MEASURES OF INEQUALITY *

PAUL D. ALLISON
Cornell University

American Sociological Review 1978, Vol. 43 (December):865–880

Although measures of inequality are increasingly used to compare nations, cities, and other social units, the properties of alternative measures have received little attention in the sociological literature. This paper considers both theoretical and methodological implications of several common measures of inequality. The Gini index is found to satisfy the basic criteria of scale invariance and the principle of transfers, but two other measures —the coefficent of variation and Theil's measure—are usually preferable. While none of these measures is strictly appropriate for interval-level data, valid comparisons can be made in special circumstances. The social welfare function is considered as an alternative approach for developing measures of inequality, and methods of estimation, testing, and decomposition are presented.

1. Defining Inequality

Although inequality has long been topic of intense interest to sociologists, few have bothered to carefully specify what they mean by the term. It is easy, of course, to distinguish perfect equality from a state of inequality. But given two different, unequal distributions of some social reward, how does one decide which distribution is the more unequal? The answer to this question would seem to be a prerequisite for any theory of the determinants and consequents of social inequality. Yet even Lenski's (1966) influential theory on the effects of economic surplus and democracy on inequality fails to include a definition of the dependent variable.

This lack of rigor created little difficulty so long as research on inequality emphasized the determinants of individual attainments. But recent efforts to test hypotheses explaining why some societies are less equal than others have necessitated the adoption of precise measures of inequality, such as the Gini index or the standard deviation (Jackman, 1974; Rubinson and Quinlan, 1977; Hewitt, 1977; Kelley and Klein, 1977). In the absence of clear criteria for choosing among the numerous measures of inequality, researchers have usually based their choice on convenience, familiarity, or on vague, methodological grounds. Nevertheless, the decision to rank one distribution as more unequal than another has theoretical as well as methodological implications. In fact, the choice of an inequality measure is properly regarded as a choice among alternative definitions of inequality rather than a choice among alternative ways of measuring a single theoretical construct.

This choice *can* make a difference. Although some have reported high correlations among different inequality measures (Alker and Russett, 1964), such correlations are relevant only for the particular populations and variables for which they are computed. And in one case of particular interest, it has been shown that the rank ordering of countries by income inequality can differ substantially with different measures of inequality (Atkinson, 1970; Yntema, 1933).

A major exception to the failure of sociological theorists to specify what they mean by greater and lesser inequality is the recent work of Peter Blau (1977a; 1977b). Blau assumes that inequality is a funda-

---

* Address all communications to: Paul D. Allison; SASS; 323 Uris Hall; Cornell University; Ithaca, NY 14853.

For helpful comments, I am indebted to Scott Feld, Lowell Hargens, J. Scott Long, Barbara Reskin and Alice Trisdorfer. This paper was written while I was supported by a postdoctoral fellowship from the Social Science Research Council.

mental characteristic of all graduated social parameters (quantitative status variables), and proposes that it be conceptualized as "the average difference in status between any two pairs relative to the average status" (1977b). Noting that the Gini index is an appropriate algebraic specification of this concept, he proceeds to consider in detail how changes in the distribution affect the Gini index (1977a:56–69).

Although Blau's formulation is certainly an improvement over previous sociological work, he makes a fundamental error in suggesting that inequality measures can be meaningfully applied to any quantitative variable. He also fails to consider recent developments in economic theory which have shed much light on the implications of alternative measures of inequality. Drawing on this literature, I shall discuss the characteristics and relative merits of some common measures of inequality. Section 2 introduces several measures of inequality and considers whether they satisfy basic theoretical criteria. Section 3 examines Blau's assumption that inequality measures can be meaningfully used with interval scales. The assumption is shown to be generally false, except under special conditions. Section 4 shows how various measures are related to the Lorenz curve. This sets the stage for section 5 which discusses the social welfare approach to inequality that is favored by some economists. Finally, section 6 reports basic results on estimation and hypothesis testing with sample data. Since much of the literature is concerned with income inequality, I will assume that income is the variable of interest. With a few notable exceptions, most of the results also can be applied to other quanitative variables.

2. *Basic Criteria for Measures of Inequality*

Suppose we have a population of n people, and each person receives annual income x_i with $i = 1, \ldots, n$. It will be convenient for some applications to assume that the incomes are arranged in ascending order so that $x_1 \leq x_2 \leq \ldots \leq x_n$. The problem is to get a single measure which characterizes every possible set of x_i's in terms of inequality. What characteristics should such a measure possess? At the very least, it should be zero when all individuals have identical incomes, and should have a positive value when two or more individuals differ. These conditions are satisfied by almost all the common measures of inequality including such familiar measures of dispersion as the range and the variance.

Many of these measures are ruled out by the criterion of *scale invariance,* which requires that multiplying everyone's income by a constant leaves the degree of inequality unchanged. This eliminates the variance, for example, since it quadruples when everyone's income is doubled. Although scale invariance is widely accepted as a desirable property of inequality measures, the justification is rarely made explicit. And since a few sociologists and economists have questioned this criterion,[1] it seems worthwhile to examine some of the arguments in its favor.

Since a change of units in which a variable is measured does not constitute any real change in the distribution of that variable, it seems appropriate to ask that a measure of inequality be invariant to such changes. Thus inequality of income should not depend on whether income is measured in dollars or yen. Obviously, scale invariant measures have this property. There is an enormous payoff in convenience since it becomes unnecessary to adjust for inflation or to deal with currency conversions.

It is less clear that *real* proportionate increases in everyone's income should leave inequality unchanged since, in absolute amounts, the rich benefit more than the poor. But note that increasing everyone's income by, say, 10% leaves the ratios of all pairs of incomes unchanged. Equivalently, each individual's proportionate share of the total national income remains the same. Hence, it is reasonable to say that the *relative* differences among individuals have not been altered. From a variety of perspectives, it is desirable that measures of inequality respond to relative

[1] Kelley and Klein (1977) and Kolm (1976) argue that proportional increases in income represent an increase in inequality. Taking the opposite point of view, Dalton (1920) and Sen (1973) suggest that inequality should *decrease* when all incomes are increased proportionately.

rather than to absolute differences (Blau, 1977a:57–9). For example, Easterlin (1974) presents evidence that individuals' self-reported happiness depends less on their absolute incomes than on their relative positions in the income distribution.

It may seem like a big jump to require that inequality be comparable across completely different sorts of quantities. Yet, it is common to hear assertions like "power is more unequally distributed than wealth in nation A." Without scale invariance, such assertions are meaningless. With scale invariance, we can, for example, compare the inequality of nations' energy consumption with the inequality of population size.

A final argument is that scale invariant measures (at least all those we shall consider here) respond in an intuitively appealing manner when a positive constant is *added* to everyone's score; specifically, they decline. Consider a three-person society with incomes of \$5,000, \$15,000, and \$25,000. Clearly, these differences would be of great importance to the individuals involved. If each were given a million dollars, however, the remaining differences would become trivial, and it would be reasonable to say that inequality had declined. Yet, the standard deviation would stay exactly the same. For all these reasons, the remainder of this paper will deal only with scale invariant measures.

Most measures of dispersion can be converted into scale invariant measures of inequality by dividing by the mean or some function of the mean. The *coefficient of variation* V, for instance, is just the standard deviation divided by the mean:

$$V = \frac{\sigma}{\mu}. \tag{1}$$

Similarly, the *relative mean deviation,* also known as Schutz's (1951) coefficient, is defined as

$$D = \frac{\frac{1}{n} \sum_{i=1}^{n} |x_i - \mu|}{2\mu}, \tag{2}$$

where $|\,.\,|$ is the absolute value function.

Perhaps the most commonly used measure of inequality is the *Gini index,* G, which is usually defined in terms of the Lorenz curve (see section 4). However, the following equivalent definition makes it clear that the Gini index is a measure of dispersion divided by twice the mean:

$$G = \frac{\frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} |x_i - x_j|}{2\mu}. \tag{3}$$

The numerator of this expression, known as Gini's coefficient of mean difference (Kendall and Stuart, 1977: 48), is the average absolute difference between all pairs of individuals. An equivalent formula (Dasgupta et al., 1973) is more mathematically tractable and computationally convenient for individual-level data:

$$G = \frac{2}{\mu n^2} \sum_{i=1}^{n} i x_i - \frac{n+1}{n}. \tag{4}$$

Notice that the first term in (4) involves a weighted sum of all the scores, where the weight applied to each score is its rank in the distribution.

Based on information theory, Theil (1967: 92) proposed the measure

$$T = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{x_i}{\mu}\right) \log \left(\frac{x_i}{\mu}\right). \tag{5}$$

(Natural logarithms are used throughout this paper.) Simple algebra reduces this to a formula that is more computationally convenient and also reveals that T is a measure of dispersion divided by the mean:

$$T = \frac{\frac{1}{n} \sum_{i=1}^{n} x_i \log x_i - \mu \log \mu}{\mu}. \tag{6}$$

When $x_i = 0$, $x_i \log x_i$ is also defined to be 0.

One scale invariant measure that cannot be expressed as a ratio of a measure of dispersion to the mean is the *variance of the logarithms* L. This is obtained by taking the logarithm of each income and computing the variance of the transformed scores. Thus, if $z_i = \log x_i$ for all i,

$$L = \frac{1}{n} \sum_{i=1}^{n} (z_i - \bar{z})^2. \tag{7}$$

L is undefined when the distribution includes incomes of zero. Although many other measures of inequality have been proposed (Alker and Russett, 1964; Martin and Gray, 1971; Ray and Singer, 1973), they either fail to satisfy the preceding criteria, have very restricted applications, or are simple monotone functions of those defined here.

How is one to choose among these five scale-invariant measures? Considerable headway can be made by applying the *principle of transfers.* Dalton (1920) argued that measures of inequality ought to increase whenever income is transferred from a poorer person to a richer person, regardless of how poor or how rich or the amount of income transferred. Not only does this principle have considerable intuitive appeal but, as we shall see, it also has important relationships with the Lorenz curve and the social welfare approach to inequality measurement.

Two of the scale-invariant measures fail to satisfy the principle of transfers (Atkinson, 1970). The relative mean deviation is not affected by transfers between persons who are both below the mean or both above it, and it will therefore be dropped from further consideration. The variance of the logarithms responds appropriately to transfers at lower income levels. But at high income levels (greater than 2.718 times the geometric mean), it actually decreases with a transfer from a (relatively) poorer to a richer individual.[2] Although this is a serious limitation, the variance of the logarithms has very desirable inferential properties which will be discussed in section 6.

Sensitivity to transfers. Among those measures of inequality which satisfy the principle of transfers, there are important differences in sensitivity to transfers at different points on the scale (Atkinson, 1970). Suppose we transfer h dollars from a person with income x_i to another person with income x_j, where $x_i \leqq x_j$. All other incomes remain the same. Let V_1 and V_2 be the coefficients of variation before and after the transfer, respectively. It can be shown (Dalton, 1920) that

$$V_2^2 - V_1^2 = c\,h(x_j - x_i) + c\,h^2, \quad (8)$$

where c is positive and depends only on the mean and the number of observations (the same is true for c′ and c″ used below). This result says that V is equally sensitive to transfers at all income levels. Thus, a transfer of $100 from a person earning $5,000 to another earning $6,000 has the same impact as a transfer of $100 from a person earning $50,000 to another earning $51,000.

The Gini index is peculiar in that its sensitivity to transfers depends on individuals' ranks rather than their numeric scores. Using formula (4), it is easily shown that, for a transfer of h from x_i to x_j,

$$G_2 - G_1 = c'h(j - i), \quad (9)$$

where G_2 and G_1 are the values of the Gini index before and after the transfer, and i and j are the ranks of incomes x_i and x_j. This result says that the change in G depends on the number of individuals with incomes lower than x_j and higher than x_i. In the U.S. today, there are many more persons in the $10,000–11,000 interval than there are in the $50,000–51,000 interval. It follows that a transfer from one person earning $10,000 to another earning $11,000 will have more affect on G than an equal transfer from a person earning $50,000 to another earning $51,000. But there are also fewer people in the $4.000–5,000 interval than in the $10,000–11.000 interval. Thus, for a typically shaped income distribution, the Gini tends to be most sensitive to transfers around the middle of the distribution and least sensitive to transfers among the very rich or the very poor.

To get a simple expression for the effect of a transfer on Theil's measure T, it is necessary to use a limiting argument. Let ΔT be the change in T resulting from a transfer of h from x_i to x_j. As h goes to 0, the limiting expression for ΔT can be shown to be

$$\Delta T = c''h \log(x_j/x_i). \quad (10)$$

Whereas the change in V depended on the differences between the incomes, the change

[2] Creedy (1977) argues that the extent to which the variance of the logarithms violates the principle of transfers is very minor for most empirical distributions.

in T depends on the ratio of the incomes. As a consequence, transferring \$100 from a person earning \$5,000 to one earning \$6,000 has approximately the same effect on T as a transfer of \$100 from a person earning \$50,000 to another earning \$60,000. This change in T is approximately nine times as large as the change resulting from a transfer of \$100 from a person earning \$50,000 to another earning \$51,000. The lower the level of income, the more sensitive T is to transfers. (The variance of the logarithms is similar in this respect.)

Do these differences in sensitivity give any basis for choosing among G, V and T? If it is assumed that income has diminishing marginal utility, then a transfer of income among low income earners would be more consequential (for them) than a transfer of an equal amount among high earners. Since Theil's index T reflects such a difference, it would seem to have the advantage. On the other hand, when the variable of interest does not have diminishing marginal utility or when its utility or value (if any) is irrelevant to the analysis, then the "flat" response of the coefficient of variation might be desirable. Thus, depending on the context, V might be a good measure of inequality of age, city size, years of schooling, etc. The sensitivity of the Gini index depends on the shape of the frequency distribution, and it is not easy to think of cases where this property would be desirable in itself. Of course, since many distributions tend to be somewhat bell-shaped, this means that G tends to be most sensitive in the middle range. If one is most concerned about changes in inequality among middle income earners, then G might be a good choice.[3]

These differences in sensitivity can have important consequences. In comparing income inequality across nations, Atkinson (1970) concluded that measures which were most sensitive in the lower range of incomes tended to show relatively less inequality in developing nations and more inequality in developed nations. On the other hand, measures which were sensitive in the higher ranges tended to favor the developed nations. He suggested that this was due to the tendency for developing nations to have a large, homogeneous population of poor together with great inequality among the rich (cf. Blau, 1977a: 9).

Upper and lower bounds. Sometimes inequality measures are chosen on the basis of their upper and lower bounds, or how they respond to changes in the population size. In infinite populations the Gini index varies between 0 and 1 while the coefficient of variation and Theil's measure vary between 0 and infinity. These bounds should imply no preference, however, since simple (nonunique) transformations can produce any desired bounds. To make the coefficient of variation have an upper bound of 1, simply take $V/(V+1)$. To make the Gini index vary between minus and plus infinity, take its logit: $\log(G/(1-G))$. The logit transformation has some merit when the Gini index is to be used as a dependent variable in a regression analysis since it avoids the usual problems associated with bounded dependent variables (Nerlove and Press, 1973).

In finite populations or samples, the Gini index has an upper bound of $1-1/n$, the coefficient of variation has an upper bound of $\sqrt{n-1}$ and Theil's measure has an upper bound of $\log n$. For all three measures, the upper bound is reached when one individual has everything and everyone else has nothing. Theil (1967:92) has argued that this dependence on n is desirable since a two-person society in which one person has everything is, intuitively, less unequal than a million-person society in which one person has everything. Nevertheless, there may be some situations in which it is desirable to have a measure whose upper bound does not depend on n (Ray and Singer, 1973).[4] Again, however,

[3] Using simulated distributions, Champernowne (1974) reached very similar conclusions about these differing sensitivities. He found that the coefficient of variation was most sensitive to inequality of extreme wealth, the Gini index was most sensitive in the middle range, and the variance was most sensitive to inequality of extreme poverty. Contrary to the results here, however, he found that Theil's index behaved more like V than like L.

[4] Ray and Singer (1973) proposed a new index CON whose range did not depend on n. It can be shown, however, that CON is just the coefficient of variation divided by its upper bound.

it is a simple matter to divide V, G and T by their upper bounds to obtain measures that vary between 0 and 1 independently of n (Martin and Gray, 1971). Alternatively, it may be preferable to use $V^* = V/(\sqrt{n-1} - V)$ and $T^* = T/(\log n - T)$ to get measures which vary between 0 and infinity.

3. *Inequality and Interval Scales*

Discussions of the coefficient of variation and the Gini index in statistics texts frequently end with a warning that the measures are only appropriate for variables measured on a ratio scale, like income or age, which have a theoretically fixed zero point (Mueller et al., 1977). As Kendall and Stuart (1977:48) put it:

> The coefficients both suffer from the disadvantage of being very much affected by . . . the value of the mean measured from some arbitrary origin, and are not usually employed unless there is a natural origin of measurement or comparisons are being made between distributions with simliar origins.

Jencks et al. (1972:352) have noted that this tends to rule out the application of inequality measures to such common variables as IQ scores and occupational prestige which are, at best, measured on interval scales.

Blau (1977a; 1977b), choosing to disregard this caveat, however, claims that "the Gini index is substantively appropriate for any status criterion" including mathematical aptitudes and intelligence. The aim of this section is to show that Blau is only partly right. In general, one cannot make valid inferences about inequality using interval-level data. But under certain conditions, some of which can be easily checked, such inferences are valid.

The presentation can be simplified by first observing a special relationship between V and G. Kendall and Stuart (1977: 41) show that the variance can be expressed as one-half the average squared difference between all pairs of individuals:

$$\sigma^2 = \frac{1}{2n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} (x_i - x_j)^2. \qquad (11)$$

Using this fact with formulas (1) and (3), we can define a general family of inequality measures that includes both the Gini index and the coefficient of variation:

$$I = \frac{\left(\frac{1}{2n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} |x_i - x_j|^r \right)^{\frac{1}{r}}}{\mu}. \qquad (12)$$

When $r = 1$, I is the Gini index. When $r = 2$, I becomes the coefficient of variation.[5]

This result enables us to consider some common properties of G and V by referring only to I. For example, I is clearly scale invariant, which can be expressed as $I(cX) = I(X)$ where c is any constant.[6] If X is an interval-scale variable, its origin (zero point) may be arbitrarily changed by adding a constant to every score. The measure I is quite sensitive to changes in origin, however; it can be readily verified that

$$I(X + c) = \left(\frac{\mu}{\mu + c} \right) I(X). \qquad (13)$$

By appropriate choice of c, one can thus make $I(X + c)$ take on any real value, a quite undesirable result.[7] Clearly these measures of inequality are useless for characterizing a single distribution if measurements are only at the interval level.

What about comparisons of inequality for two distributions, each measured from the same, arbitrary origin? Could we not, for example, compare inequality of Duncan SEI scores for the occupations of blacks and whites? Although Kendall and

[5] It can shown that I satisfies the principle of transfers for all $r \geq 1$. For another family of inequality measures which includes the coefficient of variation and Theil's measure, but not the Gini index, see Gastwirth (1975).

[6] In this notation, X is an nx1 vector of scores. Depending on the context, c is either a scalar or an nx 1 vector of equal constants.

[7] The usual (0, 1) bounds on the Gini index apply only if all scores are nonnegative. Formulas (3) and (4) make it clear that G can be readily computed even if some or all scores are negative. In that case, however, the usual interpretation in terms of Lorenz curves is invalid (see section 4). Equation (13) also holds for D, the relative mean deviation, even though D is not a special case of (12). It also can be shown that T and L decrease when a positive constant is added to all scores and, hence, these measures also have arbitrary values when applied to interval scales.

Stuart suggest that such comparisons would be meaningful, it is easy to construct counterexamples. Suppose we wanted to compare the inequality of two distributions X and Y, both measured on the same interval scale. We observe $I(X) = .50$, $I(Y) = .60$, $\mu_x = 20$, and $\mu_y = 10$. Let us change the origin of measurement by adding 10 to each score. Using formula (13), we get $I(X + 10) = .33$ and $I(Y + 10) = .30$. Thus, different origins lead to different conclusions about the relative inequality of the two distributions.

An empirical example is provided by Blau's (1977a: 211) reanalysis of data reported by Hauser et al. (1975) who give the frequency distribution among 12 occupational categories of U.S. men at selected ages in 1952, 1962, and 1972. Using the Duncan SEI scores corresponding to the occupational categories, Blau calculated the Gini index for men aged 35–44 and noted the trend in inequality. The Gini values for the three periods were .353, .300, and .318—a clear decline in inequality over time. The corresponding mean scores were 31.8, 37.5, and 40.4. But since SEI scores are only interval level at best, we can freely change the origin by adding an arbitrary constant to all scores. A constant of 50 yields Gini values of .137, .141, and .142, a lower level of inequality but increasing rather than decreasing with time. Blau's prescription is clearly unacceptable.

Nevertheless, there are circumstances in which valid comparisons of inequality can be made for two distributions measured on the same interval scale. These cases arise when we know (or believe) that there is a nonnegative ratio scale underlying the interval scale that we observe. For instance, if we wish to measure social power, an interval scale may be the best that we can obtain. Yet zero power or powerlessness is clearly a meaningful concept and negative power is not. It is therefore reasonable to believe that there is a true ratio scale underlying our measurements of power. This is important because it puts constraints on the constant that can be added to or subtracted from the interval measure.

More generally, suppose X^* is a nonnegative ratio-level variable and X is an interval-level variable obtained by adding an unknown constant to X^*, i.e., $X = X^* + c$. If we observe a set of X's, we can be sure that c is no greater than the smallest value of X. Otherwise X^* would have negative values which are disallowed by assumption. Now suppose that we also have $Y = Y^* + c$, where Y^* is another nonnegative ratio-level variable measured on the same scale and origin as X^*. It follows that X and Y also have a common scale and origin. We observe $I(X) > I(Y)$, as well as the means μ_X and μ_Y. Under what circumstances can we conclude that $I(X^*) > I(Y^*)$? The answer is that $I(X^*) > I(Y^*)$ whenever there is no permissible value of c such that $I(X^*) \leq I(Y^*)$.

The following test for this condition is proved in the Appendix. First, make sure that the origin of X and Y is such that all values are nonnegative. Then calculate

$$c^* = \frac{\mu_X \, \mu_Y \, [I(X) - I(Y)]}{\mu_X \, I(X) - \mu_Y \, I(Y)}. \qquad (14)$$

If c^* is greater than the minimum value of either X or Y, we can conclude that $I(X^*) > I(Y^*)$. If c^* is less than or equal to the minimum values of both X and Y, no conclusion may be drawn. When $\mu_X I(X) = \mu_Y I(Y)$, c^* is undefined, but this also implies that $I(X^*) > I(Y^*)$. If the minimum values of X and Y are not readily available, then $c^* > \mu_X$ or $c^* > \mu_Y$ obviously insures that it is greater than one of the minimum values.

As an example, I apply this test to data on occupational status of blacks and whites reported by Farley (1977). In 1959, black males had a mean occupational status of 18.9 and a coefficient of variation of .815; in the same year, white males had a mean of 37.2 and a coefficient of variation of .608. The value of c^*, computed from formula (14), is -20.17. Since this is necessarily less than the minimum value for both blacks and whites, no conclusion about inequality is possible. In 1969, black females had a coefficient of variation of .773 and a mean of 28.6; white females had a coefficient of variation of .491 and a mean of 42.6. This yields $c^* = 288.43$ which is greater than the mean of either group. We conclude that in this sample black females were more unequal in occupational status than white females. This presumes, of course, that there is an underlying nonnegative ratio scale of occupational status, an assumption some may wish to question. It also says nothing about possible sampling errors.

4. Measures of Inequality and the Lorenz Curve

The Gini index, as noted earlier, is usually defined in terms of the Lorenz curve (e.g., Ray and Singer, 1973). It so happens that all scale-invariant measures of inequality which satisfy the principle of transfers have a simple relationship to the Lorenz curve. This fact makes it possible to formulate a criterion for greater and lesser inequality which transcends the several measures we have been considering. It is also closely bound up with the social welfare approach to inequality which will be discussed in the next section.

Suppose we rank order all the individuals in a population by income, from lowest to highest. For each rank, we may calculate the proportion of the population at that rank or below, and also the proportion of the total income that is earned by people at that rank or below. We may find, for example, that the poorest .25 of the population earns only .05 of the nation's income. If we plot the relationship between these two proportions for every rank, we get the Lorenz curve. Three such curves are shown in Figure 1. Line A, the diagonal straight line, is the Lorenz curve under the condition of perfect equality. It is often taken as a reference point for other observed curves. If any persons have unequal incomes, the Lorenz curve will fall below the line of perfect equality, indicating that the poorest y percent of the population earns less than y percent of the income for some value(s) of y.

The Gini index is equal to twice the area between the Lorenz curve and the line of perfect equality. Clearly line C, which falls entirely below line B, corresponds to a larger Gini index than line B. As many authors have shown (Morris, 1972), this conclusion also applies to any scale-invariant measure which satisfies the principle of transfers. Let us state this more formally. Let H be a scale-invariant measure of inequality that satisfies the principle of transfers, and let X and Y be two distributions. If the Lorenz curve for Y is never above and is somewhere below the Lorenz curve for X, then $H(X) < H(Y)$. The upshot is that whenever one Lorenz curve "dominates" another, in this sense, it makes little difference whether one uses the Gini index, the coefficient of variation or Theil's measure. All three will give the same rank ordering. It is not uncommon, however, for two Lorenz curves to intersect. In such cases, the Gini index may give one rank crdering while the coefficient of variation gives another.

5. Inequality and Social Welfare

In 1920, Dalton argued that the choice of an inequality measure typically involves an implicit normative judgment as to whether one distribution of income is to be preferred, in some sense, to another. He concluded that we would be in a better position to devise and choose measures of inequality if we could make these normative criteria more explicit. The problem may be formulated as follows: Suppose that we have a fixed total income that is to be distributed among n persons. We assume that for any possible distribution of incomes there is a number W indicating the desirability of that distribution. In functional notation

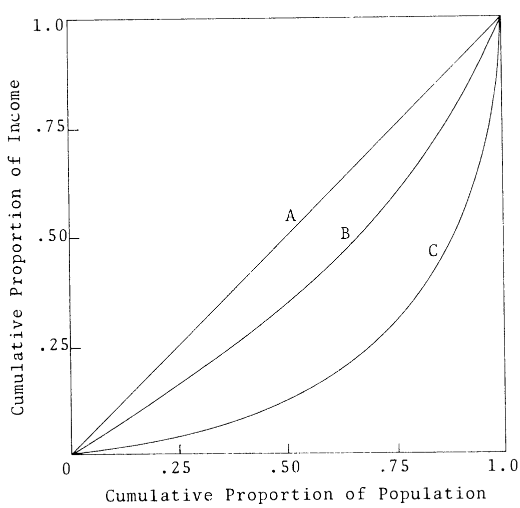

Figure 1. Lorenz Curves for Three Distributions of Income

$$W = W(x_1,x_2,\ldots,x_n) = W(X). \quad (15)$$

W is called the social welfare function. If we can specify the form of W, and if W reaches a maximum when all incomes are equal, then it is plausible to take a decreasing function of W as a measure of inequality.

This approach was largely ignored until Aigner and Heins (1967) proposed some alternative functions for W and derived corresponding measures of inequality. Atkinson (1970) greatly increased interest in the approach by demonstrating that plausible constraints on W implied an important relationship between W, the Lorenz curve, and the principle of transfers. He began by assuming that

$$W = \sum_{i=1}^{n} U(x_i). \quad (16)$$

$U(x_i)$ may be thought of as the utility of income x for individual i. Hence, the total social welfare is just the sum of the individual utilities. He further assumed that all individuals have the same utility function and that $U(x_i)$ is concave, i.e., there is diminishing marginal utility from increasing income. These last two assumptions imply that W will be maximized when all incomes are equal. Atkinson's major contribution was to show that, under these conditions, an ordering of the Lorenz curves implies an ordering of social welfare. Specifically, if we are given two distributions X and Y and the Lorenz curve for X is never below and somewhere above the Lorenz curve for Y, then $W(X) > W(Y)$. This result can also be expressed in terms of the principle of transfers. If a distribution Y can be obtained from a distribution X by a sequence of transfers from poorer to richer individuals, then $W(X) > W(Y)$.

Atkinson's conditions on W may seem rather restrictive, but they have since been greatly generalized by Dasgupta et al. (1973), Sen (1973), and Rothschild and Stiglitz (1973). The generalized versions of the theorem do not require that W be a sum of individual utility functions. What is necessary is that W be symmetric, continuous, monotonic, and "locally equality preferring." For an especially lucid account of these generalizations, see Rothschild and Stiglitz (1973).

The force of these theorems is to strengthen the theoretical foundation of the Lorenz curve and the measures of inequality that are associated with it. Given a fixed total income, if the Lorenz curve for one distribution X lies above the Lorenz curve for another distribution Y, we can be sure that X is preferable to Y under a broad class of functions defining preferability. Fixing the total income may seem like a restrictive condition, but it should not be a surprising one. Indeed, we would hardly expect that a perfectly equal distribution of $10 among ten persons would produce greater social welfare than a slightly unequal distribution of $1,000 among the same ten persons. Nevertheless, Rothschild and Stiglitz (1973) give a generalized definition of Lorenz dominance that allows for differences in total income in two distributions.

Let us return to Dalton's original aim of constructing a measure of inequality based on the social welfare function. The problem, of course, is that the choice of a social welfare function is normative, not empirical, and it is difficult to achieve any agreement on what that function should be. Perhaps the most widely acceptable class of functions is the additive, concave welfare function defined in (16). For this class, Atkinson (1970) showed that the following inequality measure was especially appropriate:

$$A = 1 - \frac{n}{1}\left[\sum_{i=1}^{n}\left(\frac{x_i}{\mu}\right)^{1-e}\right]^{\frac{1}{1-e}}, \quad (17)$$

where $e > 0$. The parameter e determines the "inequality aversiveness" of the measure. As e rises, A becomes more sensitive to transfers among lower incomes and less sensitive to transfers among top income recipients. In the limit, as $e \rightarrow 1$, A goes to $1 - M/\mu$ where M is the geometric mean and μ is the arithmetic mean. A is also scale invariant and satisfies the principle of transfers.

The advantage of A is that it provides a flexible, theoretically-based approach to the sensitivity question discussed in section 2. Thus, one may choose e to conform to one's judgment about what portions of the distribution are most relevant to the analysis. For some empirical applications of A, see Atkinson (1970), Bartels and Nijkamp (1976), and Williamson (1977). In all three cases, the authors replicate the analysis using several different values of e.

6. *Estimation and Testing*

To this point I have dealt only with measures of inequality applied to entire populations. In this section, I consider several issues related to estimation and testing with sample data, both grouped and ungrouped. As a prelude to the case of grouped data, I also present formulas for decomposing inequality into within- and between-groups components.

Ungrouped data. The usual approach to estimating measures of inequality is to apply the population formulas (1) through (7) to the sample in hand. While this produces consistent estimators in most cases, the resulting sampling distributions tend to be quite complicated, making it difficult to obtain standard errors, confidence intervals and test statistics. There is also the possibility that more efficient estimators might be available.

An alternative approach is that of maximum likelihood, which produces estimators that are approximately unbiased and efficient in large samples. To derive maximum likelihood estimators (MLEs), it is necessary to assume that the data come from a known family of distributions. While most statistical theory is built upon the normal distribution, this is an unsuitable family for the estimation of measures of inequality. The normal distribution is symmetrical and ranges from minus infinity to plus infinity. By contrast, measures of inequality are usually used only for variables which are inherently nonnegative, and the distributions are often positively skewed. The *lognormal* distribution has both these characteristics, and also has been found to provide a reasonably good fit to frequency distributions for a wide variety of positively-valued variables in several fields (Aitchison and Brown, 1957).

The lognormal distribution can be simply defined as follows. Suppose we have two random variables X and Y, and $Y = \log X$. If Y is normally distributed with mean R and variance L, then X is said to have a lognormal distribution with parameters R and L. Notice that L is just the variance of the logarithms which we considered in section 2. The fact that it is also the variance of a normally distributed random variable means that we can apply the usual normal theory for inferences about variances.

For the lognormal distribution, most of the inequality measures defined earlier can be expressed as increasing functions of L (Aitchison and Brown, 1957; Theil, 1967: 97). Thus

$$\begin{aligned} T &= L/2; \\ V &= (e^L - 1)^{1/2}; \\ G &= 2\,\Phi(\sqrt{L/2}) - 1; \end{aligned} \tag{18}$$

where $\Phi(.)$ is the cumulative distribution function for a standard normal variable. That is, $\Phi(a)$ is the probability that a normally distributed variable, with a mean of zero and a variance of one, is less than or equal to a.

To get MLEs of all these measures, we apply a basic principle of likelihood estimation (Cox and Hinkley, 1974: 302). Suppose we have a parameter A which can be expressed as a function of another parameter B, i.e., $A = f(B)$. Then if $\hat{B}$ is the MLE of B, it follows that $\hat{A} = f(\hat{B})$ is the MLE of A. This means that if we have the MLE of L, we can readily get MLEs of T, V, and G by substituting into (18). If the data consist of a simple random sample of n observations, it is well-known that the sample variance of $\log x_i$ (given in (7)) is the MLE of L.

As a numerical example, suppose that for a sample of 101 persons, the variance of the logarithm of income is .84. Then using (18) we get estimates of $\hat{T} = .42$, $\hat{V} = 1.15$, and $\hat{G} = .48$. In computing G, one can use the normal probability table found in almost any basic statistics text.

The asymptotic standard errors of $\hat{L}$, $\hat{T}$, $\hat{V}$ and $\hat{G}$ are easily derived (Kendall and Stuart, 1977: 258, 247), but it is not advisable to use them in constructing confidence intervals unless the sample is quite large. A better approach is to use the usual formula for constructing confidence intervals around the variance of a normal variaate. A $(1-\alpha)100$ percent confidence interval around L is given by

$$\frac{n\hat{L}}{\chi^2(n-1,\frac{\alpha}{2})} \leq L \leq \frac{n\hat{L}}{\chi^2(n-1,\frac{1-\alpha}{2})}, \quad (19)$$

where $\chi^2(n-1, b)$ is the value in a chi-square distribution with n–1 degrees of freedom which cuts off the upper b of sample values (Hays, 1963: 345).[8] In the example just given, a 95% confidence interval around L is $.65 \leq L \leq 1.13$. To obtain confidence interval for T, V, and G, simply substitute the upper and lower bounds for L into (18) which yields the corresponding upper and lower bounds for the derived inequality measures. In the example, this produces 95% confidence intervals of $.33 \leq T \leq .57$, $.96 \leq V \leq 1.45$, and $.43 \leq G \leq .64$.

It is also easy to test whether the level of inequality is the same in two distributions. Suppose we have two independent random samples of n_1 and n_2 observations from two different lognormal distributions. Since each of the four inequality measures is an increasing function of any of the others, $L_1 = L_2$ implies that $T_1 = T_2$, $V_1 = V_2$, and $G_1 = G_2$. Thus, we need only test the null hypothesis that $L_1 = L_2$. The test statistic is just $\hat{L}_1/\hat{L}_2$ which has an F distribution with n_1–1 and n_2–1 degrees of freedom under the null hypothesis (Hays, 1963: 351).

Although the assumption of a lognormal distribution provides an elegant and easy-to-use solution to the problem of inferences about measures of inequality, it is not without its limitations. First, it essentially does away with the differences in the various inequality measures discussed in section 2. Since any of the measures will give the same rank order for a set of lognormal distributions, one might just as well use L alone. Converting to T, G, or V by the formulas in (18) is only useful to achieve comparability with other studies. This is surely an advantage when the data truly come from a lognormal distribution. But departures from lognormality expose one to the peculiarities of L which were discussed in section 2. Second, neither the confidence interval in (19) nor the F test for comparing two populations is very robust to departures from lognormality (Hays, 1963: 352). Only very large samples can compensate for serious departures.[9] Finally, the methods fail when the sample includes values of zero since log 0 is undefined and, hence, L is undefined.

Decomposition. It is often desirable to decompose the inequality in a population into inequality between groups and inequality within groups. For example, it would be of considerable interest to decompose the inequality of household income in the U.S. into inequality between states and inequality within states. Theil (1967: 123–7) has given several decomposition formulas and I shall reproduce some of them here using slightly different notations.

Suppose the population can be divided into J mutually exclusive and exhaustive groups. For each group $j = 1,\ldots,J$ we know $\bar{X}_j$, the arithmetic mean income in group j, p_j the proportion of the population in group j, M_j the geometric mean income, and T_j, V_j, and L_j, the level of inequality in each group for three different measures. The decomposition of Theil's index for the total population is

$$T = \sum_{j=1}^{J} \left(\frac{p_j \bar{X}_j}{\bar{X}} \right) \log \left(\frac{\bar{X}_j}{\bar{X}} \right)$$

[8] For large n, it may be convenient to use the normal approximation to the chi-square distribution. See, e.g., Hays (1963:347).

[9] A plausible alternative to the lognormal distribution is the two-parameter gamma family, which also has an origin at zero and is positively skewed. Since one of the parameters α determines G, V, and T, one can take the same approach used here for the lognormal: first get the MLE of α and use that to get MLEs of G, V, and T. Estimation is considerably more difficult, however, because it requires the iterative solution of an equation involving a digamma function. Moreover, the equations relating α to G and T involve digamma and incomplete beta functions which are inconvenient to evaluate (Salem and Mount, 1974).

$$+ \sum_{j=1}^{J} \left(\frac{p_j \bar{X}_j}{\bar{X}} \right) T_j, \qquad (20)$$

where $\bar{X} = \sum_{j=1}^{J} p_j \bar{X}_j$ is the grand mean.

The first term on the right hand side is the between-groups component. It is equivalent to the value of T that would be obtained if everyone in each group received the mean income for that group. The second term on the right hand side is a weighted average of the within-group values of T. The weight for group j is the fraction of the total income that is earned by group j. For an extensive empirical application of (20), see Theil (1967: 98–114).

For L, the variance of the logarithms, we have

$$L = \sum_{j=1}^{J} p_j(\log M_j - R)^2 + \sum_{j=1}^{J} p_j L_j, \qquad (21)$$

where $R = \sum_{j=1}^{J} p_j \log M_j$. The first component is the value of L that would be obtained if everyone in group j received income M_j. The second component is again a weighted average of the within-group values of L, but here the weights are just the fractions of the population contained in each group.

Finally, we have

$$V^2 = \sum_{j=1}^{J} \frac{p_j (\bar{X}_j - \bar{X})^2}{\bar{X}^2} + \sum_{j=1}^{J} \left(\frac{p_j \bar{X}_j}{\bar{X}^2} \right) V_j^2. \qquad (22)$$

Notice that this is a decomposition of the *squared* coefficient of variation. The interpretation is similar to that for Thiel's measure, with one exception: the weights for V_j^2 do not have an obvious interpretation and they do not sum to one. Instead, they sum to one plus the between-groups component, a somewhat undesirable result. The Gini index cannot be conveniently decomposed.[10]

Grouped data. Frequently the researcher will wish to estimate the inequality in some population when individual-level data are unavailable. Either the data are grouped on some exogenous criterion, e.g., geographic units, or the grouping is by interval of the variable of interest. Let us first consider the exogenous criterion.

Suppose we wish to estimate the inequality of household income in a state, but the data are grouped by census tract. In each tract, we know the mean income $\bar{X}_j$ and p_j, the proportion of households in tract j. In order to get a lower bound on T and V, the obvious approach is to compute the between-groups components in (20) and (22). That this is a lower bound is evident from the fact that the true values of T and V must also include the within-groups components which could conceivably be quite large.

The geometric means will usually not be available to compute the between-groups components of L, the variance of the logarithms. However, we can still compute the variance of the logarithms that would be obtained if everyone in each group received the arithmetic mean income for that group. This is given by

$$\sum_{j=1}^{J} p_j (\log \bar{X}_j - R')^2, \qquad (23)$$

where $R' = \sum_{j=1}^{J} p_j \log \bar{X}_j$. Similarly, we can compute the Gini index that would be obtained if everyone in each group received the arithmetic mean income for that group. A convenient formula given by Blau (1977b) is

$$\frac{\sum_{j=1}^{J} \bar{X}_j p_j (q_j - r_j)}{\bar{X}}, \qquad (24)$$

where q_j is the proportion that are in

[10] Pyatt (1976) attempts a decomposition of the Gini index, but ends up with three components instead of two.

groups with means less than $\bar{X}_j$ and r_j is the proportion in groups with means greater than $\bar{X}_j$. Hence $p_j + q_j + r_j = 1$ for all j. As in the previous cases, this will give a lower bound on the true value of G.

The other sort of grouped data consists of a grouped frequency distribution. We are given a set of cutpoints a_j for $j = 1, \ldots, J + 1$, and p_j the proportion of observations falling in the interval $[a_j, a_{j+1})$ for $j = 1, \ldots, J$. In most cases, a_1 will be zero, and in some cases a_{J+1} will be unspecified. In some cases, we also know $\bar{X}_j$, the mean in each interval. When $\bar{X}_j$ is not available, most methods require an estimate, such as the arithmetic midpoint, $\frac{1}{2}(a_{j+1} + a_j)$. The most common method for estimating measures of inequality from grouped frequency distributions is to treat the data as if the grouping were based on some exogenous criterion. That is, the inequality measures are computed as though each person's score were the interval mean or midpoint. This is a lower bound, as we have just seen. When the data are grouped by intervals, however, it is also possible to compute sharp upper bounds on V, T and G by methods recently introduced by Gastwirth (1972; 1975) and Mehran (1975). An alternative approach begins with the assumption that the data are drawn from an underlying probability distribution. Aitchison and Brown (1957: 51) give the MLE of L under the assumption of a lognormal distribution. Aigner and Goldberger (1970) assume a Pareto distribution and give several methods for estimating the parameter α, which also determines many of the inequality measures. Kakwani and Podder (1973; 1976) specify the distribution by specifying the Lorenz curve; they then obtain efficient estimators for all inequality measures that depend on the Lorenz curve.

Conclusion

One approach to defining inequality is to choose one of the common measures of inequality. The range of choice is narrowed considerably by requiring that (a) inequality be invariant to proportionate increases or decreases in everyone's score, and (b) any transfer from an individual with a lower score to another individual with a higher score represents an increase in inequality. This leaves three measures: the Gini index, the coefficient of variation and Theil's measure.

Although the Gini index is the most popular measure of inequality, it does not appear to possess any special advantages over the other two measures. Moreover, it is the most difficult to compute and cannot be decomposed into inequality within and between groups. Because its sensitivity to transfers decreases as scores increase, Theil's measure is especially desirable for measuring inequality of income, or other social rewards having diminishing marginal utility. For variables like age, where utility is neither strictly increasing nor especially relevant, the flat sensitivity of the coefficient of variation makes it the appropriate choice. The fact that it is also easily obtained from standard computer output makes it particularly useful for exploratory work.

None of these inequality measures is appropriate for interval-level variables which lack a theoretically fixed zero point. If they are applied to such data, the numerical values are essentially arbitrary. Nevertheless, we saw in section 3 that there are special situations in which valid comparisons of inequality can be made for two interval-level distributions. These situations arise when it can be assumed that there is an underlying nonnegative ratio scale. If such an assumption cannot be made, the only recourse is to use a scale-dependent measure of dispersion such as the standard deviation.

It is a simple matter to get efficient estimates of the Gini index, the coefficient of variation, and Theil's measure under the assumption that the data come from a lognormal distribution. Simply take the logarithm of each score and calculate the variance of these transformed scores. Using formulas in section 6, we can then easily obtain maximum likelihood estimates of the three measures. Methods for confidence intervals, hypothesis tests, and grouped data also were presented.

Many economists have taken the position that instead of choosing one of the traditional inequality measures, one should first specify a social welfare function and then derive an appropriate measure of inequality. Unfortunately, there is little consensus on the general form of the social welfare function, let alone the parameter values for any particular functional form. Atkinson (1970) also has warned that a social welfare function for the distribution of income among persons may be entirely inappropriate for the social welfare of industrial concentration, another area in which measures of inequality are frequently employed (Hall and Tideman, 1967). The same may be said for such social rewards as power and prestige.

Sen (1973) has suggested that current notions of inequality may be too imprecise for any conventional measure of inequality. In his view, the best we may be able to achieve without exceeding our understanding is a partial ordering: for some pairs of distributions we can say that one is more unequal than the other, while for other pairs we suspend judgment. Such a partial ordering is produced by the criterion of Lorenz dominance discussed in section 4. If one Lorenz curve lies nowhere below and somewhere above another, the first distribution is the more equal. For distributions whose Lorenz curves cross, no ranking is possible. Although Sen's position is extreme, it is worthwhile keeping in mind that the Lorenz dominance criterion is virtually unquestioned. Thus, if a particular hypothesis about inequality can be demonstrated by Lorenz dominance, confidence in the result is greatly strengthened.

APPENDIX

Proof of Test for Differences in Inequality When Variables Are Measured on an Interval Scale

Let X be an n x 1 vector of observations on an interval scale and let Y be an m x 1 vector of observations on the same interval scale. Let X* and Y* be nonnegative vectors representing an underlying ratio scale. They are related to X and Y by

$$X = X^* + C_n; \quad Y = Y^* + C_m; \tag{A.1}$$

where C_n is a vector obtained by multiplying an n x 1 vector of ones by the constant c, and similarly for C_m. The only restriction on c is that

$$c \leq \min(x_1, \ldots, x_n, y_1, \ldots, y_m), \tag{A.2}$$

since otherwise X* or Y* would have negative elements. I(X) is defined in (12) and c* is defined in (14). The assertion to be proved is that

(a) $I(X) > I(Y)$ and $c^* > \min(x_1, \ldots, x_n, y_1, \ldots, y_m) \geq 0$;

or

(b) $I(X) > I(Y)$ and $\mu_x I(X) = \mu_y I(Y)$

implies

(c) $I(X^*) > I(Y^*)$.

Proof by contradiction. Assume that (a) or (b) is true, and assume for the moment that (c) is false, i.e., $I(X^*) \leq I(Y^*)$. From (A.1) we have

$$I(X - C_n) \leq I(Y - C_m)$$

which, by (13), is

$$\left(\frac{\mu_x}{\mu_x - c}\right) I(X) \leq \left(\frac{\mu_y}{\mu_y - c}\right) I(Y).$$

This reduces to

$$c[\mu_x I(X) - \mu_y I(Y)] \geq \mu_x \mu_y [I(X) - I(Y)]. \tag{A.3}$$

If (b) is true, (A.3) simplifies to $I(Y) > I(X)$ which is a contradiction. Assume (a) is true. The assumption that $c^* > 0$ implies that $\mu_x I(X) - \mu_y I(Y) > 0$. Formula (A.3) then becomes

$$c \geq \frac{\mu_x \mu_y [I(X) - I(Y)]}{\mu_x I(X) - \mu_y I(Y)} = c^* > \min(x_1, \ldots, x_n, \ldots, y_1 \ldots, y_m),$$

which contradicts (A.2). Therefore, $I(X^*) > I(Y^*)$.

REFERENCES

Aigner, Dennis J. and Arthur S. Goldberger
1970 "Estimation of Pareto's law from grouped observations." Journal of the American Statistical Association 65: 712–23.

Aigner, Dennis J. and A. J. Heins
1967 "A social welfare view of the measurement of income equality." Review of Income and Wealth 13:12–5.

Aitchison, J. and J. A. C. Brown
1957 The Lognormal Distribution. Cambridge, Eng.: Cambridge University Press.

Alker, H. and B. M. Russett
1964 "On measuring inequality." Behavioral Science 9:207–18.

Atkinson, Anthony B.
1970 "On the measurement of inequality." Journal of Economic Theory 2:244–63.

Bartels, C. P. A. and P. Nijkamp
1976 "An empirical welfare approach to regional income distributions." Socio-Economic Planning Sciences 10:117–28.

Blau, Peter M.
1977a Inequality and Heterogeneity. New York: Free Press.
1977b "A macrosociological theory of social structure." American Journal of Sociology 83:26–54.

Champernowne, David
1974 "A comparison of measures of inequality of income distribution." The Economic Journal 84:787–816.

Cox, D. R. and D. V. Hinkley
1974 Theoretical Statistics. London: Chapman and Hall.

Creedy, John
1977 "The principle of transfers and the variance of the logarithms." Oxford Bulletin of Economics and Statistics 39:152–8.

Dalton, Hugh
1920 "The measurement of the inequality of incomes." Economic Journal 30:348–61.

Dasgupta, Partha, Amartya Sen and David Starrett
1973 "Notes on the measurement of inequality." Journal of Economic Theory 6: 180–7.

Easterlin, R. A.
1974 "Does economic growth improve the human lot? some empirical evidence." Pp. 89–125 in P. A. David and M. W. Reder (eds.), Nations and Households in Economic Growth. New York: Academic Press.

Farley, Reynolds
1977 "Trends in racial inequalities." American Sociological Review 42:189–208.

Gastwirth, Joseph L.
1972 "The estimation of the Lorenz curve and the Gini index." The Review of Economics and Statistics 54:306–16.
1975 "The estimation of a family of measures of income inequality." Journal of Econometrics 3:61–70.

Hall, M. and N. Tideman
1967 "Measures of concentration." Journal of the American Statistical Association 62:162–8.

Hauser, R. M., P. J. Dickinson, H. P. Travis, and J. M. Koffel
1975 "Structural changes in occupational mobility among men in the United States." American Sociological Review 40:585–98.

Hays, William
1963 Statistics. New York: Holt, Rinehart and Winston.

Hewitt, Christopher
1977 "The effect of political democracy and social democracy on inequality in industrial societies." American Sociological Review 42:450–64.

Jackman, Robert W.
1974 "Political democracy and social equality: a comparative analysis." American Sociological Review 39:29–45.

Jencks, Christopher, M. Smith, H. Acland, M. J. Bane, D. Cohen, H. Gintis, B. Heyns, and S. Michelson
1972 Inequality. New York: Harper and Row.

Kakwani, N. C. and N. Podder
1973 "On the estimation of Lorenz curves from grouped observations." International Economic Review 14:278–92.
1976 "Efficient estimation of the Lorenz curve and associated inequality measures from grouped observations." Econometrics 44:137–8.

Kelley, Jonathan and Herbert S. Klein
1977 "Revolution and the rebirth of inequality: a theory of stratification in postrevolutionary society." American Journal of Sociology 83:78–99.

Kendall, Maurice and Alan Stuart
1977 The Advanced Theory of Statistics. Vol. 1. 4th ed. New York: Macmillan.

Kolm, Serge-Christophe
1976 "Unequal inequalities, I." Journal of Economic Theory 12:416–42.

Lenski, Gerhard
1966 Power and Privilege. New York: McGraw-Hill.

Martin, J. David and Louis N. Gray
1971 "Measurement of relative variation." American Sociological Review 36:496–502.

Mehran, Farhad
1975 "Bounds on the Gini index based on observed points of the Lorenz curve." Journal of the American Statistical Association 70:64–6.

Morris, Carl
1972 Measures of Relative Income Inequality. Santa Monica: Rand.

Mueller, John H., Karl F. Schuessler and Herbert L. Costner
1977 Statistical Reasoning in Sociology. 3rd ed. Boston: Houghton-Mifflin.

Nerlove, Marc and S. James Press
1973 Univariate and Multivariate Log-Linear and Logistic Models. Santa Monica: Rand.

Pyatt, Graham
1976 "On the interpretation and disaggregation of Gini coefficients." Economic Journal 86:243–55.

Ray, James Lee and J. David Singer
1973 "Measuring the concentration of power in the international system." Sociological Methods and Research 1:403–37.

Rothschild, Michael and Joseph E. Stiglitz
1973 "Some further results on the measurement of inequality." Journal of Economic Theory 6:188–204.

Rubinson, Richard and Dan Quinlan
1977 "Democracy and social inequality: a reanalysis." American Sociological Review 42:611–23.

Salem, A. B. Z. and T. D. Mount
1974 "A convenient descriptive model of in-

come distribution: the gamma density." Econometrica 42:1115–27.

Schutz, R. R.
1951 "On the measurement of income inequality." American Economic Review 41:107–22.

Sen, Amartya
1973 On Economic Inequality. New York: Norton.

Theil, Henri
1967 Economics and Information Theory. Chicago: Rand McNally.

Williamson, Jeffrey G.
1977 " 'Strategic' wage goods, prices, and inequality." American Economic Review 66:29–41.

Yntema, D. B.
1933 "Measures of inequality in the personal distribution of income or wealth." Journal of the American Statistical Association 28:423–33.

THE EFFECT OF DIRECT GOVERNMENT INVOLVEMENT IN THE ECONOMY ON THE DEGREE OF INCOME INEQUALITY: A CROSS-NATIONAL STUDY*

STEVEN STACK
Alma College

American Sociological Review 1978, Vol. 43 (December):880–888

Sociological theories of income inequality have neglected a Keynesian approach to the problem. This paper tests the Keynesian notion that the degree of direct government involvement in the economy should reduce income inequality through such means as full employment and the facilitation of economic growth. A regression analysis of data from 32 nations indicates that the degree of direct government involvement in the economy is the single most important factor associated with low income inequality. This relationship is independent of both the level of economic development and the rate of economic growth.

The sociological explanation of income inequality has focused on a number of recurrent themes such as economic development theory (Kuznets, 1963; Kerr et al., 1964; Cutright, 1967; Peters, 1973; Hewitt, 1977; Rubinson and Quinlan, 1977), and world-economy theory (Cutright, 1967; Rubinson, 1976). However, there is also a fourth possible theory, one built upon the Keynesian theory of political economy.[1] The degree of income inequality also depends upon the extent to which the state is directly involved in the national economy. Such mechanisms as employment producing expenditures, state provided employment, and the enactment of monetary and fiscal policies that will augment the rate of growth have considerable bearing on the degree of inequality in society. With data from 32 nations, this paper performs a test of two related models of income inequality, the

* Address all communications to: Steven Stack; Department of Sociology; Alma College; Alma, MI 48801.

The author would like to thank Kenneth Neubeck, William V. D'Antonio, Bill Harrell, Gerald Sazama, and the anonymous reviewers for their helpful comments, suggestions, criticisms, and encouragement on earlier versions of this work.

[1] The Keynesian model of income inequality might be viewed as a refinement and extension of the political model. For example, the political model focuses on elements of political organization and development which are presumed to have an effect on income inequality, while the Keynesian model focuses on the actual policies adopted by governments in their efforts to reduce inequality. However, there are a number of differences between the two perspectives which may serve to separate the two models. Political models of income redistribution tend to view the problem from a conflict perspective wherein the common people obtain universal suffrage and, through such mechanisms as the formation of social-democratic parties, use the state apparatus to obtain a larger share of the income at the expense of the elite. However, in a Keynesian perspective we may understand the process of redistribution through the use of a consensus frame of reference. Here the balancing of savings, consumption and investment benefits the propertied class as well as the common people. Relative to a state of recession, the achievement of such an economic equilibrium increases profits as well as wages.