-147.
- Sen, A.K. 1973. *On Economic Inequality*. Clarendon Press, Oxford.
- Sen, A.K. 1976. Poverty: an ordinal approach to measurement. *Econometrica* 44, 219-231.
- Sen, A.K. 1983. Poor, relatively speaking. *Oxford Economic Papers* 35, 153-169.
- Sheshinski, E. 1972. Relation between a social welfare function and the Gini index of inequality. *Journal of Economic Theory* 4, 98-100.
- Shorrocks, A.F. 1983. Ranking income distributions. *Economica* 50, 3-17.
- Silber, Jacques, ed. 1999. *Handbook of Income Inequality Measurement*. Kluwer Academic Publishers, Boston.
- Suits, D.B. 1977. Measurement of tax progressivity. *American Economic Review* 67, 747-752.
- Watts, H. 1968. An economic definition of poverty. In: Moynihan, D.P. (ed.), *On Understanding Poverty*. Basic Books: New York, 316-329.
- Welch, F., ed. 2001. *The Causes and Consequences of Increasing Inequality*. University of Chicago Press, Chicago.
- Yitzhaki, S. 1979. Relative deprivation and the Gini coefficient. *Quarterly Journal of Economics* 93, 321-324.
- Yitzhaki, S. 2002. Do we need a separate poverty measurement? *European Journal of Political Economy* 18, 61-85.