VALUE JUDGEMENTS IMPLIED BY THE USE OF VARIOUS MEASURES OF INCOME INEQUALITY*

YAAKOV KONDOR
Economic Planning Authority, Jerusalem

If a welfare economist wants to express income inequality in a sensible way by a single parameter, he has to make rather strong assumptions regarding the social preferences of his fellow citizens. Formulas are presented with whose aid one is able to test whether or not these assumptions hold. The standard measures used nowadays contradict prevailing preferences. If no common single measure can be found which fits the social preferences of almost all individuals concerned tolerably well, additional parameters measuring poverty and riches separately are necessary.

1. INTRODUCTION

Income inequality is essentially an n-dimensional phenomenon. In order to express it meaningfully by a scalar, we must have in mind an n-dimensional objective function which can be abbreviated as a function of a few parameters of the income distribution. One of these parameters may be an inequality measure. The nature of this measure depends obviously on the nature of our original, unabbreviated objective function. Conversely, an inequality measure may be used meaningfully only if it appears (at least approximately) in the abbreviated objective function.

In general, our objective function may express either a positive theory or normative preferences. In the present paper, however, we are interested only in normative, i.e. "social welfare," functions. Thus the theme of this paper is the correspondence which ought to exist between value judgements and measures of inequality. We shall see that such correspondence does not exist to date between prevailing value judgements and traditional inequality measures. Moreover, expressing inequality by any single number may be too drastic an abbreviation in the sense that it is liable to prevent meaningful discussion of and decision on the income distribution by individuals whose values differ.

2. AN ABBREVIATED SOCIAL WELFARE FUNCTION

If someone has consistent preferences regarding the size distribution of incomes, then, for some set of income vectors $(x_1, \ldots, x_n)$, there exists a function,

$$w = w(x_1, \ldots, x_n), \tag{1}$$

unique at least up to monotonic transformation and commutative (symmetric) in

*This paper briefly restates and develops further some ideas contained in an unpublished D.Sc. thesis [12] submitted to the Technion, Israel Institute of Technology, Haifa. The author is indebted especially to M. Schechter for valuable comments and criticisms. Opinions expressed are the author's own. So are any shortcomings.

all $x_i (i = 1, \ldots, n)$. We shall call $w(\ )$ "the social welfare function."[1] The individual whose preferences are described by $w(\ )$, will be called "the client" because we offer him our services in our capacity as welfare economists. There are as many social welfare functions (and clients), as there are people interested in the size distribution of incomes. Equation (1) is, accordingly, an individualistic or Bergson-type social welfare function, not an universalistic or Hegel–Arrow-type one.[2] The individualistic approach has the advantage that there are no *a priori* grounds to call in question the existence of $w(\ )$. On the contrary, it is an empirical fact that people do have preferences about the size distribution of incomes.

The mere existence of (1), however, is not enough. It must also be valid. That is, the client must not think in the wrong terms so that his observed $w(\ )$ fails to describe his ultimate social preferences. We therefore assume that:

(a) Incomes (and income receiving units) are defined such that they may be used as tolerable indicators of the net economic satisfaction of individuals.

(b) Relevant changes in the income distribution will not produce significant changes in relative prices.

We do *not* assume that the client is interested only in the size distribution of incomes and is indifferent to other social variables. His social preferences in general may be described by a function of the type

$$w_g = w_g[w(\ ), y_1, y_2, \ldots], \tag{2}$$

where the ys may stand for any variables. So long, however, as the client is interested in the size distribution too and has consistent preference about it, $w(\ )$ will remain an argument in his general welfare function. The importance of political preferences lies, of course, in that that they guide any one of us in political decision making (e.g. the citizen deciding how to vote).

It is, however, not practical to decide, or even to think, about the income distribution in an n-dimensional space. In order to make the problem tractable, we have to simplify it. One way of doing so is to arrange the incomes in a limited number of size groups or fractile groups and to regard each group as homogeneous. Another way is to work with a few parameters which "represent" the whole distribution. The most frequently used parameters are total income, S (or the mean income, M[3]), and an inequality measure, $\theta(x_1, \ldots, x_n)$. (On the constraints side we have here the well-known choice between "efficiency" and "equity.") Representing the distribution by these two parameters is legitimate if, and only if, it is possible to abbreviate (1) by the form $W(S, \theta)$, that is, if for every relevant

[1]A utilitarian welfare function, restricted to the form $w = \sum_i u(x_i)$, employed recently by Aigner and Heins [1], Atkinson [2], Bentzel [3] and others, is unacceptable for two reasons. First, the individual's utility is certainly not a function solely of *his* income; second, if a client is interested in the *sum* of utilities, this means that he is indifferent to their distribution which is contrary to experience. Cf. Sen [15] 15–18.

[2]The problems dealt with in this paper are relevant to an Arrowian welfare function too but I do not think that such functions, even if they exist, are necessary for good solutions of the social decision problem. (By "good solution" is meant a feasible program decided on by some democratic process.) The sharpest statement of the individualistic approach was made by Little [13]. See also Bergson [5].

[3]Since in this paper we assume that n is given, it makes no difference whether we use S or M.

distribution we have, at least approximately,

$$w(x_1, \ldots, x_n) = W(S, \theta). \tag{3}$$

When such an abbreviation exists for any inequality measure, we shall say that the measure and the welfare function fit each other. We have a special, but plausible and convenient, case if, for any given S, W is a monotonic function of θ, at least for the relevant region of distributions.[4] A client will be called an "equality lover" if his W is a decreasing function of some θ; in the opposite case he will be an "inequality lover."[5]

The functional equation (3) raises various kinds of problems, depending on the assumptions and restrictions connected with it. Only two of these problems will be dealt with at length here:

1. How to test whether a given θ fits the unknown $w(\ )$ of a client.
2. How to reveal general properties of welfare functions which fit a given θ.

An important related problem, which will be touched on briefly in the last section, is the following: Can we get a tolerable approximation of all (or almost all) the existing preferences in a society by using the same θ in a variable function $W(S, \theta)$? Or must we use different inequality measures in order to express different preferences? Clearly, if we are compelled to do so, our abbreviation (3) will be worthless, as it will make impossible meaningful social discussion and decision on the size distribution of incomes.[6]

3. Quantitative Tests of Fit

We want to test whether, using S and a given θ, it is possible to extract for a certain client all (or almost all) the relevant information contained in the n-dimensional income space. In other words, we are going to test if (3) is true, at least approximately, for the given θ and the given (but unknown) $w(\ )$.

Suppose (3) is true. For simplicity assume also that $\partial W/\partial\theta$, $\partial W/\partial S$ exist in the whole relevant region of the (S, θ) plane. In this case, the marginal rate of substitution between any two incomes, x_i, x_j (if the partial derivatives $\partial\theta/\partial X_i$, $\partial\theta/\partial X_j$ exist too) is

$$\frac{dx_i}{dx_j} = -\frac{\partial W/\partial S + (\partial W/\partial\theta)(\partial\theta/\partial x_j)}{\partial W/\partial S + (\partial W/\partial\theta)(\partial\theta/\partial x_i)} = -\frac{d\theta/dS - \partial\theta/\partial x_j}{d\theta/dS - \partial\theta/\partial x_i} \tag{4}$$

[4]As an example of non-monotonic fit, think of a client whose $W(\ )$, for a given S, has a single maximum at some strictly positive value of θ. Such a peculiar maximum point must be, however, sharply distinguished from the *optimum* "degree of inequality" which will be strictly positive for almost all clients: in the first case we are dealing with the nature of the objective function; in the second, with the nature of the optimum solution of the distribution problem.

[5]These terms are of course derived from Tobin's "risk averter" and "risk lover" [17]. However, a typical client will not evaluate the income distribution, even formally, in the same fashion as he does his portfolio. See on this Appendix A in [12] where other exponents of the functional approach to the measurement of inequality (or dispersion in general) are reviewed too, along with a critique of the traditional formalistic approach.

[6]It is evident that we may abbreviate all welfare functions on the same (S, θ) field if the relevant income distributions are all of the same two-parametric (e.g. lognormal) type. But it would be absurd to restrain the freedom of decision on the shape of the distribution in such a way. Other, more reasonable, restrictions will be mentioned in the last section.

Here $d\theta/dS$ is the marginal rate of substitution between the two parameters. By (4), it may be derived explicitly from the marginal rate of substitution between the two incomes that:

$$\frac{d\theta}{dS}=\frac{\partial\theta/\partial x_j+(\partial\theta/\partial x_i)(dx_i/dx_j)}{1+dx_i/dx_j}. \tag{5}$$

Now exhibit to the client the existing distribution of incomes. Let him consider it and then ask him (say): "By how much must an income of $10,000 be increased in order to compensate (you!) for a $100 cut in an income of $5,000?" The answer will give an approximation of the marginal rate of substitution between the two incomes. By (5), we can deduce from it $d\theta/dS$. Again, from $d\theta/dS$ we can predict, by (4), the marginal rate of substitution between any other two incomes. If repeated questions of this sort elicit answers not differing significantly from the predicted results, we can accept the hypothesis that θ fits the client's $w(\ )$. Otherwise it must be rejected.

4. Qualitative Tests of Fit

The tests dealt with in the rest of this paper are built on qualitative or semiqualitative questions and answers and are thus weaker than the test of Section 3. By the same token, however, they give more general results. Indeed, they enable us to reject some well-known measures of inequality from no more than a superficial knowledge of clients' preferences. Here, the testing questions refer to two kinds of change in the distribution: (a) transfers between two incomes; (b) increments to a given income, other incomes remaining equal. The client is asked to give one of three answers: "The change is desirable," "undesirable" or "uninteresting."

Transfers leave S unchanged. Therefore, if a transfer alters $w(\ )$, it must alter θ too, otherwise (3) would not hold. This condition alone enables us to reject certain measures of inequality, namely those which are constant under transfers in a wide region in which the client is not indifferent to the distribution of incomes. For instance, the measures $(\max x_i - \min x_i)$ and $(\max x_i)/(\min x_i)$ $(i=1,\ldots,n)$, can fit only those (presumably rare) clients who are indifferent to any transfer which does not alter the lowest or the highest income. More important examples will be given later.

Suppose now that the preferences of a client are such that any transfer from a higher to a lower income is desirable for him, provided that the transfer does not widen the initial difference between them [the last condition follows from the symmetry and transitivity properties of $w(\ )$]. It is obvious that, if the client is to be considered as an equality lover [i.e., when for a given S, $W(\ )$ is a decreasing function of θ], then θ ought to decrease as a result of every such transfer.[7] The first to use this property (that, for all $x_i > x_j$, any transfer from x_i to x_j reduces θ), as a criterion of a "good" inequality measure, was Dalton.[8] We shall therefore refer to it as the Dalton condition.

[7] It may be shown that the same conclusion holds for all clients, equality-loving or not, provided that θ is a continuous function of the incomes.

[8] [7] 5. In the source this requirement was called "the principle of transfers." It was derived from the assumption of a decreasing marginal utility of the incomes when θ functioned as a measure of the inefficiency of the distribution in a utilitarian sense.

If one distribution is derived from another by a series of Daltonian transfers, it is fairly clear (and has been proven in the literature)[9] that the Lorenz curve representing the new distribution will lie entirely within the old one. It follows that distributions which can be represented by non-intersecting Lorenz curves are ranked identically by these curves and by all (but only those) inequality measures which observe the Dalton condition and are homogeneous of zero degree in all incomes.

If the partial derivative, $\partial\theta/\partial x_i$ exists for all x_i, then the Dalton condition may be written as

$$d\theta = \frac{\partial\theta}{\partial x_i}(-dx_j) + \frac{\partial\theta}{\partial x_j}\, dx_j < 0. \tag{6}$$

That is, for all $x_i > x_j$,

$$\frac{\partial\theta}{\partial x_i} > \frac{\partial\theta}{\partial x_j}. \tag{7}$$

If θ is a smooth function of the incomes, this means simply that for all x_i,

$$\frac{\partial(\partial\theta/\partial x_i)}{\partial x_i} > 0, \tag{8}$$

where all symmetrical functions of the income vector appearing in the first derivative are taken as constants.[10]

The Dalton condition has considerable cutting power: in the next section we shall see that some of the traditional inequality measures do not fulfil it. The reader should remember, however, that it is a legitimate condition only for a "Daltonian" client. It is not at all self-evident that all equality lovers (let alone other clients) have such preferences. For any transfer which diminishes the difference between two incomes in general increases the difference between them and some other incomes,[11] the evaluation of these counteracting effects depending on the form of $w(\ )$.[12] More important, in its strong form the Dalton condition is not useful for clients who are virtually indifferent to the distribution *within* small regions of the income space. For example, a measure of income inequality which is computed from grouped data may closely approximate the preferences of a client who is indifferent to the distribution within the income groups; but evidently measures so computed violate the Dalton condition. In such cases, if the client does not consider it positively desirable to widen the difference between any given pair of incomes, a weak Dalton condition may be required:

$$\frac{\partial(\partial\theta/\partial x_i)}{\partial x_i} \geq 0, \qquad i = 1, \ldots, n \tag{9}$$

[9] Cf. Atkinson [2] 245–249; Kolm [11] 188–193.

[10] For in (8) we are dealing with the result of an (infinitesimal) uphill motion *on* a given distribution, not with a change *in* it.

[11] This point has already been made by Blum and Kalven [6] 97–98.

[12] A symmetric $w(\ )$ is Daltonian if, and only if, it is strictly quasi-concave. See Sen [15] 52–53. On non-Daltonian welfare functions and on the possible rationale for them see Gorman [9]. Cf. also the case of λ^2 in the next section.

where the second derivation is performed as in (8). This is, it must be conceded, a very weak test. For example, the measure (max x_i − min x_i) will pass it (though some other measures, including, as we shall see, the logarithmic variance, will not). It is possible, however, to combine (9) with a modified (7) where only distant incomes are compared.

We proceed now to the other type of elementary change in the distribution. The test here is based on the question: "Is it desirable to increase some specified income, x_i, all other incomes remaining the same?" If the answer is "yes" for all i and for all conceivable distributions, we shall say that the welfare function (and the client) is Paretian. Of course, even a non-Paretian client will, in general, desire increases in some incomes and, for certain regions of the income space, in all of them. Non-Paretianity therefore needs additional specification regarding its boundary both within a given distribution and in the relevant income space.

There is no *a priori* reason to reject a non-Paretian welfare function, nor does our everyday experience of human judgements make it implausible. To begin with, one may accept the Pareto principle in its original meaning, i.e., with respect to utilities, but not, on account of external effects, with respect to incomes. Moreover, a client may be a non-Paretian even with respect to utilities if, for example, an increase in the pleasures of the very rich goes against his sense of justice.[13] We know from experience that there are such clients, perhaps a lot of them.[14] On the other hand, there are certainly many Paretians too. Both kinds of clients have certain restrictions on the inequality measures eligible for them under (3). The mathematical expression of Paretianity is, of course,

$$\text{(10)} \qquad \frac{\partial w}{\partial x_i} > 0, \qquad (i = 1, \ldots, n).$$

For simplicity, let us suppose from now on that the client is an equality lover, $\partial W/\partial\theta < 0$, and is also interested (for a given θ) in raising the total income, $\partial W/\partial S > 0$.[15] Replacing w by $W(S, \theta)$ in (10), developing the left-hand side by the function of a function rule, as in (4), dividing by $-\partial W/\partial\theta$ and regrouping gives, on these assumptions, the alternative expression of Paretianity:

$$\text{(11)} \qquad \frac{d\theta}{dS} > \frac{\partial\theta}{\partial x_i} \qquad (i = 1, \ldots, n).$$

This states no more than that the increase in the inequality measure necessary to make the client indifferent to a unit increase in S is greater than the actual increment of θ following from a (ceteris paribus) unit increase in any income. This is as it should be, since the client is *not* indifferent to the actual change.

In order to demonstrate a restriction on θ implied by Paretianity, consider three incomes, $x_i > x_j > x_k$, such that the client has a clearcut preference for a

[13]Some possible cases of non-Paretianity are dealt with by Vickrey [18] 533.

[14]Discarding the Pareto principle as an axiom ought not alter our approach to allocation policy. It is a pity that economists are not always aware of the fact that the Pareto principle is superfluous for the derivation of the classical optimal conditions of allocation, even though Bergson derived them without it over thirty years ago [4] 318.

[15]Parallel, but perhaps less interesting, conclusions may be drawn in the opposite cases. Complications will arise if W is not a monotonic function of S and θ. However, such cases seem to be rare, at least for the relevant sets of distributions.

transfer from x_j to x_k. Suppose that θ fulfils the Dalton condition. Then, inserting (11) and (7) in (4), we have

$$(12) \qquad 1<\left|\frac{dx_j}{dx_k}\right|<\frac{\partial\theta/\partial x_i-\partial\theta/\partial x_k}{\partial\theta/\partial x_i-\partial\theta/\partial x_j}.$$

Now, if θ is such that $\partial\theta/\partial x_i$ is not bounded from above then, as x_i grows indefinitely, the marginal rate of substitution in (12) will approach unity. But this means that the client is indifferent to a transfer between x_j and x_k, which contradicts our assumption. Such a θ must therefore be disqualified. A traditional group of inequality measures, which on these grounds does not fit Paretian clients, is the variance and related indexes, as we shall see in the next section.

Non-Paretians too have their restrictions. For such a client the opposite of (11) holds, that is

$$(13) \qquad 0<\frac{d\theta}{dS}<\frac{\partial\theta}{\partial x_i},$$

for all x_i in the non-Paretian region. (The left-hand inequality is always true for $\partial W/\partial\theta<0$, $\partial W/\partial S>0$.) The condition, that $\partial\theta/\partial x_i$ ought to be strictly positive in all the relevant non-Paretian region attached to $w(\ )$, may in certain cases disqualify *inter alia* the Gini coefficient of concentration, as we shall see later. Expression (13) may also be interpreted in the reverse direction: if θ is an inequality measure fitting the welfare function, then for those sections of the distribution where $\partial\theta/\partial x_i$ is positive, a strong enough dislike of inequality ($d\theta/dS<\partial\theta/\partial x_i$) implies non-Paretianity of $w(\ )$.

5. Value Judgements Implied by Four Conventional Inequality Measures

In this section we inquire into the general properties of functions of the form $W(S, \theta)$ which fit one of the following four conventional measures of inequality: (a) the variance, σ^2; (b) the logarithmic variance, λ^2; (c) the Gini coefficient of concentration, R; (d) the relative mean deviation, T. The definitions of these parameters, their range and their first and second derivatives [the second in the sense of (8)] with respect to any income, x_i, are exhibited in Table 1.[16] For simplicity, we shall again assume $\partial W/\partial\theta<0$, $\partial W/\partial S>0$.

a. *The Variance*

It must be emphasized at the outset that our conclusions regarding the variance will stand for the whole family of inequality measures connected with it: for the standard deviation, which is its monotonically increasing function; for the

[16]Since we are dealing with finite populations of individual incomes, it seemed natural to adapt the parameters to a continuous n-dimensional income space instead of defining them as functionals of a two-dimensional continuous distribution function.

Pearson coefficient of variation, since any function of the form

$$W\left(S, \frac{\sigma}{M}\right), \quad \frac{\partial W}{\partial S} > 0, \quad \frac{\partial W}{\partial(\sigma/M)} < 0 \tag{14}$$

may be also written as

$$W^*(S, \sigma^2), \quad \frac{\partial W^*}{\partial S} > 0, \quad \frac{\partial W^*}{\partial \sigma^2} < 0 \tag{15}$$

(the converse, as we shall see, is not true); and, on the same grounds, for the standardized measure (with range 0 through 1), $\sigma^2/(\sigma^2+M^2)$.

From Table 1 we note that the variance fulfils the Dalton condition (8). (We shall see below that it is nevertheless insensitive to many significant transfers.) From Table 1 and from (13) we see also that a welfare function of the form $W(S, \sigma^2)$ may be non-Paretian for any income above the mean. By (12), it *must* be non-Paretian for *some* incomes (unless the client is indifferent to any transfer), since $\partial\sigma^2/\partial x_i$ is not bounded from above. Moreover, in the case of σ^2 non-Paretianity is implied not only for some (perhaps non-existent) incomes. It is binding upon at least the upper tail of empirical distributions. Otherwise, virtual indifference to the distribution of at least all incomes below the mean is implied. Consider the marginal rate of substitution between an income which equals the mean, x_M, and an income which equals zero, x_0, in the case of a client who is Paretian for an income that equals K times the mean, x_{KM}. Substitution in (12) gives:

$$1 < \left|\frac{dx_M}{dx_0}\right| < \frac{\partial\sigma^2/\partial x_{KM} - \partial\sigma^2/\partial x_0}{\partial\sigma^2/\partial x_{KM} - \partial\sigma^2/\partial x_M} = \frac{(K-1)M+M}{(K-1)M} = \frac{K}{K-1}. \tag{16}$$

If, for example, the client is Paretian up to incomes of 10 times the mean or less, this means that he consents to taking away the last \$9 of a poor person, in order to increase an income which equals the mean by \$10. It is easily verified that the same client is fully compensated for taking away the last \$8 of a poor person by adding \$10 to an income which is double the mean. In practice, this means that moderately Paretian clients must not use σ^2 as a measure of inequality, at least not in the sense of (3). Actually, it is doubtful whether even boldly non-Paretian clients may use it. Suppose that for a given distribution the function is non-Paretian at the income $\alpha = KM$. Now, any function of the form $W(S, \sigma^2)$, $\partial W/\partial\sigma^2 < 0$, may be written as $W^*(S, \sigma^2/M^2)$, $\partial W^*/\partial(\sigma^2/M^2) < 0$, because a transfer increases the coefficient of variation if, and only if, it increases the variance. By assumption,

$$\begin{aligned} \frac{\partial W}{\partial x_{KM}} = \frac{\partial W^*}{\partial x_{KM}} &= \frac{\partial W^*}{\partial S} + \frac{\partial W^*}{\partial(\sigma^2/M^2)} \frac{\partial(\sigma^2/M^2)}{\partial x_{KM}} \\ &= \frac{\partial W^*}{\partial S} + \frac{\partial W^*}{\partial(\sigma^2/M^2)} \frac{2}{S}\left(K - 1 - \frac{\sigma^2}{M^2}\right) < 0. \end{aligned} \tag{17}$$

Since $\partial W^*/\partial(\sigma^2/M^2)$ is negative and S is positive, this is possible only if $K-1 > \sigma^2/M^2$ or if $\partial W^*/\partial S < 0$. In other words, if $K-1$ is smaller than σ^2/M^2, the client

TABLE 1

FOUR MEASURES OF INEQUALITY

	Symbol	Definition	Range for non-negative incomes	$\frac{\partial\theta^*}{\partial x_i}$	$\frac{\partial(\partial\theta/\partial x_i)^*}{\partial x_i}$ for constant parameters
Variance	σ^2	$\frac{1}{n}\sum_{i=1}^{n}(x_i-M)^2$	0, ∞	$\frac{2}{n}(x_i-M)$	$\frac{2}{n}$
Logarithmic variance	λ^2	$\frac{1}{n}\sum_{i=1}^{n}(\log x_i-\mu)^2$ $\mu=\frac{1}{n}\sum_{i=1}^{n}\log x_i,\ x_i>0$	0, ∞	$\frac{2}{n}\frac{\log x_i-\mu}{x_i}$	$\frac{2}{n}\frac{1+\mu-\log x_i}{x_i^2}$
Coefficient of concentration	R	$\frac{1}{2M}\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\lvert x_i-x_j\rvert^{\dagger}}{n(n-1)}$	0, 1	$\frac{1}{S}\left[\frac{2}{n-1}\left(i-\frac{n+1}{2}\right)-R\right]^{\dagger}$	0
Relative mean deviation	T	$\frac{1}{2M}\frac{\sum_{i=1}^{n}\lvert x_i-M\rvert}{n}$	0, 1	$-\frac{1}{S}\left(\frac{n-m}{n}+T\right),\ x_i<M$‡ $\frac{1}{S}\left(\frac{m}{n}-T\right),\ x_i>M$	0

*In certain cases, as will be clear from the following notes, R and T may have only one-sided derivatives. This is, however, enough for our purpose.

†Incomes are here ordered by size. If there exist several incomes equal to x_i, i will be the highest serial numver in this group (i.e. n times the distribution function) for a marginal increase in x_i; for a marginal decrease in x_i, i will be the lowest serial number in the group.

‡The letter m denotes the number of incomes not exceeding the mean (for a marginal increase in x_i) or the number of incomes smaller than the mean (for a marginal decrease in x_i). In the case of a transfer, however, m may be taken uniquely in the first sense (since a marginal increase of one income increases M too). If $x_i=M$ then the first expression holds for a marginal decrease, the second for a marginal increase in x_i.

will reject a proportional increase in *all* incomes. A client who is non-Paretian for incomes of twice the mean ($K-1=1$) or more will find himself in such a queer situation as regards many empirical distributions. It therefore seems that we may not as a rule use σ^2 and its derived indexes to abbreviate welfare functions by (3).

b. The Logarithmic Variance

The derivative in Table 1 shows that a client with $W(S, \lambda^2)$ may be non-Paretian for all incomes above the geometric mean. Paretianity is also allowed, in this case with no special trouble.

Trouble may (and in most cases will) nevertheless arise. For, as follows from the second derivative of λ^2, it does not fulfil the Dalton condition, even in its weak form (9), for any two incomes which are both higher than 2.718 ($=e$) times the geometric mean. That is, λ^2 will *increase* if income is transferred from a very rich man to a less rich one. It follows from what has been said above that λ^2 and non-intersecting Lorenz curves will in general assign different ranks to a set of distributions by their "degree of inequality."[17] We may infer that the use of λ^2 fits only those equality loving clients who believe that, if there must be poverty and riches, it is better to concentrate riches in as few hands as possible. It goes without saying that all these conclusions apply to the measure antilog λ, which may be considered as the geometrico-quadratic mean of the ratios between the incomes and their geometric mean.[18]

c. The Gini Coefficient of Concentration

The first thing to note is that R is a strictly Daltonian measure. Table 1 shows that, for any given distribution, $\partial R/\partial x_i$ is a strictly increasing function of i, the incomes being here numbered according to their size.[19] Thus a transfer from a higher to a lower income will always decrease R (if it is computed from the original, ungrouped, data).[20] Table 1 also shows that Paretians, as such, will have no problem with R: $\partial R/\partial x_i$ is bounded from above and its rate of growth does not

[17]Another example of deviation from the Lorenz-curve-ranking gives the measure $(\log M)/\mu$. This measure clearly fulfils the Dalton condition (indeed, it was recommended by Dalton himself [7] 9–10), but a proportional increase in all incomes decreases it. A simple welfare function which fits it is $w=\prod_i x_i$.

[18]They apply also to the simple geometric mean of ratios (taken always in the downward direction), antilog $1/n \sum_{i=1}^{n} |\log x_i - \mu|$. This measure is non-Daltonian above the geometric mean. If a client is more interested in the ratios than in the differences between incomes and is nevertheless a Daltonian, then his welfare function may be perhaps approximately abbreviated by one of Theil's indexes. These are the logarithms of the geometric means

$$\sqrt[n]{\prod_{i=1}^{n} M/x_i}, \qquad \sqrt[\Sigma x_i]{\prod_{i=1}^{n} (x_i/M)^{x_i}}.$$

Cf. [16], Ch. 4. The second index has the interesting and unique property that the increment in it effected by a unit transfer between two incomes depends only on the ratio between them. This "entropy" measure has therefore, contrary to Sen's opinion, [15] 35–36, a clear (though very special) intuitive basis.

[19]The second derivative is of no use here, since R is not a smooth function of the incomes.

[20]Obviously, transfers within income groups will not alter R. It follows that R (together with σ^2 and λ^2) computed from groups is not too well defined as an inequality measure. There are as many types of R as there are ways of grouping, and different types may behave very differently—a fact which has sometimes led to misinterpretation of income statistics. Cf. [12] 81, 84.

depend on the size of the incomes. Non-Paretians, however, have such problems, since (for a sufficiently great n) $\partial R/\partial x_i$ will be positive only if $\frac{1}{2}(1+R)$ is not greater than the distribution function at x_i. Thus, for example, if the client is non-Paretian with respect to the upper third of the incomes and R is greater than $\frac{1}{3}$ (as it will be in most contemporary societies), then by (13), it is impossible to describe his welfare function by any $W(S, R)$.

Now suppose that our client is both Daltonian and Paretian. It is not nonetheless unlikely that R will fit his $w(\ )$ tolerably, because of the strange property of $\partial R/\partial x_i$ that it depends only on the rank of the income in the distribution, and not on its size. As a result, the desirability (dw) of a small transfer between any two incomes must here too depend only on their rank and be independent of their size.[21,22] A convenient test of fit for R is therefore to ask the client: "The first quartile income is $4,000; the median is $5,000; and the third quartile is $10,000. You may transfer $100 from the third quartile to the median or from the median to the first quartile. Do you decisively prefer one transfer over the other or are you (nearly) indifferent between them?" Most clients, I think, will have clearcut preferences in such cases, so that R will not fit their $w(\ )$ even approximately.[23]

d. The Relative Mean Deviation

This measure is today better known as the Maximum Equalization Percentage, being indeed that percentage of the total income which would have to be transferred from incomes above the mean to incomes below it in order to achieve perfect equality.[24]

Both Paretians and non-Paretians may use T and in this connection there does not seem to be much to say against it. Transfer tests, however, will in general disqualify this measure. For, as is widely known and as is clear from Table 1, T is absolutely insensitive to transfers which do not pass through the mean, that is to transfers between any two incomes which are both below or both above the mean. This is a striking violation of the strong Dalton condition (7) and much more than that. $W(S, T)$ may be a rough description of the values of only those people who distinguish in the population between only two broad groups: the "rich" and the "poor."

6. Conclusion

In this paper various formulas were suggested for testing approximations to an individual's welfare function by the abbreviated form $W(S, \theta)$, θ being a given

[21] Cf. Atkinson [2] 256, Bentzel [3] 262.

[22] There are several formulas for R which exhibit this property directly. Perhaps the most striking of them is

$$R = \frac{2}{(n-1)S} \sum_{i=1}^{n} |i-m||x_i - M|$$

which gives R as a weighted relative mean deviation about the mean (m/n being the distribution function at the mean). An analogous formula, due to Gini [8] 237–238, defines R as a weighted relative mean deviation about the median.

[23] Newbery [14] proved that there exists no utilitarian welfare function which fits R. This is, however, beside the point as such welfare functions are implausible. For welfare functions which fit R, see Kats [10].

[24] For the rather curious literature on T, see [12] 88–92.

inequality measure. It can be shown, although it was not done here,[25] that, on fairly reasonable assumptions, such abbreviation is always possible, for *some* θ peculiar to the individual's $w(\ )$. This possibility is, however, not enough to make our abbreviated welfare functions usable in social discussion and planning of the income distribution. For if it is to be discussed at all, a θ must be *common* to the various welfare functions confronted. Is it possible to find such a common measure of inequality? The results of the previous section are not encouraging in this respect. They show that any one of the four widely used measures investigated fits only the welfare functions of special groups of clients.

There exists, of course, an infinity of other measures, and many of them may not be disqualified by general, qualitative tests. Whether such a measure is suitable for abbreviation of all, or almost all, relevant welfare functions is an empirical problem which can be investigated by the method described in Section 3.[26] Obviously, the various welfare functions must have much in common in order to make the use of a common measure of inequality plausible. However, this restriction on the shape of the welfare functions will be weaker, the more we constrain the relevant set of distributions. Especially important here are constraints on the tails of the distribution (e.g. an institutionally rigid minimum income), since the tails are likely to be the principal source of differences in the weighting of incomes in $w(\ )$ (and in θ).

More generally, if we cannot hope to proceed with $W(S, \theta)$, we can use a less severe abbreviation, $W(S, \theta_1, \theta, \theta_2)$, where θ_1, θ_2 are parameters of poverty and riches, while θ remains a measure of "general" inequality (or, alternatively, a measure of inequality in the truncated, tailless population). There are good prospects of finding tolerable common parameters for such an abbreviation. In this broader setting, however, our findings here are only partly valid.

References

[1] Aigner, D. J. and Heins, A. J., "A Social Welfare View of the Measurement of Income Inequality," *The Review of Income and Wealth*, 13 (1967), 12–25.

[2] Atkinson, A. B., "On the Measurement of Inequality," *Journal of Economic Theory*, 2 (1970), 244–263.

[3] Bentzel, R., "The Social Significance of Income Distribution," *The Review of Income and Wealth*, 16 (1970), 253–264.

[4] Bergson, A., "A Reformulation of Certain Aspects of Welfare Economics," *Quarterly Journal of Economics*, 52 (1938), 310–334.

[5] Bergson, A., "On the Concept of Social Welfare," *Quarterly Journal of Economics*, 68 (1954), 233–252.

[6] Blum, W. J., and Kalven, H., *The Uneasy Case for Progressive Taxation*, Univ. of Chicago Press, Chicago, 1953.

[7] Dalton, H., "The Measurement of the Inequality of Incomes," *Economic Journal* (1920), as reprinted in Dalton, H., *The Inequality of Incomes* (2nd Ed.), Routledge and Kegan Paul, London, 1925.

[25] It was done in [12] 32–35.

[26] Atkinson [2], who rightly criticized the conventional inequality measures, proposed a well-known new one, I, depending on a parameter ε, called the "degree of inequality aversion." This is not a fortunate term, for in fact ε measures the sensitivity of I to poverty *and* its insensitivity to riches. However, I is strongly "poverty biased" for any ε in the following sense: given various pairs of incomes with the same ratio between them, an equalizing unit transfer within a pair will decrease I the more, the lower that pair is located in the distribution. Hence, not only will clients having different ε's be devoid of a common language, but it is unlikely that I, for any ε, fits widely held preferences at all.

[8] Gini, C., *Variabilità e Concentrazione*, Sec. Ed., Veschi, Roma, 1955.
[9] Gorman, W. M., "Are Social Indifference Curves Convex?" *Quarterly Journal of Economics*, 73 (1959), 485–496.
[10] Kats, A., "On the Social Welfare Function and the Parameters of Income Distribution," *Journal of Economic Theory*, 4 (1972), 377–382.
[11] Kolm, S. Ch., "The Optimal Production of Social Justice," in J. Margolis and H. Guitton (Eds.), *Public Economics*, Macmillan, London, 1969, 145–200.
[12] Kondor, Y., *Inequality in the Size Distribution of Incomes; Its Measurement and Planning*, D.Sc. Thesis, Technion, Haifa, 1970 (Hebrew, with English summary).
[13] Little, I. M. D., "Social Choice and Individual Values," *The Journal of Political Economy*, 60 (1952), 422–432.
[14] Newbery, D., "A Theorem on the Measurement of Inequality," *Journal of Economic Theory*, 2 (1970), 264–266.
[15] Sen, A., *On Economic Inequality*, Clarendon Press, Oxford, 1973.
[16] Theil, H., *Economics and Information Theory*, North-Holland Publ. Co., Amsterdam, 1967.
[17] Tobin, J., "Liquidity Preference as Behaviour Towards Risk," *The Review of Economic Studies*, No. 67 (Feb. 1958), 65–86.
[18] Vickrey, W., "Utility, Strategy and Social Decision Rules," *Quarterly Journal of Economics*, 74 (1960), 507–535.