JOURNAL OF ECONOMIC THEORY 12, 416-442 (1976)

Unequal Inequalities. I

SERGE-CHRISTOPHE KOLM

EHESS and CEPREMAP, Paris, France

Received March 7, 1974; revised December 2, 1975

INTRODUCTION

Many people consider the reduction of economic inequalites as a basic aim of society. Such ideas are, however, largely nonoperational, sterile, and even meaningless, as long as what is called inequality is not stated with precision. This is so because, as well appear below, different measures of inequality give widely different, and even opposite, results. Such policy which diminishes some apparently reasonable measure increases other ones. And I can take any country and prove that in some period (whatever it is) inequality has increased or decreased in it, or any two countries and prove that inequality is higher in the one or in the other, by choosing different inequality measures, all of which would probably seem good and valuable at first sight. Furthermore, this embarrassing situation happens not only when we consider the complex social, or socioeconomic, or economic position of persons, or even their consumption–labor bundles over time or even timeless, but even in the simplest case where the variable is a unique magnitude (and, even, a quantity) per person, such as its income or its wealth.[1]

It thus seems essential to appraise the economic and, if we dare say, ethical, implications of the inequality measures, and to build measures embodying the economic and ethical properties we feel inequality means. Several economists, among them Pigou [1], Dalton [2, 3], Taussig [4], Cannan [5], and Loria [6], have stated their views about whether some well-defined changes in incomes increase or decrease their inequality; these are implicit properties of inequality measures or indices. Dalton also added more specific but less valid structures for them (see [7, Sect. VI]). A number of other properties and their relations were introduced in

[1] I wish to thank Professor A. Atkinson for having provided the motivation for writing down these ideas, invaluable linguistic corrections, very valid (if not met) criticism, and important ideas which will be mentioned in the text.

[7, Sects. VI and VII],[2] along with some explicit inequality indices. Atkinson ([8]) then presented anew some of these relations, singled out one of these indices, and initiated the empirical application of it, which was followed in a number of other studies. This will be the starting point of this paper, Part II of which will appear in the next issue of this journal.

In Part I, we first consider two specific measures labeled "rightist" and "leftist," their opposite and common properties (Sect. I), and we derive them axiomatically from sets of basic properties (Sect. II). They happen to be extreme cases of a more general, "centrist," class of measures (Sect. III). We then discuss the question of "per person" versus "per pound" inequality (Sect. IV), and we find how these measures vary with their parameters and with equal absolute and relative variations in all incomes (Sect. V). Part II will then consider broader measures and further properties. Inequality invariance under these two latter kinds of income variations brings out the difference between the rightist and leftist measures; but if we abandon some property common to both ("independence"), these invariances will be reconciled by a general class of measures, among which are the standard deviation and coefficient of variation (Sect. VI). Pigou and Dalton's "principle of transfers" (i.e., a transfer of one pound from a richer person to a poorer one diminishes inequality) will be extended into a "principle of diminishing transfers" (Sect. VII), and relations between inequality measures and Lorenz and "concentration" curves will be analyzed (Sect. VIII). We will then turn to the important question of how inequality is affected by addition of incomes (or taking from them, or growth in them), and to a law of "diminishing returns in equality" (Sect. IX). And the lumping together of different populations also implies a specific relation between inequality measures (Sect. X). Section XI will finally present the relations between the most general properties of inequality measures (and of the "social welfare functions" they imply).

The properties and results are presented here for income distributions. But almost all of them interestingly apply to distributions of other things. Most notable is their immediate interpretation in terms of risk and uncertainty analysis, which we do not state explicitly because it is thoroughly straightforward and would be cumbersome (furthermore, a part of this translation appeared in [9]). The application to the com-

[2] Unfortunately, a systematic misprint in the beginning of Section VI of this paper makes its reading difficult: The preference relations $\succ$ and $\succcurlyeq$ have everywhere been written as $>$ and $\geqslant$, and they sometimes appear in the same definitions and theorems as "larger than." This paper was a contribution to a conference (I.E.A., Biarritz, 1966), and the author was not presented with the galley proofs of the publication to check.

parison of growth paths (distribution of dated magnitudes) is also fruitful. And the results can of course also be applied to other interpersonal distributions (wealth, consumption, etc.). The next step consists in passing from unidimensional to multidimensional (or multivariate) distributions; this is done in [10], which also reviews the history of the literature about the various economic applications of the ideas of "Schur-concavity," "rectifiance," "isophily," "concentrations preferences," "averages preferences," "stochastic dominance," "transfers principle," etc.

I. Rightist vs Leftist Measures of Inequality

I.a. *Two Conceptions of Inequality*

Inspired by a recent paper by Atkinson [8], a rapidly increasing number of studies use the index of inequality

$$I_r = 1 - \left[\sum (x_i/\bar{x})^{1-\epsilon} f(x_i)\right]^{1/(1-\epsilon)},$$

or

$$I_r = 1 - \prod (x_i/\bar{x})^{f(x_i)},$$

which is its limit when ϵ tends to zero ($\prod$ being the product sign), to measure inequalities in the distribution of income. x_i is then an income level, $f(x_i)$ is the proportion of persons whose income has this level ($\sum f(x_i) = 1$), $\bar{x}$ is average income ($\bar{x} = \sum x_i f(x_i)$), ϵ is a nonnegative coefficient (if $\epsilon = 0$, $I = 0$, if ϵ tends to infinity, I tends to $1 - (\underline{x}/\bar{x})$, where $\underline{x}$ is the smallest of the x_i's); $I_r = 0$ when there is no inequality (all x_i's are equal and equal to $\bar{x}$), and only in this case if $\epsilon > 0$. Professor Atkinson himself provides in the same paper an extremely interesting application of this index to an intercountries comparison of income distribution inequalities.[3] Jakobsson and Normann have used this index to analyze the redistributive effects of tax policies in Sweden [11]. A very important work by Bruno and Habib uses this index to evaluate the effect of income tax-transfer schemes on income inequality, and apply it to data of Israel [12].[4] Another very important work by Bruno uses it to measure the effects of education policies on income inequality, with, again, numerical application with data from this country [13]. This index

[3] It should be recalled that Professor Atkinson also discusses properties of other measures of inequality in this article (among them what will be called I_l below).

[4] Their qualitative results remain the same, however, for the other measures of inequality discussed below.

is also being used for figures of Canada and of various other European countries. It thus is timely to recall the social implications of this measure and its alternatives.

In May 1968 in France, radical students triggered a student upheaveal which induced a workers' general strike. All this was ended by the Grenelle agreements which decreed a 13% increase in all payrolls. Thus, laborers earning 80 pounds a month received 10 pounds more, whereas executives who already earned 800 pounds a month received 100 pounds more. The Radicals felt bitter and cheated; in their view, this widely increased incomes inequality. But this would have left unchanged an inequality index I_r computed according to the above formula.

This is so because this formula implies that, if all incomes x_i are multiplied by the same number, I_r does not change. In other countries (I have been quoted examples from England and The Netherlands), trade unions are more clever and often insist on equal absolute, rather than relative, increases in remuneration, so as to avoid the above effect. And I have found many people who feel that it is an equal absolute increase in all incomes which does not augment inequality,[5] whereas an equiproportional increase makes income distribution less equal or more unequal—and these were people of moderate views. Now, the index

$$I_l = (1/\alpha) \log \sum e^{\alpha \cdot (\bar{x} - x_i)} f(x_i)$$

does not change when all incomes x_i are increased by the same amount. $\bar{x}$ and $f(x_i)$ have the same meaning as previously, α is a nonnegative parameter ($I_l = 0$ if $\alpha = 0$, and when α tends to infinity I_l tends to $\bar{x} - \underline{x}$, i.e., the gap between average and minimum income); $I_l = 0$ when there is no inequality (all x_i's are equal and equal to $\bar{x}$), and only in this case if $\alpha > 0$.

When all incomes x_i are multiplied by the same number, whereas I_r does not change, I_l is multplied by this number. Therefore, if we study variations of I_l over time in an inflationary country, we must call x_i the real incomes, discounted for inflation; or if we make international comparisons of I_l, we must use the correct exchange rates. This need not be done if we use I_r. But these problems are exactly the same ones which are traditionally encountered in the comparisons of national or *per capita*

[5] The topic was an equal increase in all incomes rather than an equal decrease in them. But it is the first point which is relevant in our progressive societies. Anyway, all that is said here is that it is no less legitimate to attach the inequality between two incomes to their difference than to their ratio. One view must not be judged from the other's prejudice. The term "leftist" used below must not be taken too literally: the measure corresponds to this view of society in very important cases, not in all imaginable ones.

incomes ($\bar{x}$), and they can be given the same traditional solutions. Anyway, convenience could not be an alibi for endorsing injustice.

These two measures are reconciled, in the sense that they are equal to each other for all income distributions, if and only if ϵ and α are zero, which entails $I_r = I_l = 0$ for all distributions, that is, when no inequality is felt (this is also a necessary and sufficient condition for each measure to remain the same for all distributions). When ϵ and α tend to infinity, both measures classify distributions with equal $\bar{x}$ according to $\underline{x}$. Of course, when all x_i's are equal, these two measures are equal whatever ϵ and α since they are zero. Otherwise, the very properties mentioned (I_r invariant when all incomes are multiplied by the same number, I_l invariant when the same amount is added to all incomes) show that these measures, and more significantly their variations over time or societies (nations, regions, professions, etc.), will differ widely.

The economic literature is, of course, relatively rich in opinions about the effects of equal or equiproportional variations in incomes on the inequality of their distribution. They roughly tend to support Abba Lerner's contention that economic science tends to shift its servents to the right. For instance, Taussig [4, p. 485] feels that a variation of all incomes in the same proportion does not change inequality; this is I_r's property. Loria [6, p. 369], Cannan [5, p. 137], and Dalton [2, 3] feel that an equal addition to all incomes decreases inequality; I_l of course does not satisfy this condition, whereas it will be shown in Section V below that I_r does. For Dalton [2, 3] again, an increase of all incomes in the same proportion increases inequality; I_r of course does not satisfy this property, whereas it will be shown in Section V below that I_l does. From this we see that Dalton would have liked neither I_r nor I_l. But the "centrist" measures of inequality presented in Section IV below might suit his taste, since Section V will show that they satisfy both his requirements. Finally, [7, Sects. VI and VII] presents a systematic exposition of properties of inequality measures (there called "unjustice") and of their implications; in fact, a large part of the present article is only an elaborate exposition of some of its points.

I.b. *Differences and Similarities between These Two Measures*

The two initially mentioned properties may be called "inequality invariance under equiproportional, resp. equal absolute, variations" of incomes. The first one also says that I_r is an "intensive" magnitude in the physicists' sense.[6] Measures I_r and I_l differ from these viewpoints, but they also have in common several interesting properties, whereas

[6] This property was thus called "intensive justice" in [7].

there exist still other significiant properties according to which these measures differ and which give an interest to one or the other. We shall now mention the properties in these two categories; that they hold will be obvious for some but not for others. In the latter case, the proofs will be delayed to the following sections. Also, this presentation will exclude the trivial cases $\epsilon = 0$ and $\alpha = 0$,

Among the properties I_r and I_l have in common, they are zero when all incomes are equal and positive otherwise (ϵ and α being positive). In both these measures, too, incomes are distinguished only by their level, and not in any other way by their recipients; these measures in this sense respect principles of "equal treatment of equals," or "horizontal equity," or "impartiality" (the label we shall keep); in other words, a permutation of incomes between income recipients does not affect the inequality of the distribution, that is, the inequality measure is a symmetric function of incomes.

Both measures also satisfy a fundamental series of equivalent properties. One is that a transfer of one pound from a richer person to a poorer one decreases inequality; this is Dalton's "principle of transfers" ([7]'s "rectifiance," also previously mentioned by Pigou in [1]). Another one is that inequality is lower when the Lorenz curve is everywhere higher, for two distributions of the same total income (or, more generally, [7]'s "isophily," which also applies when the latter condition does not hold and consists in comparing concentration curves, i.e., the sums of the jth smallest incomes for all j's; cf. Part II, Sect. VIII). Other properties equivalent to these two are presented in [7, 8, 10]. In mathematical terms, all this just means that I_r and I_l are Schur-convex functions of the set of individual incomes.[7]

[7] We say that an inequality measure function $I(x)$ is *rectifiant* when $x_i < x_j$ and $0 < h \leqslant (x_j - x_i)/2$ imply that the replacement of x_i by $x_i + h$ and of x_j by $x_j - h$ decreases I, for all pairs i, j and all x's. This is strict rectifiance; for weak rectifiance, I would not be increased by this transfer. If I is symmetric, the condition on h can equivalently be replaced by $0 < h < x_j - x_i$. If I is differentiable, weak and strict rectifiance are respectively equivalent to $x_i < x_j$ implying $\partial I/\partial x_i \leqslant \partial I/\partial x_j$, and $\partial I/\partial x_i < \partial I/\partial x_j$ almost everywhere. Taking an h tending to zero shows the necessity, applying to successive small variations in h shows the sufficiency. Rectifiance plus symmetry is equivalent to $I(Bx) \leqslant I(x)$ for all x and all bistochastic matrix B (i.e., $B = [b_{ij}]$, $b_{ij} \geqslant 0$, $\sum_i b_{ij} = \sum_j b_{ij} = 1$, for all i, j's) for the weak form, and to a similar relation with strict inequality if Bx's coordinates are not a permutation of x's for the strict one.

These properties were used in [7 Sects. VI and VII] in justice theory and in [10, p. 110–119] in risk and uncertainty theory and portfolio choice analysis (x_i then being wealth in eventuality i with a definition of the i's such that they all have the same probability of occurrence), both for I and for evaluation functions $V(x)$ with a reversal of the inequality signs—these properties of these two functions are equivalent when

A refinement of this "transfers principle" is that the decrease in inequality obtained by a one pound transfer to an income lower by a given amount is larger when these two incomes are smaller. Both I_r and I_l satisfy this "principle of diminishing transfers" (cf. Part II, Sect. VII).

Both measures also satisfy another of Dalton's principles, the "principle of proportionate additions to persons," i.e., the inequality does not change when the numbers of income recipients in each income class vary in the same proportion (hence, total population also varies in this proportion). For instance, if society is exactly duplicated, the numbers of persons in each income class and in total being doubled, inequality in the income distribution remains the same. The reason is that the numbers of persons only enter in I_r and I_l through the $f(x_i)$'s which are their *ratios* to total population.

But, on the other hand, I_r and I_l differ for the following interesting properties, although we shall find that they still have common points in some of these questions.

One property refers to the addition of incomes of several kinds (the "composition" of their distribution, in statistical jargon). For instance, how does income inequality relate to inequalities in earned and unearned incomes? Or, how does the inequality in income increments from the last to the present year make income inequality vary? Or, how does fiscal inequality transform before tax into after tax income inequalities? Or, how does the inequality in government transfers affect the inequality in incomes? Or, how do the inequalities in "private income" and in the individuals' benefits from government activities combine to determine the inequality in the distribution of a more general income concept encompassing both? Generally, if an income distribution is a sum ("composition") of several distributions, what relation is there between its inequality and its components'? The property, which holds for I_r, is the following: *Inequality in a sum of distributions is lower than the sum of these distributions' inequalities* weighted by total or average incomes in each, except when these distributions are proportional (in which case all the inequality measures mentioned are equal). If we consider another index

---

the relations between V and I are the ones presented below (for the risk theory case, I would be a risk-premium or an insurance premium). The equivalence between rectifiance and symmetry on the one hand and the property mentioned on the other hand was first proved by Ostrowski in the differentiability case ([14, Theorem VII]). Rectifiance plus symmetry is Schur-convexity of I or Schur-concavity of V, since functions satisfying the equivalent property for the concavity case and positive variables were introduced by I. Schur in his generalization of Hadamard, Julia, and Parodi's inequalities derived from positive Hermitian forms [15, 16, 17] (Schur also showed the necessity of the marginal rectifiance condition).

$\bar{x}I_r$ or $n\bar{x}I_r$ (n being the total number of income recipients), the property says that inequality in the sum is lower than the (unweighted) sum of inequalities, except when the constituting distributions are proportional, in which case inequality in the sum is the sum of inequalities (cf. Part II. Sect. IX). This property may be called "nonincreasing inequality under additions of distributions," or "subadditivity" of the inequality measure. It might be an appropriate description because one may feel that the composition of differently unequal distributions evens out the inequality in some sense and in some degree. No such property holds for I_l.

There is, on the other hand, another, somehow related, property, which I_l has and I_r has not, although I_r has it if we consider only distributions of the same total amount, and $\bar{x}I_r$ and $n\bar{x}I_r$ have it. It appears when we consider several additions of the same distribution to another given one (for instance a sequence of income increments), and more generally what is "added" could now include some decrease (a "negative addition") for some (or all) incomes. The property then says that *inequality increments are larger and larger* (*or decrements smaller and smaller*). There is a limiting exception which we already know: When added incomes are the same for all individuals, I_l does not change (also, when they are proportional to the initial distribution, the increments in $\bar{x}I_r$ and $n\bar{x}I_r$ are constant). For I_l, if in particular we take as the initial distribution the one in which all incomes are zero (the "null distribution"), this property shows that *a doubling of all incomes more than doubles inequality* if it is not zero (and a similar property for a multiplication by any positive number λ). The same property can be expressed on operations of proportional bridging of gaps between distributions: If we bridge half the gap between two distributions, for all incomes, the resulting inequality is less than the (arithmetic) average of the two respective inequalities. The exception for I_l then is the case when the gap is the same for all incomes: The two initial distributions and the halfway one all have the same inequality (for $\bar{x}I_r$ and $n\bar{x}I_r$, the exception is the case when the two distributions are proportional: inequality of halfway is halfway inequality). In particular, bridging half the gap from an unequal distribution to an equal one diminishes inequality by more than half (the equal one could for instance be a completely equalizing redistribution of the initial one). These properties constitute a "law of increasing average and marginal inequality," or of "diminishing returns in equality." Mathematically, they are a case of convexity of the inequality measure as a function of all incomes (cf. Part II, Sect. IX).

A consequence of these properties is that, for I_l, for $\bar{x}I_r$, and also for I_r when we consider distributions of same total or average income, if we bridge part of the gap between two distributions, in the same proportion for all incomes (for instance each income is the arithmetic average of what it is in these two distributions), the inequality is not larger than the largest one for the two initial distributions (this is a property of "quasi-convexity"; cf. Part II, Sect. IX).

Instead of adding persons' incomes, we may add persons. That is, we may consider lumping together several populations, or, conversely, partitioning a society into several subpopulations. How does total inequality relate to the inequalities of the various subpopulations? How does European inequality relate to inequalities in the various European countries? This question requires an extension of Dalton's "principle of proportionate additions to persons," which applies to the special case when all constituting populations are identical except for the multiplication of the number of people in each income class and in total by some number. It must generally be untrue that if all constituting populations present the same degree of inequality, the global population also shows this degree of inequality. Since if each of the constituting populations has a perfectly equal income distribution, but personal incomes differ from one to the other, the distribution will not be equal in the aggregated society. In brief, inequality in the global population will have two sources: inequalities *within* the constituting populations, and inequalities *between* them. From this viewpoint, I_r and I_l are more similar than different. For each of them, total inequality is not lower than a weighted sum of the subpopulations' inequalities, but the weights are the number of persons for I_l and total incomes for I_r. For both, if the subpopulations present the same degree of inequality, the aggregated population also has this degree if and ony if the subpopulations also have the same average income (cf. Part II, Sect. X).

We have noted earlier that both I_l and I_r have the same relation to the disposition of Lorenz curves for distributions of the same total amount. But when the total amounts of the distributions differ, so do these measures in this respect. With I_r as the measure, a distribution whose Lorenz curve is nowhere under and somewhere above another distribution's curve has a smaller inequality, whatever their total or average incomes. This is not so for I_l. What can be said for this measure is that it is lower if both the Lorenz curve is nowhere under and somewhere above, and total or average income is not larger (cf. Part II, Sect. VIII).

Finally, the formulas show that we cannot write negative incomes (x_i) into I_r, whereas this poses no problem in I_l. Now, business failures make negative incomes a reality.

I.c. *Similar Properties of Other Inequality Measures*

One then naturally wonders how the other measures of inequality fare in front of these properties. We shall see this in detail for the "centrist" measures presented below, which are intermediate cases between I_r and I_l and contain them as special limiting cases. As for the more conventional measures, an interesting result is that the standard deviation of incomes $\sigma = [\sum (x_i - \bar{x})^2 \cdot f(x_i)]^{1/2}$, and its division by average income which is the coefficient of variation $\sigma/\bar{x}$, satisfy almost all properties in the following sense (some of these properties are obvious, and the other ones will be proved in the following sections).

σ and $\sigma/\bar{x}$ are zero at equality and positive elsewhere. They satisfy the "impartiality" or symmetry property. The multiplication of all incomes by the same number does not change $\sigma/\bar{x}$. The addition of the same amount to all incomes does not change σ. σ and $\sigma/\bar{x}$ satisfy Dalton's "principle of transfers" and its equivalent properties, such as lower value for distributions of the same amount with uniformly higher Lorenz curves. Furthermore, for $\sigma/\bar{x}$ this latter property holds even if the distributions do not have the same total income. However, their decrease for a one pound transfer to an income smaller by a given amount is proportional to this amount and thus independent of the income levels. Both satisfy Dalton's "principle of proportionate additions to persons." The σ of a sum (composition) of nonproportional unequal distributions with $n > 2$ is lower than the sum of their σ's. σ, and $\sigma/\bar{x}$ given $\bar{x}$, satisfy the "diminishing returns properties" because they are convex functions of the set of incomes: Successive identical variations in incomes give these measures increasing increments, or decreasing decreases, except, for σ, when they are proportional to the initial distribution or when the variation is the same for all incomes; the measure for halfway between two distributions neither proportional nor with the same difference for all incomes is less than the average of their measures; to bridge half the gap from an unequal distribution to an equal (nonnull) one diminishes the measure by more than a half. Furthermore, the values of σ and $\sigma/\bar{x}$ for a union of populations are higher than the sums of their values for these populations, respectively weighted by the numbers of persons and total incomes, except when all these populations have the same σ's and $\bar{x}$'s and $\sigma/\bar{x}$'s, in which case the union's corresponding values are also the same. Finally, negative incomes can be perfectly included in the computations of σ and $\sigma/\bar{x}$ (we assume positive average $\bar{x}$, however).

II. An Axiomatic of These Two Measures

To really see what these measures imply, it is necessary to build an axiomatic of them, i.e., to find for each one a set of properties which are equivalent to its adoption. If these properties are small in number and as intuitive as possible, they will display the implicit assumptions made by this choice.

We shall use the following notations. i will now be the index of an income recipient unit (called a "person"), rather than of an income class (but the difference is only formal). There are n i's. x_i is i's income. x is the n-vector of the x_i's. $\bar{x} = (1/n) \sum x_i$ is the average income. e is the n-vector each coordinate of which is 1. λ and μ are scalars (real numbers). A measure or index of inequality is a function $I(x)$. We consider, for this function, the following properties.

II.a. *Properties*

(1) When all incomes are equal, $I = 0$. This is a natural requirement for a measure of inequality. It can be written as $I(\lambda e) = 0$ for all admissible λ's.

(2) When incomes are unequal, $I > 0$. This also is a natural requirement for I. At the limit, we would have $I \geqslant 0$.

(3) "Impartiality": I is left unchanged by a permutation of the x_i's, i.e., it is a symmetrical function of the x_i's. This property is unavoidable as long as income recipients are not distinguished by anything else but their income. It is akin to the old "equal treatment of equals" and "horizontal equity" principles of public finance.

(4) "Transfers principle" (Pigou, Dalton, etc.): The transfer of a pound from a richer person to a poorer one decreases inequality. More precisely, this is I's "rectifiance": $(\partial I/\partial x_i) - (\partial I/\partial x_j)(x_i - x_j) > 0$ if $x_i \neq x_j$. At the limit, this transfer could have no effect and we would write the inequality as $\geqslant$ for all x_i and x_j's (that is, rectifiance can be strict or weak).

(5) The addition (or subtraction) of the same amount to all incomes does not change I. That is, $I(x + \mu e) = I(x)$ for all admissible μ's and x's.

(6) "Intensive inequality": The multiplication of all incomes by the same scalar (an equiproportional variation in all incomes) does not change I. That is, $I(\lambda x) = I(x)$ for all admissible λ's and x's.

(7) $[\partial(\bar{x} - I)/\partial x_i]/[\partial(\bar{x} - I)/\partial x_j]$ does not depend upon x_k for all k's either i or j, all i's and j's, and all x's.

(8) $\{\partial[(1 - I)\,\bar{x}]/\partial x_i\}/\{\partial[(1 - I)\,\bar{x}]/\partial x_j\}$ does not depend upon x_k for all k's either i or j, all i's and j's, and all x's.

To evaluate (7) and (8), we observe that $\bar{x} - I$ or $(1 - I)\,\bar{x}$ can be considered as indices of social welfare, with two different concepts of inequality measurement, one which is in some sense absolute, and the other one which is in some sense relative. Those two concepts will be discussed in detail in Section IV below. Presently, it suffices to remark that in the first case inequality is the number of pounds sterling which must be deducted from average income in order to obtain a measure of global welfare which takes into account both the average per capita income and the inequality in distribution, whereas in the second case inequality is the proportion of average income which must be deducted from it in the same intent. It may be said that in these two concepts inequality respectively cuts down or scales down average income to obtain a measure of social welfare. Then (7) and (8) say: If we think that one pound more to a person who earns 90 pounds per month increases social welfare as much as one pound and a half more to a person who earns 80 pounds per month, this opinion does not depend upon the incomes of other people; and this must hold for all pairs of incomes and all distributions. This property may be labeled "welfare independence." Well-known results in economics show that it is equivalent to saying that there exists a function of this social index which can be written as a sum of functions of each of the x_i's. In other words, social welfare is of course an ordinal concept, but we consider here that its ordinal index (i.e., defined up to an increasing transformation) has two interesting specifications: one is an absolute specification, measured in pounds sterling, which is $\bar{x} - I$ or $(1 - I)\bar{x}$, and the other one is a cardinal specification (defined up to a linear transformation) which is a sum of functions each of only one x_i.

II.b. *Result*

Now, the result[8] is:

(1°) (a) Properties (1), (3), (5), and (7) hold altogether if and only if I is of the form

$$I_t = (1/\alpha) \log \left[(1/n) \sum e^{\alpha\cdot(\bar{x}-x_i)}\right].$$

(b) These properties, plus (2) or (4) (which can thus replace (3)), hold altogether if and only if, furthermore, $\alpha > 0$ ($\alpha \geqslant 0$ if we choose the weak form of (2) or (4)).[9]

[8] It is a straightforward application of [7, theorems 13–17] (where the proofs are not reported) or of the equivalent results in the theory of choice under uncertainty (more specifically, of portfolio choice since the random variable is a unidimensional quantity) first presented (at least for (1°)) in [10, p. 128].

[9] Sections IX and XI in Part II will show that conditions (2) or (4) can also be replaced

(2°) (a) Properties (1), (3), (6), and (8) hold altogether if and only if I is of the form

$$I_r = 1 - \left[(1/n) \sum (x_i/\bar{x})^{1-\epsilon}\right]^{1/(1-\epsilon)}$$

or

$$I_r = 1 - \prod (x_i/\bar{x})^{1/n}.$$

(b) These properties, plus (2) or (4) (which can thus replace (3)), hold altogether if and only if, furthermore, $\epsilon > 0$ ($\epsilon \geqslant 0$ if we choose the weak form of (2) or (4)).[10]

II.c. *Proofs*

Here, as in the rest of this article, the reader who does not feel like going through a bit of mathematics can skip the proofs or the mathematical remarks he does not feel at home with. These are not however to be considered as an "appendix" (and eventually materially reported in one), because their development generally shows significant and interesting economic and logical properties.[11]

We first need a preliminary remark. We shall meet functions of the x_i's of the form $\varphi^{-1}[(1/n) \sum \varphi(x_i)]$ where φ is a function of one variable. This is the special case of the "generalized mean" of the x_i's with function φ where all the weights $1/n$ are equal. Now, we shall want to use theorems established for the general form $\varphi^{-1}[\sum q_i\varphi(x_i)]$ where the q_i's are any weights ($q_i \geqslant 0$ for all i's and $\sum q_i = 1$) and about properties valid for all q_i's. Such a property is also valid for the special case where all q_i's are equal and thus are $1/n$. But the important point is that the reverse is also true if the property for equal weights is valid for all n's and all admissible x_i's. This is so because one can consider m (let us say) i's with equal x_i's

in (1°)(b) and (2°)(b) by any of four other properties which are *convexity*, *quasi-convexity*, *constant-sum convexity*, *constant-sum quasi-convexity of I_1 or of $\bar{x}I_r$*, or *constant-sum convexity*, *constant-sum quasi-convexity of I_r*. Thus, with the properties of (1°)(a) or (2°)(a), all these properties are equivalent to each other and to $\epsilon > 0$ or $\alpha > 0$ ($\geqslant 0$ for the weak forms); in particular, the transfers principle or merely nonnegativity (or positivity out of equality) then implies the convexities mentioned.

[10] See Footnote 9.

[11] This warning to the less mathematically oriented readers must be supplemented by another one to the more mathematically oriented ones. We want to keep the mathematical apparatus to the lowest possible level, and to this end we do not mention explicitly some of the properties' conditions when this gap can be easily and straightforwardly filled in by the reader; this often happens in particular for the domains of variation of the variables.

and lump them in the equal weights case so as to form a term $(m/n)\ \varphi(x_i)$, and m/n can be made to approach any q_i as close as we want to, by increasing the number n if necessary. We note that this m/n is nothing but the $f(x_i)$ of Section 1.

We now prove the result (1°) (a). Let us call here $\bar{\bar{x}} = \bar{x} - I$. This $\bar{\bar{x}}$ is thus a function of x, as are $\bar{x}$ and I. Property (7) shows that it can be written as a function of a sum of functions of only one x_i each. And $\bar{\bar{x}}$ is a symmetrical function of the x_i's since $\bar{x}$ and I are (property (3)). These functions of each x_i must thus be the same functions. We can then write

$$\bar{\bar{x}} = F\left[(1/n) \sum \varphi(x_i)\right].$$

When all x_i's are equal, they are equal to $\bar{x}$, and so is $\bar{\bar{x}}$ from property (1). Then,

$$\bar{x} = F[\varphi(\bar{x})]$$

for all $\bar{x}$'s, which means that $F = \varphi^{-1}$, the inverse function of φ. $\bar{\bar{x}}$ is thus of the form

$$\bar{\bar{x}} = \varphi^{-1}\left[(1/n) \sum \varphi(x_i)\right].$$

From a well-known theorem (see for instance [18, Theorem 83]) and the above preliminary remark, two functions φ give the same value to $\bar{\bar{x}}$ for all x's in this expression if and only if they are in a linear relation. We thus choose to replace $\varphi(y)$ (where y is the current variable) by $\varphi(y) - \varphi(0)$, so that, with this new φ, $\varphi(0) = 0$. Since adding the same constant μ to all x_i's transforms $\bar{x}$ into $\bar{x} + \mu$, property (5) means that it transforms $\bar{\bar{x}}$ into $\bar{\bar{x}} + \mu$, that is,

$$\bar{\bar{x}} = \varphi^{-1}[(1/n)\ \varphi(x_i + \mu)] - \mu.$$

If we consider the new function defined by $\psi(y) \equiv \varphi(y + \mu)$, this expression is

$$\bar{\bar{x}} = \psi^{-1}\left[(1/n) \sum \psi(x_i)\right],$$

which shows that the functions φ and ψ are in a linear relation, which can be written as

$$\psi(y) = \varphi(y + \mu) = a(\mu) \cdot \varphi(y) + b(\mu),$$

where a and b, constant in y, are functions of μ, and $a(\mu) \neq 0$. Then, $y = 0$ gives $\varphi(\mu) = b(\mu)$, so that

$$\varphi(y + \mu) = a(\mu) \cdot \varphi(y) + \varphi(\mu).$$

Interchanging y and μ gives

$$\varphi(y + \mu) = a(y) \cdot \varphi(\mu) + \varphi(y),$$

and the equality of the right-hand sides can be written as

$$[a(y) - 1]/\varphi(y) = [a(\mu) - 1]/\varphi(\mu),$$

which shows that this expression is a constant c, so that $a(y) = c\varphi(y) + 1$. Then

$$\varphi(y + \mu) = c \cdot \varphi(y) \cdot \varphi(\mu) + \varphi(y) + \varphi(\mu).$$

If $c = 0$, this expression reduces to

$$\varphi(y + \mu) = \varphi(y) + \varphi(\mu),$$

which implies that $\varphi(y) = k \cdot y$ where k is a constant. If $c \neq 0$, writing $\chi(y) = c\varphi(y) + 1$ shows that

$$\chi(y + \mu) = \chi(y) \cdot \chi(\mu),$$

the general solution of which is $\chi(y) = e^{-\alpha y}$ where α is a constant, and the corresponding φ is $\varphi(y) = (e^{-\alpha y} - 1)/c$ which is equivalent to $e^{-\alpha y}$ to compute $\bar{\bar{x}}$:

$$\bar{\bar{x}} = -(1/\alpha) \log \left[(1/n) \sum e^{-\alpha x_i}\right],$$

and I is

$$I_l = \bar{x} - \bar{\bar{x}} = (1/\alpha) \log \left[(1/n) \sum e^{-\alpha \cdot (x_i - \bar{x})}\right],$$

where we see that the linear case is the limit of this expression when α tends to zero.

We now prove (1°) (b). $I_l = 0$ for all x's if and only if $\alpha = 0$. For $\alpha \neq 0$, $e^{-\alpha y}$ is a convex function of y. From Jensen's inequality [19] this is equivalent to $e^{-\alpha \bar{x}} < (1/n) \sum e^{-\alpha x_i}$ for all x_i's not all equal. Thus, $\bar{x} > \bar{\bar{x}}$ out of equality if and only if $\alpha > 0$. Besides, I is Schur-convex if and only if $\bar{\bar{x}} = \bar{x} - I$ is Schur-concave (the transfer does not change $\bar{x}$). And $e^{-\alpha \bar{\bar{x}}}(\partial \bar{\bar{x}}/\partial x_i) = (1/n) e^{-\alpha x_i}$ shows that the derivatives $\partial \bar{\bar{x}}/\partial x_i$ have their magnitude classified in order inverse to the x_i's, which is $\bar{\bar{x}}$'s strict Schur-concavity if and only if $\alpha > 0$.

Let us now prove (2°) (a). We now call $\bar{\bar{x}} = (1 - I)\bar{x}$. This $\bar{\bar{x}}$ thus is a function of x, as are $\bar{x}$ and I. Property (8) shows that it can be written as a function of a sum of functions of only one x_i each. And $\bar{\bar{x}}$ is a symmetrical function of the x_i's since $\bar{x}$ and I are (property (3)). These functions of each x_i must thus be the same functions. We can then write

$$\bar{\bar{x}} = F\left[(1/n) \sum \varphi(x_i)\right].$$

As previously shown, property (1) then implies that $F = \varphi^{-1}$ and thus

$$\bar{\bar{x}} = \varphi^{-1}\left[(1/n) \sum \varphi(x_i)\right].$$

Since multiplying all x_i's by the same scalar λ multiplies $\bar{x}$ by λ, property (6) and the present definition of $\bar{\bar{x}}$ shows that it also multiplies $\bar{\bar{x}}$ by λ. But this property, the structure of $\bar{x}$ as a generalized mean with equal weights, and the preliminary remark, give the result from a well-known theorem (see, for instance, [18, Theorem 84]): φ can be written as a power or logarithmic function $\varphi(y) = y^{1-\epsilon}$ or $\varphi(y) = \log y$. Then,

$$\bar{\bar{x}} = \left[(1/n) \sum x_i^{1-\epsilon}\right]^{1/(1-\epsilon)}$$

or

$$\bar{\bar{x}} = \left[\prod x_i\right]^{1/n},$$

and I is

$$I_r = 1 - \left[(1/n) \sum (x_i/\bar{x})^{1-\epsilon}\right]^{1/(1-\epsilon)}$$

or

$$I_r = 1 - \left[\prod x_i/\bar{x}\right]^{1/n},$$

which is the limiting case of the previous one when ϵ tends to one.

We finally Prove (2°) (b). $\bar{x} = \bar{\bar{x}}$ and $I_r = 0$ for all x_i's if and only if $\epsilon = 0$. For $\epsilon = 1$, $\bar{\bar{x}} < \bar{x}$ if x_i's are not all equal from the well-known relation between arithmetic and geometric means. Using Jensen's inequality for x_i's not all equal, we see that: If $\epsilon < 0$, the function $y^{1-\epsilon}$ is strictly convex increasing, $\bar{\bar{x}}^{1-\epsilon} > \bar{x}^{1-\epsilon}$, $\bar{\bar{x}} > \bar{x}$; if $0 < \epsilon < 1$, $y^{1-\epsilon}$ is strictly concave increasing, $\bar{\bar{x}}^{1-\epsilon} < \bar{x}^{1-\epsilon}$, $\bar{\bar{x}} < \bar{x}$; if $\epsilon > 1$, $y^{1-\epsilon}$ is strictly convex decreasing, $\bar{\bar{x}}^{1-\epsilon} > \bar{x}^{1-\epsilon}$, $\bar{\bar{x}} < \bar{x}$. Thus, $I_r > 0$ out of equality if and only if $\epsilon > 0$. Besides, I is Schur-convex if and only if $\bar{\bar{x}} = (1 - I)\bar{x}$

is Schur-concave (the transfer does not change $\bar{x}$). And $\bar{\bar{x}}^{-\epsilon}(\partial\bar{\bar{x}}/\partial x_i) = (1/n)\, x_i^{-\epsilon}$ shows that the derivatives $\partial\bar{\bar{x}}/\partial x_i$ have their magnitudes classified in order inverse to the x_i's, which is $\bar{\bar{x}}$'s strict Schur-concavity, if and only if $\epsilon > 0$.

II.d. *Relations between the Basic Properties of Inequality Measures*

The basic properties of inequality measures mentioned are not independent of each other. Although the most general properties of inequality measures will be discussed in more detail in Part II, let us mention these relations here and give a sketch of the proof of those which will not be proved under their present form. The following relations hold between these properties, with respective correspondence between weak and strict forms of the properties when this is relevant (when this distinction is meaningful for the hypotheses only, the implication is true for either form), and with the eventually required assumptions about differentiability and domain of variation.

(i) Zero at equality (1) and impartiality (3) and rectifiance (4) imply nonnegativity or positivity out of equality (2).

(ii) Independence ((7) or (8)) and zero at equality (1) and nonnegativity or positivity out of equality (2) imply impartiality (3) and rectifiance (4).

(iii) Independence ((7) or (8)) and rectifiance (4) imply impartiality (3).

(iv) Independence ((7) or (8)) and rectifiance (4) and zero at equality (1) imply nonnegativity or positivity out of equality (2).

(v) Given independence ((7) or (8)) and zero at equality (1), nonnegativity or positivity out of equality (2) and rectifiance (4) are equivalent and they imply impartiality (3).

(v) is a synthesis of (ii) and (iv). Relation (iii) shows that, with independence, rectifiance suffices to define Schur-convexity (rectifiance plus impartiality). Relation (i) will be proved in Part II, Section XIa.

We remark that properties (1), (2), (3), and (4) can be expressed equivalently on an inequality measure $I(x)$ or on $\bar{x}$. $I(x)$ or on $I(x)/\bar{x}$ (with $\bar{x} > 0$). Then, independence means that one of these functions can be written as $I(x) = x - \Phi[\sum \varphi^i(x_i)]$. Independence plus impartiality means that it can be written as $I(x) = \bar{x} - \Phi[\sum \varphi(x_i)]$. Independence plus impartiality plus zero at equality means that it can be written as $I(x) = \bar{x} - \varphi^{-1}[(1/n) \sum \varphi(x_i)]$.

This latter form is thus implied both by the hypotheses of (ii) plus impartiality (3), and by the hypotheses of (iv) if (iii) is true. Rectifiance then means that φ is concave if it is increasing and convex if it is decreasing, and Jensen's inequality shows that these structures are equivalent to nonnegativity (weak forms) or positivity out of equality (strict forms) of I.

It remains to show that independence plus either (1) and (2), or (4), implies impartiality (3). We consider two neighboring distributions of the same $\bar{x} = \xi$. In one, all x_k are equal to ξ. The second one differs from it only by $x_i = \xi + \epsilon$ and $x_j = \xi - \epsilon$ with $\epsilon > 0$ and tending to zero. We choose a Φ increasing at the first point. To pass from the second situation to the first one is both to pass from inequality to equality and to make a transfer from a richer to a poorer person. Thus, if either (1) and (2), or (4), holds, it must not decrease $\varphi^i(x_i) + \varphi^j(x_j)$:

$$\varphi^i(\xi + \epsilon) + \varphi^j(\xi - \epsilon) \leqslant \varphi^i(\xi) + \varphi^j(\xi).$$

If φ^i and φ^j are differentiable at ξ this implies $\varphi^{i\prime}(\xi) - \varphi^{j\prime}(\xi) \leqslant 0$. Reversing the roles of i and j similarly implies $\varphi^{j\prime}(\xi) - \varphi^{i\prime}(\xi) \leqslant 0$. Therefore, $\varphi^{i\prime}(\xi) = \varphi^{j\prime}(\xi)$. Integrating and letting j go from 1 to n shows that $\varphi^k(y) = \varphi(y) + c_k$ for $k = 1,..., n$. And changing the function Φ into $\Psi(z) = \Phi(z + \sum c_i)$ shows I's impartiality.

*

We thus have necessary and sufficient conditions for an inequality measure to be I_r or I_l. But there exist both intermediate measures, between these two, and other measures which synthesize most of their important properties.

III. Centrist Measures of Inequality

III.a. *The Most Specific Common Generalization*

Many people feel that an equal augmentation in all incomes decreases inequality, whereas an equiproportional increase in all incomes increases it. Dalton, for one, was of this opinion [2, 3]. Neither I_l nor I_r suits these "centrists" as a measure of inequality. They will thus feel comforted by the existence of a class of measures which has the property they require, is the closest extension of both I_r and I_l, and contains them as special, limit, cases. By "closest extension" of I_r and I_l we mean that all the properties—suitably defined—of these two measures are retained, except the criticized invariances for equal or equiproportional variations in all incomes. In particular, we retain both the general properties of inequality measures—value zero for equal distributions (1), positivity otherwise (2), impartiality-symmetry (3)—and the "welfare independence" property which is more specific to I_r and I_l; they happen to also possess the Schur-convexity (and thus "transfers principle") and even convexity properties, as will be shown further. Both I_r and I_l are one-parameter families of functions (the parameters being ϵ and α). A generalization of both must therefore be a several-parameter family. And the form of inequality measure which is a necessary and sufficient condition for all the properties mentioned to hold will turn out to be a family of functions with only two parameters.

To see at best how these measures are intermediaries between I_r and I_l, it is enlightening to cast a geometrical glance in the n-dimensional vector space of the income distributions x. To multiply the vector x by the scalar λ is a "blowing-up" of x from the origin (if $\lambda > 1$; it is a "blowing-down" if $\lambda < 1$); this leaves I_r unchanged, and it thus multiplies $\bar{x}I_r$ by λ. To add μ to each x_i, i.e., the vector μe to x, is a shift in the direction of the vector e; this leaves I_l unchanged. This shift can be considered as a blowing-up from a point which is at infinity in the direction of vector e (or rather $-e$ if we consider oriented directions and if we want an addition, i.e., $\mu > 0$, to be the limit case of a blowing-up, i.e., $\lambda > 1$, and a subtraction, i.e., $\mu < 0$, to be the limit case of a blowing-down, i.e., $\lambda < 1$). These two transformations are both special cases of a blowing-up from a point in the x vector space. More precisely, they are special cases of the case when this point is on the "generalized bisector" which is the straight line Δ passing through the origin and bearing vector e. But if we consider a property only based upon a blowing-up operation, this latter, specific position will certainly be imposed by the symmetry property. If we call X the point, or n dimensional vector, which is the center of this blowing-up, this operation transforms vector $x - X$ into vector $\lambda \cdot (x - X)$, and therefore vector x into vector $\lambda \cdot (x - X) + X$. The property will say that this transforms the inequality measure I into λI. More precisely, the property is

$$(9) \quad I[\lambda \cdot (x - X) + X] = \lambda \cdot I(x)$$

for some X and all admissible x's and λ's.

Let us first check that this includes the two properties studied above as special cases. When $X = 0$ (the origin), (9) is $I(\lambda x) = \lambda \cdot I(x)$, which is a property that the measure $I = \bar{x}I_r$ has. When X goes to infinity in the direction of vector $-e$, let us write $X = -\xi e$ where ξ is a scalar which tends to infinity; then $\lambda \cdot (x - X) + X = \lambda x + (\lambda - 1)\,\xi e$; given any number μ, we choose a $\lambda = 1 + \mu/\xi$; then $(\lambda - 1)\,\xi = \mu$, and when ξ tends to infinity λ tends to 1; at the limit, $\lambda \cdot (x - X) + X = x + \mu e$, and (9) becomes $I(x + \mu e) = I(x)$, i.e., property (5).

The result then is[12]:

[12] A similar structure was introduced in the analysis of choice under uncertainty for portfolio selection theory in [9, p. 129].

(a) Properties (1), (3), (7), and (9) hold altogether if and only if I is of the form

$$I_c = \bar{x} + \xi - \left[(1/n) \sum (x_i + \xi)^{1-\epsilon}\right]^{1/(1-\epsilon)}$$

or

$$I_c = \bar{x} + \xi - \prod (x_i + \xi)^{1/n},$$

where ϵ and ξ are numbers.

(b) These properties plus (2) or (4) (which can thus replace (3)) hold altogether if and only if, furthermore, $\epsilon > 0$ ($\epsilon \geqslant 0$ if we choose the weak form of (2) or (4)).[13]

III.b. *Proof and Differential Characterization*

This result is deduced from the result for I_r (Sect. II.b, 2°) by a mere change in variables from x to $x - X$ following a redefinition of I into $\bar{x}I$. More specifically, this last change transforms property (8) into property (7) and property (6) into $I(\lambda x) = \lambda \cdot I(x)$ for all admissible λ's and x's. A change of variables from x into $x - X$ transforms the latter relation into $I[\lambda \cdot (x - X)] = \lambda \cdot I(x - X)$. A new change in the definition of I from $I(x)$ into $I(x + X)$ transforms it into property (9). Equations (9) and (3) imply that X is on Δ if I is not a function of $\bar{x}$ only (since the family of x-space manifolds defined by $I(x) =$ constant then has as centers of homotheticity all points derived from X by permutations of its coordinates, and it can have only one). Besides, when x is on Δ (i.e., all x_i's are equal), $I(x) = 0$ from (1); (9) then gives $I[\lambda x + (1 - \lambda)X] = 0$, which (1) and (2) (in strict form) show to hold if and only if $\lambda x + (1 - \lambda)X$ is on Δ, which is equivalent to saying that X is on Δ, i.e., that all its coordinates are equal. Call, then, $-\xi$ the coordinates of X. Then, $x - X$ is on Δ if and only if x is on Δ, permutations of the x_i's and of the coordinates $x_i + \xi$ of $x - X$ are equivalent, the x_i's and the $x_i + \xi$'s are classified in the same order, so that properties (1), (2), (3), and (4) are conserved in the transformations of functions and variables mentioned. The results found for I_r thus imply the ones mentioned for I_c.

The second form of I_c is of course the limit of the first one when ϵ tends to zero. Besides, these forms are defined only for $x_i + \xi \geqslant 0$ for

[13] The remarks which were presented for I_r and I_t should be repeated here. With (7) and (1), (2) and (4) imply each other and imply (3). In result (b), (2) or (4) could be replaced by either convexity, or quasi-convexity, or constant-sum convexity, or constant-sum quasi-convexity, of I. Properties of (a) thus make these properties equivalent to each other, and equivalent to $\epsilon > 0$ ($\epsilon \geqslant 0$ for the weak forms); in particular, the transfers principle or merely nonnegativity (or positivity out of equality) then implies these convexities.

all i's. (We shall remark below that $\epsilon \geqslant 1$ and $x_i + \xi = 0$ for one i at least imply $I_c = \bar{x} + \xi$.) If $\xi \geqslant 0$, this is implied by $x_i \geqslant 0$ for all i's. But if $\xi < 0$, this condition restricts the domain of variation of x to $x_i \geqslant -\xi$ for all i's. However, we shall see in Section V below that this latter case implies that if all x_i's are increased in the same proportion, $I_c/\bar{x}$ decreases; such a measure would thus be in this sense "more rightist" than I_r rather than centrist.

Evidently, if we redefine the index i so as to represent an income class, and if $f(x_i)$ is the proportion of persons who have income x_i, I_c is rewritten as

$$I_c = \bar{x} + \xi - \left[\sum (x_i + \xi)^{1-\epsilon} f(x_i)\right]^{1/(1-\epsilon)}$$

or

$$I_c = \bar{x} + \xi - \prod (x_i + \xi)^{f(x_i)}.$$

I_c is a two-parameter measure (ϵ and ξ). When $\epsilon = 0$, $I_c = 0$ whatever x (and ξ). When ξ tends to infinity, whereas ϵ remains finite, I_c tends to 0, whatever x (and ϵ). When ϵ tends to infinity whereas ξ remains finite, I_c tends to $\bar{x} - \underline{x}$, i.e., the difference between average and minimum income.

We may now check that $\bar{x}I_r$ and I_l are special cases of I_c. I_c is obviously $\bar{x}I_r$ when $\xi = 0$. And we shall show that I_c tends to I_l when both ϵ and ξ tend to infinity whereas their ratio ϵ/ξ tends to a finite value which is α. In fact, this limit can be found by the following casual remark.

We first notice that $\bar{x}I_r$, I_l, and I_c are all three of the form $\bar{x} - \bar{\bar{x}}$ where $\bar{\bar{x}}$ has the form $\bar{\bar{x}} = \varphi^{-1}[(1/n) \sum \varphi(x_i)]$ where φ is a function and φ^{-1} is its inverse function: Writing y for the current variable, $\varphi(y)$ is respectively $y^{1-\epsilon}$ or $\log y$ for $\bar{x}I_r$, $e^{-\alpha y}$ for I_l, and $(y + \xi)^{1-\epsilon}$ or $\log(y + \xi)$ for I_c. A remark of Section (II.c) above shows that the class of functions φ which give the same $\bar{\bar{x}}$ for all x's is of the form $a\varphi + b$ if φ is one of them and a and b are constants. This constitutes a two-parameter (a and b) family of functions φ, and the φ considered are twice differentiable. This family is thus characterized by its second-order differential equation which contains neither a nor b (they will be determined by the integration variables). Since φ'/φ'' does not depend upon the transformation of φ into $a\varphi + b$, this equation can be written as $\varphi'/\varphi'' = h(y)$ where h is some function, unless it is $\varphi'' = 0$, i.e., $\varphi = a'y + b'$ (a' and b' are constants) which is the φ for I_r or I_c with $\epsilon = 0$ or for I_l with α very close to zero. We find $h = -(1/\epsilon)\, y$ for $\varphi = y^{1-\epsilon}$ and $h = -y$ for $\varphi = \log y$, $h = -(1/\alpha)$ for $\varphi = e^{-\alpha y}$, $h = -(1/\epsilon)(y + \xi)$ for $\varphi = (y + \xi)^{1-\epsilon}$ and $h = -(y + \xi)$ for $\varphi = \log(y + \xi)$. Clearly, $-(1/\epsilon)(y + \xi)$ becomes $-(1/\epsilon)\, y$ when $\xi = 0$, and it becomes $-1/\alpha$ when both ϵ and ξ tend to infinity with ϵ/ξ tending to α.

This remark has shown us in passing that "welfare independence" and $I = 0$ at equality, which were seen to imply the forms $\bar{x} - \bar{\bar{x}}$ and the "mean" form for $\bar{\bar{x}}$, and φ'/φ'' of general linear form, characterize the measures I_c, with the special cases of φ'/φ'' homogeneous linear for I_r and constant for I_l.[14 15]

IV. Inequality Per Pound or Inequality Per Person?

The above remarks show that I_r and I_l can be said to differ for two reasons: They are not based upon the same function φ, and they are not derived in the same way from it since $I_l = \bar{x} - \bar{\bar{x}}$ and $I_r = (\bar{x} - \bar{\bar{x}})/\bar{x} = 1 - (\bar{\bar{x}}/\bar{x})$.

If $\sum \varphi(x_i)$ were a "social evaluation ("welfare") function," $\bar{\bar{x}}$ defined by $n\varphi(\bar{\bar{x}}) = \sum \varphi(x_i)$ would be the "equal equivalent" income([7, Sect. VI]). Generally, if $V(x)$ is an ordinal (i.e., $V(x)$ can be replaced by any $F[V(x)]$ where F is any increasing function) "social evaluation function," the "equal equivalent" income is $\bar{\bar{x}}$ defined by $V(\bar{\bar{x}}, \bar{\bar{x}},..., \bar{\bar{x}}) = V(x_1, x_2,..., x_n)$ [7, Sect. VI]. It is uniquely defined if, for all x's, $\partial V/\partial x_i \geqslant 0$ ("nonmalevolence") for all i's and $\partial V/\partial x_i > 0$ ("benevolence") for at least one i. In other words, the "equal equivalent" income $\bar{\bar{x}}$ is the income level such that, if all persons had the same income at this level, society's welfare would be considered as as good (or bad) as it effectively is. $\bar{\bar{x}}$ is a function of x and a functional (function of function) of $V(\ )$. It is a special specification of the ordinal index V.[16]

[14] The consideration of general linear φ'/φ'' and its integration were first introduced into the theory of choice under uncertainty and in portfolio theory in [10, p. 129] as the generalization of the cases of proportional and constant φ'/φ''. This structure of φ'/φ'' was then used by H. Leland in his work in dynamic portfolio analysis and then by Borch, Mossin and Hagen in their analysis of financial market efficiency. The decomposability of portfolio choice into choice between a riskless asset and a risky portfolio and choice of the latter's composition, which is associated to this structure of the utility function, belongs more generally to utility functions in the contingent incomes (the x_i's) having the structure which will be called $-\xi e$-homogenity or e-translatedness in Part II, Sect. XI.

[15] On practical grounds, if we choose to use a centrist measure of inequality to compare income distributions, apart from the question of the choice of ϵ, which arises also for I_r (as that of α for I_l), that of the choice of ξ is raised. If we compare distributions with the same average $\bar{x}$, the choice $\xi = \bar{x}$ seems reasonable. If not, a ξ which is the average of averages weighted by populations (i.e., the average income for a population which is the gathering of the compared ones) may also be suggested as suitable.

[16] As will be discussed below, "impartiality" (symmetry of V and the "transfers principle" are altogether characteristic of Schur-concavity of a differentiable V, and they imply $\bar{\bar{x}} \leqslant \bar{x}$ ($<$ out of equality for strict form) (cf. Part II, Sect. XI). $\bar{\bar{x}}$ can then be defined as the smallest average income which allows the same welfare level V as x.

More precisely, $\bar{\bar{x}}$ is in some sense average welfare per person measured in pounds sterling. $\bar{x}$ is average income per person. $\bar{\bar{x}}$ differs from $\bar{x}$ because all x_i's are not equal. Therefore, $\bar{x} - \bar{\bar{x}}$ can be taken as a "per person" pound measure of inequality. We could also use the "per pound" pound measure of inequality $(\bar{x} - \bar{\bar{x}})/\bar{x}$ (we could also have divided by $\bar{\bar{x}}$ rather then $\bar{x}$, replacing a per "income pound" measure by a per "equal welfare pound" measure). Total pound measure of inequality $n \cdot (\bar{x} - \bar{\bar{x}})$ may also be significant. We shall call $I^a = \bar{x} - \bar{\bar{x}}$ and $I^r = (\bar{x} - \bar{\bar{x}})/\bar{x} = 1 - (\bar{\bar{x}}/\bar{x})$ (a and r stand for "absolute" and "relative," but if I^r, expressed in pounds per pound, is, as implied, a pure dimensionless number, I^a is absolute only in pounds since it is otherwise expressed per person).

I_r is a per pound I^r, and I_l and I_c are per person I^a's. We can thus also consider a per person rightist measure $I_r{}^a = \bar{x} I_r$ and per pound leftist and centrist measures $I_l{}^r = I_l/\bar{x}$ and $I_c{}^r = I_c{}^r/\bar{x}$.

All measures which are not "per pound" raise the problem of measuring real, comparable, "pounds," which was mentioned earlier about I_l. The solutions are still the same, and also the same as the traditional ones for any comparison of incomes.

We must, in addition, present a remark about the domain of variation of the x_i's. The functions used may imply some restrictions on it. For I_l, no such restriction exists. But we of course define I_r only for $x_i \geqslant 0$, and $I_c{}^a$ and $I_c{}^r$ only for $x_i \geqslant -\xi$, for all i's. When x_i's are income or wealth, there exists negative x_i's: business failures make negative incomes, and a net indebtedness is a negative wealth. I_r would thus be unacceptable to measure the inequality in such magnitudes. I_l, on the other hand, can take all these cases into account, as can I_c if we choose ξ large enough. It may however be that one might neglect incomes or wealths which are negative or smaller that $-\xi$ ($\xi > 0$). But this certainly cannot be done for other interesting uses of inequality measures. (For income or wealth, however, $\bar{x} > 0$ is certainly the relevant case, and we thus shall assume it.)

V. Variations of Inequality with Its Parameters and with Equal Proportional and Absolute Variations of Incomes

V.a. *General Properties*

We now have six inequality measures: the rightist, leftist, and centrist ones, and in each case a per person ("absolute") and per pound ("relative") measure. These indices depend upon three parameters, ϵ, α and ξ, and, of course, upon the income distribution. It is interesting to know how they vary with these parameters, and with the two kinds of variations in incomes which we have considered: equal proportional or absolute increase (or decrease) in all incomes. These two questions turn out to be closely related, and they must therefore be studied together.

We shall deal separately, at the end of this section, with three kinds of special cases: (1°) $\xi < 0$, which will turn out to give a measure which not really "centrist," (2°) $\epsilon \geqslant 1$ and one $x_i + \xi$ is zero (including one x_i zero for $\xi = 0$), (3°) the measures take a zero value. The following results hold for the other cases.

The inequality measures vary as ϵ or α and inversely with ξ. An equal absolute increase (resp., decrease) in all incomes diminishes (resp., increases) all measures except I_l, which remains unchanged. An equiproportional increase (resp., decrease) in all incomes increases (resp., diminishes) all measures except I_r which remains unchanged, and I_r^a's variations are proportional, and I_l and I_c's more than proportional, to the incomes'.

We now prove these propositions. Each inequality index will be written as a function of its parameters and of x. λ and μ are numbers ($\lambda > 0$). When x is transformed into λx or $x + \mu e$, $\bar{x}$ is transformed into $\lambda \bar{x}$ or $\bar{x} + \mu$.

We first observe that:

$$\begin{aligned} I_r(\epsilon, \lambda x) &= I_r(\epsilon, x), \\ I_r^a(\epsilon, \lambda x) &= \lambda \cdot I_r^a(\epsilon, x), \\ I_l^r(\alpha, \lambda x) &= I_l^r(\alpha\lambda, x), \\ I_l(\alpha, \lambda x) &= \lambda \cdot I_l(\alpha\lambda, x), \\ I_c(\epsilon, \xi, x + \mu e) &= I_c(\epsilon, \xi + \mu, x), \\ I_l(\alpha, x + \mu e) &= I_l(\alpha, x). \end{aligned}$$

V.b. *Demonstrations*

This last property shows that I_l^r varies inversely with μ (i.e., in the opposite direction).

How does I_l vary with α? $I_l = \bar{x} - \bar{\bar{x}}$, and $\bar{\bar{x}}$ only depends upon α. But $\bar{\bar{x}} = -(1/\alpha) \log((1/n) \sum e^{-\alpha x_i})$ is of the form $\varphi^{-1}[(1/n) \sum \varphi(x_i)]$ with $\varphi(y) = e^{-\alpha y}$, i.e., it is a "generalized mean" with this function. As α' is another α, we call $\psi(y) = e^{-\alpha' y}$. The function $\psi\varphi^{-1} = y^{\alpha'/\alpha}$ is convex or concave according as $\alpha' \gtrless \alpha$. And φ and ψ are strictly monotonic and ψ is decreasing (since $\alpha' > 0$). Hence, applying the demonstration of [18, Theorem 83] shows that $\bar{\bar{x}}$ is smaller, α the larger. I_l thus varies in the same direction as α.

$I_l^r = I_l/\bar{x}$ also varies as α. And since to multiply α or x by λ gives the same I_l^r, I_l^r varies as λ. I_l thus also varies as λ, and more than proportionally.

As a function of ϵ, I_c varies in the direction opposite to that of

$((1/n) \sum \eta_i^{1-\epsilon})^{1/(1-\epsilon)}$ where $\eta_i = x_i + \xi$. A well-known result (for instance, [18, Theorem 16]) says that this "mean" of the η_i's varies as $1 - \epsilon$, i.e., in the direction opposite to that of ϵ. I_c then varies as ϵ. And so do $I_c^r = I_c/\bar{x}$ and their special cases for $\xi = 0$, I_r^a and I_r.

How does I_c vary with its second parameter, ξ? We have

$$dI_c/d\xi = 1 - \left((1/n) \sum \eta_i^{1-\epsilon}\right)^{\epsilon/(1-\epsilon)} (1/n) \sum \eta_i^{-\epsilon}.$$

This expression has the sign of

$$\left((1/n) \sum \eta_i^{-\epsilon}\right)^{-1} - \left((1/n) \sum \eta_i^{1-\epsilon}\right)^{\epsilon/(1-\epsilon)},$$

which has the sign of

$$\left((1/n) \sum \eta_i^{-\epsilon}\right)^{-1/\epsilon} - \left((1/n) \sum \eta_i^{1-\epsilon}\right)^{1/(1-\epsilon)}.$$

These two terms are "means" which differ only by their exponents $-\epsilon$ and $1 - \epsilon$. From the above-mentioned result, they compare as these exponents. Therefore, $dI_c/d\xi < 0$.

$I_c^r = I_c/\bar{x}$ also varies inversely with ξ.

A variation in ξ, and an equal variation in all x_i's, have the same effect on I_c. I_c thus varies inversely with an equal variation in all x_i's. So a fortiori does $I_c^r = I_c/\bar{x}$, and so do their special cases for $\xi = 0$, I_r^a and I_r.

To see the effect on I_c^r of an equiproportional variation in all x_i's, we replace a multiplication of all x_i's by λ (> 0) by a multiplication of all x_i's and ξ by λ followed by the subtraction of $(\lambda - 1)\,\xi$ from ξ. The first operation leaves I_c^r unchanged. The second one increases or decreases it according as $(\lambda - 1)\,\xi \gtrless 0$. Thus, for $\xi > 0$, I_c^r varies as an equiproportional variation in all x_i's, and for $\xi < 0$ it varies inversely with it. For $\xi > 0$, $I_c = \bar{x} I_c^r$ a fortiori varies as λ, and more than proportionally.

The variation of I_c with an equiproportional variation in all x_i's when $\xi < 0$ remains the only effect the direction of which is not a priori determined. In fact, the two extreme cases would be $\xi = 0$ where $I_c = I_r^a$ varies proportionally to λ, and $\xi \to -\infty$, which transforms an equiproportional increase (resp., decrease) in all incomes into an equal absolute increase (resp., decrease) in all incomes, which we have seen to decrease (resp., increase) I_c. Thus, for $\xi < 0$, an equiproportional variation in all x_i's sometimes increases, and sometimes decreases, I_c. It all depends upon where x, ξ and ϵ stand. However, the case $\xi < 0$ is not the most interesting, since, as we have just seen, the inequality measure I_c^r then decreases when all x_i's are increased in the same proportion, and this happens only in this case. It thus hardly deserves its adjective of "centrist" and is rather "far right."

Finally, we notice that the monotonicity of the variations for ϵ and α proves that all per person ("absolute") measures are between 0 and $\bar{x} - \underline{x}$, and all per pound ("relative") measures are between 0 and $1 - (\underline{x}/\bar{x})$.

V.c. *Special Cases*

The special cases when the inequality measure is zero could come either from $\epsilon = 0$ for I_r, I_r^a, I_c, or I_c^r, or from $\alpha = 0$ for I_l or I_l^r, or from the equality of all x_i's for all these measures. In all these cases, an equal absolute or relative variation in all incomes does not change the value of the measure, which remains zero: If $\epsilon = 0$ or $\alpha = 0$ the index is always zero, and if all the x_i's are equal the variation keeps them equal and the measure remains zero. Furthermore, if $\epsilon = 0$ in I_c or I_c^r, a variation in the parameter ξ does not change the measure, which remains zero.

A last category of special cases remains to be considered.

$\bar{x} - I_l = \bar{\bar{x}}$ is zero for $\alpha > 0$ if and only if $x_i = 0$ for all i's. Similarly, $\bar{x} - I_r^a = \bar{\bar{x}}$ is zero for $\epsilon < 1$ if and only if $x_i = 0$ for all i's. But for $\epsilon \geqslant 1$, $\bar{x} - I_r^a = \bar{\bar{x}}$ is zero if and only if one x_i (at least) is zero. This is ethically meaningful: When ϵ is high enough, one zero income suffices to bring the equal equivalent income down to zero and thus the inequality measure I_r up to 1 (if not all x_i's are zero), which is its maximum value ($\underline{x} = 0$). Then, for $\epsilon \geqslant 1$ and one x_i zero, I_r^a and I_r do not depend upon ϵ (but their variations for equiproportional variations, or equal increases, in all x_i's are as the general case).

More generally, for $\epsilon \geqslant 1$ and $x_i + \xi = 0$ for one i (at least), $[(1/n) \sum (x_i + \xi)^{1-\epsilon}]^{1/(1-\epsilon)} = 0$ and $I_c = \bar{x} + \xi$. Note that $x_i + \xi = 0$ cannot happen for $\xi > 0$, which is the most interesting case of $\xi \neq 0$. We thus consider $\xi < 0$. This I_c does not depend upon ϵ, and neither does $I_c^r = I_c/\bar{x}$. It satisfies $dI_c/d\xi = 1 > 0$, and I_c^r is also an increasing function of ξ. It is increased or decreased by amount μ when all the x_i's are, and $I_c^r = 1 + (\xi/\bar{x})$ also varies in this direction. It varies as an equiproportional variation in all x_i's, more than proportionnally, and I_c^r also varies in this direction.

References

1. A. Pigou, "Wealth and Welfare," Macmillan, London, 1912.
2. H. Dalton, The measurement of the inequality of incomes, *Econ. J.* **30** (1920), 348–361.
3. H. Dalton, "Inequality of Incomes," London, 1925.
4. Taussig, "Principle of Economics," 4th ed., Macmillan, New York, 1939.

5. CANNAN, "Elementary Political Economy," London, 1930.
6. LORIA, "La sintisi economica" (Politica e economica), Nuova collana di economisti stranieri e italiani, No. 12, Torino, 1934.
7. S.-CH. KOLM, The optimal production of social justice, paper presented at the 1966 conference of the International Economic Association on Public Economics in Biarritz. Published in the proceedings of this conference "Économie Publique," CNRS, Paris, 1968, and "Public Economics," Macmillan).
8. A. B. ATKINSON, On the measurement of inequality, *J. Econ. Theory* **2**, No. 2, 1970.
9. S.-Ch. KOLM, "Les choix financiers et monétaires (théorie et technique modernes)," Dunod, Paris, 1966.
10. S.-CH. KOLM, "More Equal Distributions of Bundles of Commodities," mimeographed, CEPREMAP, Paris, 1973.
11. V. JAKOBSSON AND G. NORMANN, "Redistributive Effects of Automatic and Discretionary Tax Policies," Industrial Institute for Social and Economic Research, Stockholm, Sweden.
12. M. BRUNO AND J. HABIB, "Alternative Tax-Transfer Systems and Income Distribution," mimeographed, The Maurice Falk Institute for Economic Research in Israel, Jerusalem, July 1973.
13. M. BRUNO, Economic consequences of education policies. Paper presented at the 1974 I.E.A. Conference on the Economics of Public Services, Torino.
14. A. OSTROWSKI, Sur quelques applications des fonctions convexes et concaves au sens de I. Schur. *J. Math. Pures Appl.* **31** (1952).
15. I. SCHUR, Ein Beitrag zur Hilbertschen Theorie der vollstetigen quadratischen Formen, *Math. Z.* **12** (1922), 287–97.
16. I. SCHUR, Ueber eine Klasse von Mittelbildungen mit Anwendungen auf die Determinantentheorie, *Sitzb. Ber. Math. Ges.* (1923), 9–20.
17. I. SCHUR, Ueber einige Ungleichungen im Matrizenkalkül, *Prace Mat.* **44** (1936), 353–370.
18. G. H. HARDY, J. E. LITTLEWOOD, AND G. PÓLYA, "Inequalities," 2nd ed., Cambridge Univ. Press, Cambridge, 1964.
19. J. L. W. V. JENSEN, Sur les fonctions convexes et les inégalités entre les valeurs moyennes, *Acta Math.* **30** (1906), 175–193.