JOURNAL OF ECONOMIC THEORY 13, 82–111 (1976)

Unequal Inequalities. II

SERGE-CHRISTOPHE KOLM

School for Higher Studies in the Social Sciences, and CEPREMAP, Paris, France

Received March 7, 1975; revised December 2, 1975

SUMMARIES

This paper analyzes properties of measures of inequality, applied to income inequalities but meaningful for practically any measure of dispersion in economics. We call n the number of persons, i the person's index, x_i person i's income, $\bar{x} = \sum (x_i/n)$ the average income, x the vector of the x_i's or income distribution, $I(x)$ a real-valued function of x which is the measure (or index) of inequality.

Part I (Sects. I–V), which appeared in the last issue of this journal, analyzed several structures or properties, and specific forms, of I. We distinguished several I's: the measures of inequality per person (or "absolute") I^a, per pound (or "relative") $I^r = I^a/\bar{x}$, and total nI^a. We presented several possible properties of inequality measures, such as: $I = 0$ if all x_i's are equal ("zero at equality"), $I > 0$ otherwise ("positivity out of equality"), symmetry of I for x ("impartiality"), $((\partial I/\partial x_i) - (\partial I/\partial x_j))(x_i - x_j) > 0$ for $x_i \neq x_j$ ("rectifiance" of the function I, or "transfers principle," this being the strict form whereas the weak one is with sign $\geqslant$), the fact that

$$\frac{(\partial(\bar{x} - I^a)/\partial x_i)}{(\partial(\bar{x} - I^a)/\partial x_j)}$$

does not depend upon x_k for $k \neq i, j$ ("welfare independence," or, for short, "independence"). Rectifiance plus symmetry is Schur-convexity. Independence plus symmetry plus zero at equality implies that $\bar{\bar{x}} = \bar{x} - I^a = \varphi^{-1}[(1/n) \sum \varphi(x_i)]$ where $\bar{\bar{x}}$ is the "equal equivalent income"; and we will show that, these three properties being satisfied, the following ones are equivalent to each other: positivity out of equality, rectifiance, quasi-convexity, φ's concavity.

Part I largely focused on the study of six related specific measures of inequality, which in particular possess all the above properties: ϵ, α, and ξ being positive parameters, they are

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

or

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

$$I_c^r = I_c^a/\bar{x}, \quad I_r = I_c^r \quad \text{for} \quad \xi = 0, \qquad I_r^a = \bar{x}I_r = I_c^a \quad \text{for} \quad \xi = 0,$$

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

and $I_l^r = I_l/\bar{x}$. Lower indices c, r, l respectively stand for "centrist," "rightist," and "leftist" measures of inequality. I_r and I_l are invariant under respectively equiproportional variation in, or equal addition to, all incomes; measures which have the first of these two properties are said to be "intensive."

We now consider different and more general measures, and other properties. We first reconcile the last two properties by dropping the "independence" one (Section VI). Then, we analyze another mildly equalitarian property, the "principle of diminishing transfers" (Section VII). Section VIII turns to the relations between inequality measures and Lorenz and concentration curves. We then consider the effect on inequality of additions of incomes, and we analyze the properties of "diminishing equality" (Section IX). The effect of unions of populations is the topic of Section X. Finally, the last section (XI) presents the more general relations between the various structural properties of inequality measures.[1]

VI. Synthetic Measures of Inequality

VI.a. *General Form*

We started by considering two properties: an inequality per pound which is invariant when all incomes are multiplied by the same number, and an inequality per person which is invariant when the same amount is added to all incomes. Is it not possible to find a measure which satisfies both, i.e., which encompasses all at a time the "rightist" and the "leftist" position instead of being a "centrist" compromise which may betray both of them? True, the measures I_r and I_l (and I_r^a and I_l^r) we found in answer to these two basic requirements differed from each other. But to obtain them we added further conditions and most notably the "independence" condition. If we accept dropping the latter for the sake of reconciling the two first requirements, we may find a solution to this problem.

[1] A number of this paper's results were already presented—but without the proofs—at the 1966 Biarritz conference of the International Economic Association on Public Economics ([1, Sects. VI, VIII]).

In fact, we know that there exists at least one such measure: good old standard deviation of incomes σ taken as a measure of inequality per person, and its companion the coefficient of variation $\sigma/\bar{x}$ for the corresponding measure of inequality per pound, since $\sigma = [(1/n) \sum (x_i - \bar{x})^2]^{1/2}$ is invariant when the same amount is added to all x_i's, and $\sigma/\bar{x} = [(1/n) \sum ((x_i/\bar{x}) - 1)^2]^{1/2}$ is invariant when all x_i's are multiplied by the same scalar λ. They also satisfy the requirements of being zero if all x_i's are equal and positive otherwise, and of symmetry. And if we transfer one penny from person i to person j, $\bar{x}$ does not change and $\sum (x_i - \bar{x})^2$ is increased by $2(x_j - x_i)$, which is a decrease if $x_j < x_i$, and σ and $\sigma/\bar{x}$ vary in the same direction: both these measures thus satisfy the "transfer principle," i.e., they are "rectifiant," and also Schur-convex since they are symmetrical. And if we change variables from x_i into $x_i/\bar{x}$, the average of which is 1, this property for σ shows that it also holds for $\sigma/\bar{x}$ as function of the $x_i/\bar{x}$'s. Therefore, if a distribution has its Lorenz curve uniformly above that of another one, it also has a smaller $\sigma/\bar{x}$ (and thus a smaller σ if the two distributions have the same average and total incomes).

Let us now find the most general form of inequality measure which satisfies the required properties. We now call $I(x)$ the measure of per person ("absolute") inequality. Per pound ("relative") inequality is $I/\bar{x}$. We assume $I = 0$ when all incomes are equal. We recall that average income $\bar{x}$ is transformed into $\bar{x} + \mu$ when number μ is added to all x_i's, and into $\lambda\bar{x}$ when all x_i's are multiplied by number λ. e is the n-vector each coordinate of which is 1. Constancy of inequality per person when everyone receives the same amount μ is

$$I(x + \mu e) = I(x).$$

Constancy of inequality per pound when all incomes are multiplied by the same number λ, which we assume is positive, implies

$$I(\lambda x) = \lambda \cdot I(x).$$

These two properties imply

$$I[\lambda \cdot (x + \mu e)] = \lambda \cdot I(x),$$

or, calling $\nu = \lambda\mu$,

$$I(\lambda x + \nu e) = \lambda \cdot I(x),$$

which contains each of them as special cases ($\lambda = 1$ and $\nu = 0$), and is thus equivalent to the set of these two conditions.

Now, we can put $\mu = -\bar{x}$ and, when all x_i's are not equal and thus $\sigma \neq 0$, $\lambda = 1/\sigma$. This transforms the condition into

$$I(x) = \sigma \cdot I((x - \bar{x}e)/\sigma),$$

that is, I/σ is a function of the "reduced income discrepancies" $(x_i - \bar{x})/\sigma$. Conversely, given any function of n variables F,

$$I(x) = \sigma \cdot F[\{(x_i - \bar{x})/\sigma\}]$$

satisfies $I(x + \mu e) = I(x)$ and $I(\lambda x) = \lambda \cdot I(x)$. This form is thus the most general one satisfying the two conditions [1, Theorem 10].

The other properties required from an inequality measure $I(x)$ impose properties of function F. Clearly, $I(x)$ is symmetrical, positive when not all x_i's are equal, zero when all x_i's are equal if, and only if, respectively, F is symmetrical, F is positive when all its arguments are not equal, and σF tends to zero when all x_i's tend to be equal (i.e., to $\bar{x}$). If F is linearly homogeneous, $I = F[\{x_i - \bar{x}\}]$, and therefore I is zero at equality or Schur-convex if and only if F is respectively zero when all its arguments are, or Schur-convex (since a transfer does not change $\bar{x}$ and the $x_i - \bar{x}$'s are classified as the x_i's). Of course, if F is one, I is σ; if it is a constant, I is proportional to σ; if it is a standard deviation or the square root of an arithmetic average of squares, F is one and I is σ; if it is any function of such a form, it is constant and I is proportional to σ; we finally notice that when there are only two persons, 1 and 2, $\sigma = |x_1 - x_2|/2$, $(x_i - \bar{x})/\sigma = \mathrm{sgn}(x_i - x_j)$ and, with symmetrical F, $I(x) = k \cdot |x_1 - x_2|$ where k is a positive constant.

VI.b. *Inconveniences of "Independence"*

All the measures $I(x)$ found violate the "welfare independence" property, since, if it were not so, all conditions but $I(x + \mu e) = I(x)$ would give $\bar{x}I_r$, and all conditions but $I(\lambda x) = \lambda \cdot I(x)$ would give I_l, whereas these two functions are inconsistent with each other (when they are not identically zero). We may for instance check that, k being any positive constant,

$$\frac{\partial(\bar{x} - k\sigma)/\partial x_i}{\partial(\bar{x} - k\sigma)/\partial x_j} = \frac{\sigma - k(x_i - \bar{x})}{\sigma - k(x_j - \bar{x})}$$

depends upon x_l for l neither i nor j by the intermediary of $\bar{x}$ and σ. There thus exists no functions Φ and φ such that $\bar{x} - k\sigma = \Phi[\sum \varphi(x_i)]$. Therefore, if there is "welfare independence," σ (or $k\sigma$) cannot be $I \equiv \bar{x} - \bar{\bar{x}}$ where $\bar{\bar{x}}$ is the equal equivalent. But may we then use it instead of I, along with $\bar{x}$, to classify distributions by an ordinal index $U(\bar{x}, \sigma)$ more general than $\bar{x} - k\sigma$ or increasing functions of this expression? There would thus exist functions U, Φ, and φ such that the welfare index is $U(\bar{x}, \sigma) \equiv \Phi[\sum \varphi(x_i)]$. This can happen if and only if φ is a quadratic function, which can be checked as is done in the theory of choice under uncertainty for a similar result (the formal difference is the existence of probabilities in the latter case). Now, an equal increase in all x_i's gives to $I = \bar{x} - \bar{\bar{x}}$ the marginal variation $I' = 1 - \sum (\partial\bar{\bar{x}}/\partial x_i)$. But $\varphi(\bar{\bar{x}}) = (1/n) \sum \varphi(x_i)$ by definition, and thus $\varphi'(\bar{\bar{x}}) \cdot (\partial\bar{\bar{x}}/\partial x_i) = (1/n)\, \varphi'(x_i)$, from which

$$\varphi'(\bar{\bar{x}}) \cdot \sum (\partial\bar{\bar{x}}/\partial x_i) = (1/n) \sum \varphi'(x_i) = \varphi'(\bar{x})$$

where the last equality holds because φ' is linear because φ is quadratic. φ'' is constant, and it has to be negative if $\bar{x} > \bar{\bar{x}}$ (i.e., $I > 0$) out of equality if we take φ increasing, since this is equivalent to strict concavity of φ, i.e., to $\varphi(\bar{x}) > (1/n) \sum \varphi(x_i) = \varphi(\bar{\bar{x}})$. Thus, φ' must be a decreasing function, and $\bar{x} > \bar{\bar{x}}$ out of equality implies $\varphi'(\bar{x}) < \varphi'(\bar{\bar{x}})$, and $\sum (\partial\bar{\bar{x}}/\partial x_i) < 1$, and finally $I' > 0$: an equal increase in all incomes increases per person inequality.[2] This is an "ultra-leftist" position, which can be objected to. But it also requires the "independence" property. Why not rather drop the latter? If we do that, we know that σ is from this respect a valid companion to $\bar{x}$ to classify distributions, since it is even much more: a satisfactory absolute measure of inequality per person.

Standard deviation and coefficient of variation have also been criticized, as measures of inequality, because they give the same weight to incomes symmetrically distributed around the mean (i.e., x_i and x_j such that $x_i - \bar{x} = \bar{x} - x_j$) whereas one is larger than the other. But they do not give the same importance to *variations* in such incomes since we have noticed that a small transfer from a richer person to a poorer one decreases σ and $\sigma/\bar{x}$. This decrease even appeared to be proportional to the difference of these two incomes; but this opens the way to another possible criticism of these measures, which is the topic of the next section.

[2] But it does not necessarily increase the per pound ("relative") inequality $I/\bar{x}$. This increases or decreases according as the *relative* variation in $\bar{\bar{x}}$ is smaller or larger than that of $\bar{x}$, i.e., as $\bar{x} \cdot \varphi'(\bar{x}) \lesseqgtr \bar{\bar{x}} \cdot \varphi'(\bar{\bar{x}})$. Since φ' is linear, decreasing, and positive, this expression is of the form $(a - \bar{x})\bar{x} \lesseqgtr (a - \bar{\bar{x}})\bar{\bar{x}}$ with $a > 0$ and $x_i < a$ for all i's (which implies $0 < \bar{\bar{x}} < \bar{x} < a$). But if we choose all x_i's between 0 and $a/2$, we also have $0 < \bar{\bar{x}} < \bar{x} < a/2$ and the sign $>$ holds in the inequality, whereas if we choose all x_i's between $a/2$ and a, we also have $a/2 < \bar{\bar{x}} < \bar{x} < a$ and the sign $<$ holds in the inequality. This notwithstanding the fact that both [1]'s "marginal injustice" and "relative marginal injustice" $(-\varphi''/\varphi' = 1/(a - y)$ and $-y\varphi''/\varphi' = y/(a - y))$ are increasing with y — A. Atkinson [2] calls them "inequality-aversion" and "relative inequality-aversion" by analogy with risk theory vocabulary in English (the "risk-aversion measure" was called "prudence" in French) and suggests that "while the objections to this property are less strong than the corresponding objections in the uncertainty case, it may be grounds for rejecting the quadratic."

VII. The Principle of Diminishing Transfers

A mild equalitarian will certainly appreciate a small transfer from a richer person to a poorer one ("transfers principle," Schur-concavity of the evaluation function, "rectifiance," "isophily"). But he may go one step further and value more such a transfer between persons with given income difference if these incomes are lower than if they are higher. Thus, he would prefer to transfer one pound from a person who earns 500 pounds a month to another one who earns only 100, than to transfer one pound from a 900 pounds earner to a person who already earns 500 pounds. None of these operations changes total social nor average income. Thus, their effect on an evaluation function, which shows on that specification of it, the equal equivalent $\bar{\bar{x}}$, appears with reverse ordering on inequality measures $\bar{x} - \bar{\bar{x}}$ and $(\bar{x} - \bar{\bar{x}})/\bar{x}$. Extending H. Dalton's vocabulary, we may call this property the "principle of diminishing transfers." As the "transfers principle," this concept is an ordinal one since it is defined by a classification of differences between derivatives of the index for each given distribution (if $J(x)$ is an index with $J_i = \partial J/\partial x_i$, the inequality $J_i - J_j > J_k - J_l$ does not change when J is transformed into $F(J)$ where F is any increasing function.[3]

We have noticed that the effect of a marginal transfer from person i to person j on standard deviation σ is proportional to $x_i - x_j$. For a given discrepancy between these two incomes, it does not depend upon their level. As a measure of inequality, σ thus violates the above "principle." And so does any function of σ and $\bar{x}$, such as $k\sigma$, the coefficient of variation $\sigma/\bar{x}$, the variance σ^2, $\sigma^2/\bar{x}^2$, $\sigma^2/\bar{x}$ and any of these multiplied by k.

A. Atkinson [2] pointed out this property of the coefficient of variation, and suggested that it could be a shortcoming of this measure. But, how fare, in this respect, the other measures mentioned in the precedent sections?

For "welfare independent" and "impartial" measures, of the form $\bar{x} - \bar{\bar{x}}$ or $(\bar{x} - \bar{\bar{x}})/\bar{x}$ with $\varphi(\bar{\bar{x}}) = (1/n) \sum \varphi(x_i)$ (cf. Section XI.d. below), we consider the equivalent property on the equal equivalent income $\bar{\bar{x}}$. For x_i, x_j, x_k, x_l such that $x_i < x_j$, $x_k < x_l$, $x_j - x_i = x_l - x_k$, $x_i < x_k$, $x_j < x_l$, we want to know whether $(\partial\bar{\bar{x}}/\partial x_i) - (\partial\bar{\bar{x}}/\partial x_j) \gtrless (\partial\bar{\bar{x}}/\partial x_k) - (\partial\bar{\bar{x}}/\partial x_l)$. From the definition $\varphi(\bar{\bar{x}}) = (1/n) \sum \varphi(x_i)$ and thus $\varphi'(\bar{\bar{x}}) \cdot (\partial\bar{\bar{x}}/\partial x_i) = (1/n)\, \varphi'(x_i)$, this is equivalent to $\varphi'(x_i) - \varphi'(x_j) \gtrless$

[3] A still more egalitarian concept would be a "principle of relatively diminishing transfers" saying that a small transfer is more equalizing from j to i than from l to k (i.e., it decreases inequality more, or, since $\bar{x}$ remains unchanged, it is preferred) when $x_j/x_i = x_l/x_k$ and $x_i < x_j$, $x_k < x_l$, $x_i < x_k$, $x_j < x_l$. It also is an ordinal concept.

$\varphi'(x_k) - \varphi'(x_l)$ if $\varphi'(\bar{x})$ is positive and the reverse inequalities if it is negative. This is in turn a way of writing that φ' is convex or concave, another way being, if φ''' exists, $\varphi''' \gtrless 0$ (almost everywhere for the strict inequalities). And $(\partial\bar{\bar{x}}/\partial x_i) > 0$ for all possible income levels x_i (for $(\partial\bar{\bar{x}}/\partial x_i) \neq 0$) imposes that φ' has the same sign everywhere. Therefore, the "principle of diminishing transfers," or its opposite, is true, according as φ' and φ''' have the same or opposite signs. And since φ' and φ'' have opposite signs (seen to be necessary for $\bar{x} > \bar{\bar{x}}$ out of equality), the condition is that φ', φ'' and φ''' alternate in signs or not.

Now, for $\varphi(y) = (y + \xi)^{1-\epsilon}$, including the special case where $\xi = 0$, φ', φ'', and φ''' have the respective signs of $1 - \epsilon$, $(1 - \epsilon)(-\epsilon)$, $(1 - \epsilon)(-\epsilon)(-\epsilon - 1)$. They alternate if and only if $\epsilon > 0$, which is required by the condition that φ' and φ'' alone differ in sign. These derivatives also alternate in sign for $\varphi(y) = \mathrm{Log}(y + \xi)$, including the special case $\xi = 0$. And for $\varphi(y) = e^{-\alpha y}$, the derivative alternate in sign if and only if $\alpha > 0$, i.e., if φ' and φ'' alone differ in sign. Therefore, all inequality measures I_r, I_l, I_c, I_r^a, I_l^r, I_c^r satisfy the principle of diminishing transfers.

The effect of a transfer on the measures derived from σ still poses another problem. We have $\partial\sigma/\partial x_i = (x_i - \bar{x})/n\sigma$ and

$$\frac{\partial(\sigma/\bar{x})}{\partial x_i} = \frac{1}{n\bar{x}}\left(\frac{x_i - \bar{x}}{\sigma} - \frac{\sigma}{\bar{x}}\right),$$

which shows the Schur-convexity, and $(\partial\sigma/\partial x_i) - (\partial\sigma/\partial x_j) = (x_i - x_j)/n\sigma$ and

$$\frac{\partial(\sigma/\bar{x})}{\partial x_i} - \frac{\partial(\sigma/\bar{x})}{\partial x_j} = \frac{x_i - x_j}{n\sigma\bar{x}},$$

which shows the proportionality to the difference of incomes. We then see that a small transfer from j to i ($x_i < x_j$) decreases σ and $\sigma/\bar{x}$ more when σ is smaller, for given x_i and x_j (and $\bar{x}$ for $\sigma/\bar{x}$); i.e., the transfer decreases inequality more when inequality is smaller. But the effects of this transfer on variance σ^2 (unchanged by an equal variation in all incomes), or $\sigma^2/\bar{x}^2$ (unchanged by an equiproportional variation in all incomes), and on $\sigma^2/\bar{x}$ (proportional to an equiproportional variation in all incomes) are respectively $(2/n)(x_i - x_j)$, $(2/n\bar{x}^2)(x_i - x_j)$ and $(2/n\bar{x})(x_i - x_j)$: none depends upon the inequality measure. However, the last two and the effect on $\sigma/\bar{x}$ are smaller when $\bar{x}$ is larger; we may find this objectionable, because then the fixed x_i and x_j become in some sense smaller relative to the rest of the distribution.

VIII. Inequality Measures and Lorenz and Concentration Curves

We mentioned several times the equivalence between the "transfers principle" and smaller inequality for distributions of the same total whose Lorenz curve is everywhere above. The transfers principle itself only compares distributions of the same total amount. What can be said for distributions which do not have the same total and average amount?[4]

We shall have to consider the "concentration curve" of a distribution. By this name we call the graph of the sum of the m smallest incomes as a function of m. More precisely, $x = \{x_i\}$ $(i = 1,..., n)$ being an income distribution, we reorder the x_i's into the $x_i{}'$'s with $x_1{}' \leqslant x_2{}' \leqslant \cdots \leqslant x_n{}'$ and each $x_i{}'$ is an x_j.[5] We then call $y_j = \sum_{i=1}^{j} x_i{}'$. The concentation curve is y_j as a function of j. We obviously have $y_n = \sum x_i = X$, and $y_1 = \text{Min}_i\, x_i$.

We also define $\eta_i = y_i/y_n = y_i/X$. Of course, $\eta_n = 1$. The Lorenz curve is obtained by plotting η_i against the figure i/n. We call $x' = \{x_i{}'\}$, $y = \{y_i\}$, and $\eta = \{\eta_i\} = y/X$ the n-vectors of the $x_i{}'$'s, $y_i{}'$'s and η_i's. By the relation $\geqslant$ between two vectors of same dimension, we mean $\geqslant$ for each coordinate and $>$ for at least one. Superscript 1 and 2 will refer to two distributions which we compare. $y^1 \geqslant y^2$ means that x^1's concentration curve is "nowhere under and somewhere above" x^2's. $\eta^1 \geqslant \eta^2$ means that x^1's Lorenz curve is "nowhere under and somewhere above" x^2's.[6]

The following relations hold. Of course, $x^1 \geqslant x^2$ or $y^1 \geqslant y^2$ implies $X^1 \geqslant X^2$. And if $X^1 = X^2$, $y^1 \geqslant y^2$ and $\eta^1 \geqslant \eta^2$ imply each other. Also, $\eta^1 \geqslant \eta^2$ and $X^1 \geqslant X^2$ imply $y^1 \geqslant y^2$ (but $y^1 \geqslant y^2$ does not imply $\eta^1 \geqslant \eta^2$). $y^1 = y^2$, $x'^1 = x'^2$, and x^1 and x^2 are a permutation of each other (i.e., their coordinates are), are equivalent properties. $x'^1 \geqslant x'^2$ implies $y^1 \geqslant y^2$, and $x^1 \geqslant x^2$ implies $x'^1 \geqslant x'^2$ (see, for instance, [3, pp. 108–109]) and $y^1 \geqslant y^2$.

Writing $y(x)$ and $\eta(x)$ for the vector functions by which y and η are derived from x, one has $\eta(x) = y(x)/X = y(x/X)$ since each y_j is a linear homogeneous function of the x_i's. This shows that *for intensive inequality measures* (i.e., $I(\lambda x) = I(x)$ for all admissible λ's and x's), *relations with Lorenz or concentration curves are equivalent.* Among these measures are I_r and $\sigma/\bar{x}$, and I_r is the only "welfare independent" one which has this

[4] Cf. [1, Theorems 1–6].

[5] i as a function of $x_i{}'$ thus is the number of persons whose income is not larger than $x_i{}'$. i/n as function of $x_i{}'$ is therefore the cumulative distribution function of the x_i's.

[6] In [1, Sect. VI], a preference for a higher concentration curve is called "isophily," and a preference for a higher Lorenz curve for distributions of the same amount is called "constant-sum isophily" ("isophile" = who likes equality).

property. For other inequality measures, the relation between their relations with Lorenz and concentration curves depends upon the effect on them of a multiplication of x by a scalar (even though X is a special one). Results of Part I Section V, show that all the other measures which have been considered vary in the same direction as such an equiproportional variation in all x_i's, with the exceptions there mentioned: we shall say that inequality is *subintensive* when $I(\lambda x) \gtreqless I(x)$ depending on $\lambda \gtreqless 1$ for all admissible nonequal x_i's and λ's ("sub" is here because this property may be considered as more moderate than intensiveness).

We call again $V(x)$ an ordinal, differentiable, increasing ("benevolence"), strictly Schur-concave ("impartiality"-symmetry and "rectifiance"-"transfers principle") evaluation function, $\bar{\bar{x}}$ defined by $V(x) = V(\bar{\bar{x}}e)$ the equal equivalent income, and $I^a = \bar{x} - \bar{\bar{x}}$ and $I^r = 1 - (\bar{\bar{x}}/\bar{x})$ the inequalities per person and per pound. These inequality measures are zero when all x_i's are equal ($x = \bar{x}e = \bar{\bar{x}}e$). Obviously, Schur-concavity of V and $\bar{\bar{x}}$ and Schur-convexity of I^a and I^r are all equivalent properties. They are both symmetry of these four functions of x and the "rectifiance" conditions: $x_i < x_j$ implies $\partial V/\partial x_i > \partial V/\partial x_j$, $\partial\bar{\bar{x}}/\partial x_i > \partial\bar{\bar{x}}/\partial x_j$, $\partial I/\partial x_i < \partial I/\partial x_j$ for I being I^a or I^r (we now consider only the "strict" forms). If V is "independent," i.e., if $V = \Phi[\sum \varphi(x_i)]$, these latter conditions are equivalent to strict concavity of φ if it is increasing (Φ increasing), to its strict convexity if it is decreasing (Φ decreasing). We recall that all the specific inequality measures previously considered are Schur-convex (we exclude the trivial identically zero cases).

For constant-sum comparisons, i.e., comparisons between x^1 and x^2 such that $X^1 = X^2$, $y^1 \geqslant y^2$ and $\eta^1 \geqslant \eta^2$ are equivalent, and since $\bar{x}^1 = \bar{x}^2$, I^a and I^r vary in the same direction (I will be I^a or I^r) and V in the opposite one. Then, $\eta^1 \geqslant \eta^2$ and the transfers principle are equivalent in the following sense. I is Schur-convex if and only if $I(x^1) < I(x^2)$ for all x^1 and x^2 such that $\eta^1 \geqslant \eta^2$; and $\eta^1 \geqslant \eta^2$ if and only if $I(x^1) < I(x^2)$ for all Schur-convex I.[7]

We consider now the more general case where $X^1 \geqslant X^2$ and $\bar{x}^1 \geqslant \bar{x}^2$. This inequality is implied by $y^1 \geqslant y^2$. We now have the more general

[7] This relation contains three propositions: (1°) $I(x^1) < I(x^2)$ if I is Schur-convex and $\eta^1 \geqslant \eta^2$, (2°) I is Schur-convex if $I(x^1) < I(x^2)$ for all x^1 and x^2 such that $\eta^1 \geqslant \eta^2$, (3°) $\eta^1 \geqslant \eta^2$ if $I(x^1) < I(x^2)$ for all Schur-convex I. The two first ones result from Ostrowski's characterization of strict Schur-convexity by $I(Bx) < I(x)$ for all bistochastic matrix B when Bx is not a permutation of x [4], and from the equivalence between $\eta^1 \geqslant \eta^2$ and altogether there exists a bistochastic matrix B such that $x^1 = Bx^2$ (a direct result from Hardy–Littlewood–Pòlya's Theorem 46 [5]) and x^1 is not a permutation of x^2. The third one results from an easy to prove strict form of Karamata's inequality [6] by considering the Schur-convex functions I of the form $\sum \Psi(x_i)$ with convex Ψ.

result that $y^1 \geqslant y^2$ and the transfers principle are equivalent in the following sense. *I is Schur-convex if and only if* $V(x^1) > V(x^2)$ *for all* x^1 *and* x^2 *such that* $y^1 \geqslant y^2$; *and* $y^1 \geqslant y^2$ *if and only if* $V(x^1) > V(x^2)$ *for all Schur-convex I* (and thus Schur-concave V).

Let us prove these latter results. If $V(x^1) > V(x^2)$ for all x^1 and x^2 such that $y^1 \geqslant y^2$, it is in particular so for all x^1 and x^2 for which, in addition, $X^1 = X^2$, and I is thus Schur-convex (V Schur-concave) from a previous result. The property $V(x^1) > V(x^2)$ if V is Schur-concave and $y^1 \geqslant y^2$ is Ostrowski's Theorem V of [4]. If $V(x^1) > V(x^2)$ for all (strictly) Schur-concave V, by continuity $V(x^1) \geqslant V(x^2)$ for all weakly Schur-concave V (i.e., V such that $x_i < x_j$ implies $V_i \geqslant V_j$); but y_j is a weakly Schur-concave function of x; thus, $y_j^1 \geqslant y_j^2$ for all j; and $y_j^1 = y_j^2$ for all j would imply $x'^1 = x'^2$ and thus $V(x^1) = V(x^2)$ for a (strictly) Schur-concave V; therefore $y^1 \geqslant y^2$.

These are results about concentration curves' "dominance." For Lorenz curves and non-constant-sum comparisons, the following results hold.

(1°) *If* $\eta^1 \geqslant \eta^2$ *and* $X^1 \geqslant X^2$, $V(x^1) > V(x^2)$ *for all Schur-concave V's.*

(2°) *For intensive Schur-convex inequality measures,* $I(x^1) < I(x^2)$ *if* $\eta^1 \geqslant \eta^2$.

(3°) $\eta^1 \leqslant \eta^2$ *and* $X^1 \geqslant X^2$ *imply* $I(x^1) > I(x^2)$ *for subintensive Schur-convex inequality measures.*

We note that, among the specific inequality measures considered above, I_r and $\sigma/\bar{x}$ are intensive, and all others are subintensive (except I_c for $\xi < 0$; and I_r is the only "welfare independent" intensive measure).

Property (1°) holds because $\eta^1 \geqslant \eta^2$ and $X^1 \geqslant X^2$ imply $y^1 \geqslant y^2$. Property (2°) holds because, for inequality measures such that $I(\lambda x) = I(x)$, and noting that η is the y of x/X,

$$I(x^1) = I(x^1/X^1) < I(x^2/X^2) = I(x^2).$$

Property (3°) holds because, for inequality measures such that $I(\lambda x) \gtreqless I(x)$ out of equality depending on $\lambda \gtreqless 1$, and since $\eta^1 \leqslant \eta^2$ is identical to $y^1X^2/X^1 \leqslant y^2$ whereas y^1X^2/X^1 is the y of x^1X^2/X^1,

$$I(x^2) < I(x^1X^2/X^1) \leqslant I(x^1).$$

IX. Additions of Incomes and Diminishing Equality

IX.a. *General Properties*

When we add incomes of several kinds, how does inequality in the global income depend upon the inequalities of the various components? This certainly is a useful question. The "incomes" added could for instance be: earned and unearned incomes, or yearly increments in incomes (which shows what inequality variation is induced by the growth of incomes), or private incomes and government transfers (plus, possibly, the value to persons of free government services), or after tax incomes and taxes (the sum of which gives before tax incomes, so that we relate taxes' effect on income inequality to fiscal inequality), etc. Or we may want to consider inequalities in wealth and in various kinds of wealth (nonhuman and human, etc.), or to relate the variations of inequality in wealth holding with the inequality in net savings (wealth increments), etc. Note that our "addition" would be the statisticians' "composition" of income distributions. We shall obtain the property that, for some per person or total measures, inequality in the sum is less than the sum of inequalities; we shall call this the *subadditivity* property. And, for these measures, growth of inequality is less than the inequality in growth. For the corresponding per pound measures, these relations hold when the inequalities are weighted by average or total incomes (i.e., a weight is one of these incomes divided by their sum); we shall call this the *weighted subadditivity* property.

When we consider weighted sums of incomes, rather than unweighted ones, we are introduced to the related properties of "convexity." This answers questions such as the following. Suppose we progressively, regularly, and proportionally bridge the gap from some income distribution to a more equally distributed one. Will we meet some sort of diminishing returns so that the decrease in inequality index is slower and slower? In mathematical terms, this property would be the convexity of the inequality index $I(x)$ in the set of all x_i's. Since $I = 0$ when all x_i's are equal, it implies that if we bridge some proportion of the gap from some unequal distribution to any equality, inequality is reduced by more than this proportion.

Convexity implies quasi-convexity, i.e., a distribution in which each income is the same average of what it is in several other ones is not more unequal than the most unequal of the latter. This is equivalent to saying that if the latter have the same degree of inequality, the average's is not higher. By average we mean a weighted linear one, but in this definition we can restrict ourselves to such an average with given weights—for instance to arithmetic averages—and also to the consideration of only pairs of averaged distributions. Quasi-convexity is "strict" when "not more unequal" can be replaced by "less unequal."

Quasi-convexity, in turn, plus symmetry, imply Schur-convexity which implies the "transfers principle" (rectifiance) and has interesting equivalent properties such as the mentioned one between inequality measure and Lorenz, or concentration, curves ("isophily").

All these convexity, quasi-convexity and Schur-convexity properties could be either valid for any distributions, or restricted to relations between distributions of the same total amount and which hence have the same average income (but they would then hold for all levels of these total or average incomes); in this latter case, we will add the adjective "*constant-sum*" to the property.

IX.b. *Results*

The results will roughly be the following. *I_r^a and σ are subadditive. I_r and $\sigma/\bar{x}$ present weighted subadditivity. I_l, I_c, I_r^a, σ are convex. I_r, I_c^r, I_l^r, $\sigma/\bar{x}$ are constant-sum convex.*

Let us be more precise.

For I_r^a and σ: a sum of inequalities is not larger than the inequality of the sum; a sum of inequalities of nonproportional distributions is smaller than the inequality of their sum. For I_r and $\sigma/\bar{x}$: a sum of inequalities weighted by average or total incomes is not larger than the inequality of the sum; a sum of inequalities of nonproportional distributions, weighted by their average or total incomes, is smaller than the inequality of their sum.

I_r^a and σ are strictly convex for distributions which are not all proportional to each other. I_l is strictly convex for distributions which do not all differ from each other by the same difference in all incomes.

The properties of the exceptions to all these cases are already known. Sums and weighted sums of proportional distributions are also proportional to them. They thus all have the same I_r and $\sigma/\bar{x}$. And their I_r^a and σ are as their proportions are: the sum's is the sum and the weighted sum's is the weighted sum. As for I_l, sums and weighted sums of distributions which differ from each other by an equal difference in all incomes also belong to this class, and they all have the same I_l.

Finally, we shall also find for I_c and I_c^r properties which extend the subadditivities of I_r^a and I_r and which we shall call "pseudo subadditivity" and its weighted form. Also, the cases where I_c is not strictly convex will appear.

IX.c. *Demonstrations*

IX.c.1. *Rightist and centrist measures* [1, Theorem 22]. Let x^k be several distributions with index k, $\bar{x}^k$ their average incomes, $X^k = n\bar{x}^k$ their total incomes.

For $\bar{\bar{x}}(x) = \bar{x} - I_r^a(x) = [(1/n)\sum x_i^{1-\epsilon}]^{1/1-\epsilon}$ or $(\Pi x_i)^{1/n}$ with $\epsilon > 0$, Minkowski's inequality gives $\bar{\bar{x}}(\sum x^k) \geqslant \sum \bar{\bar{x}}(x^k)$ and thus

$$I_r^a\left(\sum x^k\right) \leqslant \sum I_r^a(x^k)$$

with equality if and only if all x^k's are proportional: inequality measure I_r^a is *subadditive.*

Since $I_r^a = \bar{x} I_r$, this gives

$$I_r\left(\sum x^k\right) \leqslant \sum \left(\bar{x}^k / \sum \bar{x}^k\right) I_r(x^k) = \sum \left(X^k / \sum X^k\right) I_r(x^k)$$

with equality if and only if all the x^k's are proportional (they then all have the same I_r which is also their sum's). This is the property of *wieghted subadditivity* of I_r.

Subadditivity and linear homogeneity of I_r^a show that

$$I_r^a\left(\frac{x^1 + x^2}{2}\right) = \frac{1}{2} I_r^a(x^1 + x^2) \leqslant \frac{I_r^a(x^1) + I_r^a(x^2)}{2},$$

i.e., I_r^a is a *convex* function of x, with strict convexity for nonproportional x's.

$I_r = I_r^a/\bar{x}$ is thus convex at given $\bar{x}$ and $X = n\bar{x}$, whatever these levels of $\bar{x}$ and X. We may call this property *constant-sum convexity* of I_r.

The change in variables which transforms I_r^a into I_c does not affect the convexity properties. I_c is thus *convex* in x. And $I_c^r = I_c/\bar{x}$ is therefore convex at given $\bar{x}$ or $X = n\bar{x}$: it also has the *constant-sum convexity* property. Furthermore, I_r^a's subadditivity and this change of variables show that I_c has the following *pseudo subadditivity* property where m is the number of k's

$$I_c\left(\sum_{k=1}^{m} x^k + m\xi\right) \leqslant \sum_{k=1}^{m} I_c(x^k + \xi),$$

and that I_c^r thus has the *weighted pseudo subadditivity* property

$$I_c^r\left(\sum x^k + m\xi\right) \leqslant \sum \left(\bar{x}^k / \sum \bar{x}^k\right) I_c^r(x^k + \xi) = \sum \left(X^k / \sum X^k\right) I_c^r(x^k + \xi).$$

IX.c.2. *Standard deviation and coefficient of variation.* Call $\rho(x) = (\sum x_i^2)^{1/2}$. From Minkowski's inequality, $\rho(\sum x^k) \leqslant \sum \rho(x^k)$ with equality if and only if all x^k's are proportional. That is, $\rho(x)$ is subadditive. As shown above, this implies it is convex, strictly for nonproportional x's. Changing variables from x_i into $x_i - \bar{x}$ and ρ into $\rho/n^{1/2}$ does not change concavity for given $\bar{x}$: $\sigma = [(1/n) \sum (x_i - \bar{x})^2]^{1/2}$ is thus convex for given $\bar{x}$, i.e., for x^k's with same average $\bar{x}^k$. And since such x^k's cannot be proportional without being equal, σ is strictly convex for given $\bar{x}$. But it is also linearly homogeneous. Therefore, the hypersuface graph of $I(x)$ is a cone whose summit is the origin and having a strictly convex base in the hyperplane $\sum x_i = n\bar{x}$ for this given $\bar{x}$: it is in this sense a "strictly" convex cone. Therefore, the results are that σ is both a *convex* and a *subadditive* function of x, strictly for nonproportional x's. In particular,

$$\sigma\left(\sum x^k\right) \leqslant \sum \sigma(x^k)$$

with equality if and only if all x^k's are proportional.

Consequently, the coefficient of variation $\sigma/\bar{x}$ has the properties of *constant-sum convexity* and of *weighted subadditivity* with weights proportional to average or total incomes and equality if and only if all x^k's are proportional (they then all have the same $\sigma/\bar{x}$ which is also their sum's).

IX.c.3. *Leftist measures.* For I_l, $\bar{\bar{x}}(x) = \bar{x} - I_l(x) = (1/\alpha) \times \operatorname{Log}((1/n) \sum e^{-\alpha x_i})$ with $\alpha > 0$ will be shown to be concave by the method of directional derivatives: we choose n numbers z_i, replace x_i by $x_i + z_i t$, and compute the derivatives $\bar{\bar{x}}'$ and $\bar{\bar{x}}''$ of $\bar{\bar{x}}$ for t at $t = 0$. $\bar{\bar{x}}(x)$ is concave if and only if $\bar{\bar{x}}'' \leqslant 0$ for all z_i's. Let us thus write

$$e^{-\alpha\bar{\bar{x}}} = (1/n) \sum e^{-\alpha\cdot(x_i+z_i t)}.$$

Differentiating twice gives

$$e^{-\alpha\bar{\bar{x}}}\bar{\bar{x}}' = (1/n) \sum e^{-\alpha\cdot(x_i+z_i t)} z_i\,,$$

and

$$e^{-\alpha\bar{\bar{x}}}\bar{\bar{x}}'' = \alpha e^{-\alpha\bar{\bar{x}}}\bar{\bar{x}}'^2 - (\alpha/n) \sum e^{-\alpha x_i} z_i^2$$

for $t = 0$. For this t, carrying $\bar{\bar{x}}'$ from the second equation into the third one and using the first one, we finally obtain

$$(n^2/\alpha)\, e^{-\alpha\bar{\bar{x}}}\bar{\bar{x}}'' = \left(\sum e^{-\alpha x_i} z_i\right)^2 - \left(\sum e^{-\alpha x_i} z_i^2\right)\left(\sum e^{-\alpha x_i}\right).$$

Let us now apply Cauchy's theorem ("the square of a sum of products is smaller than the product of the sums of squares, unless the variables are proportional") to the two series of numbers $e^{-\alpha x_i/2} z_i$ and $e^{-\alpha x_i/2}$:

$$\left(\sum e^{-\alpha x_i} z_i\right)^2 \leqslant \left(\sum e^{-\alpha x_i} z_i^2\right)\left(\sum e^{-\alpha x_i}\right)$$

with equality if and only if all z_i's are equal. Thus, $\bar{\bar{x}}'' \leqslant 0$, with equality for equal variations of all x_i's and only in this case. And finally $I_l(x) =$

$\bar{x} - \bar{\bar{x}}(x)$ is *convex*, strictly except on the directions of equal changes in all x_i's.

$I_l^r = I_l/\bar{x}$ is thus *constant-sum convex*.

IX.d. *Applications of Sub-Additivity of Inequality*

Per person and total inequalities are linearly homogeneous and convex if and only if the corresponding per pound inequality is intensive and constant-sum convex. This is equivalent to respectively subadditivity and weighted subadditivity of these measures. (These relations straightforwardly result from the precedent demonstrations). This subsection deals with such inequality measures. It thus applies in particular to the per person I_r^a and σ and the corresponding per pound I_r and $\sigma/\bar{x}$. Its object is to show examples of applications of the subadditivity properties.

If for instance Y is national income, I_{global} the inequality in its distribution, and Y_{earned}, Y_{unearned}, I_{earned}, I_{unearned} respectively the total earned and unearned incomes and the inequalities in their distributions,

$$I_{\text{global}} \leqslant I_{\text{earned}} + I_{\text{unearned}}$$

for per person and total inequalities, and

$$I_{\text{global}} \leqslant (Y_{\text{earned}}/Y)\, I_{\text{earned}} + (Y_{\text{unearned}}/Y)\, I_{\text{unearned}}$$

for per pound inequalities. We know too well that the condition for equality in these relations does not hold.

We can also consider the effect of growth on income inequality. If the x^k's are successive income increments, the relations of the precedent subsection are between the inequality in some incomes $\sum x^k$ and the inequalities in their successive increments. But let us rather call now x^t the income distribution at time t, and write $I_t = I(x^t)$ for inequality at time t.

For per person or total inequality, the following relation holds between income inequalities in years t and $t + 1$ and the inequality in the yearly increments $\Delta x^t = x^{t+1} - x^t$:

$$I_{t+1} \leqslant I_t + I(\Delta x^t),$$

or

$$I_{t+1} - I_t \leqslant I(\Delta x^t),$$

with equality in the relation if and only if all incomes grow in the same proportion; that is: *the increment of inequality is lower than the inequality* *of the increment*, except when all incomes increase in the same proportion, in which case they are equal. θ being a time interval, we can also write

$$I_{t+\theta} \leqslant I_t + I(x^{t+\theta} - x^t)$$

or

$$\frac{I_{t+\theta} - I_t}{\theta} \leqslant \frac{I(x^{t+\theta} - x^t)}{\theta} = I\left(\frac{x^{t+\theta} - x^t}{\theta}\right)$$

the last equality holding because of the linear homogeneity of these measures. Letting θ tend to zero, and using Newton's dot to indicate time derivatives, this inequality becomes

$$\dot{I} \leqslant I(\dot{x}).$$

Equality again holds when $\dot{x}$ and x are proportional, and it does not hold when they are not.[8] This again says, but now for the time derivatives, that the increment in equality is not higher than the inequality in the increment, both being equal if and only if all incomes grow at the same rate. This result is also conveniently written as between the growth rate of inequality $\dot{I}/I$ and the relative inequality of the growth tendency $I(\dot{x})/I$ (if $I > 0$), as $\dot{I}/I \leqslant I(\dot{x})/I$: *the inequality growth rate is not higher higher than the relative inequality of the growth tendency*, and they are equal if and only if all incomes have the same growth rate.

For per pound inequality, we similarly have

$$I_{t-1} \leqslant \frac{X^t}{X^{t+1}} I_t + \left(1 - \frac{X^t}{X^{t-1}}\right) \cdot I(\Delta x^t)$$

with equality if and only if x^t, x^{t+1}, and Δx^t are proportional, in which case their three inequalities are equal. Writing

$$\frac{X^{t+\theta} I_{t+\theta} - X^t I_t}{\theta} \leqslant \frac{X^{t+\theta} - X^t}{\theta} \cdot I\left(\frac{x^{t+\theta} - x^t}{\theta}\right)$$

(since $I(x^{t+\theta} - x^t) = I((x^{t+\theta} - x^t)/\theta)$ for these measures), and letting θ tend to zero, we obtain

$$I\dot{X} + X\dot{I} \leqslant \dot{X} \cdot I(\dot{x}).$$

The same remark as above holds for the equality and strict inequality cases. But, now, proportionality between x and $\dot{x}$ means $I = I(\dot{x})$, which

[8] The passage to the limit does not guarantee this assertion. But it holds because if, in R^{n+1} space (I, x), we consider the convex half-cone graph of $I(x)$, and the half-cone translated from it and whose summit is point (I, x), the latter half-cone lies completely in the interior of the former one except for its ray on the line from origin to this point, which they have in common.

the relation, with equality, shows to be equivalent to $\dot{I} = 0$ (since $\dot{X} > 0$). Apart from these cases and for $I > 0$, the relation can be written as

$$\frac{\dot{I}/I}{\dot{X}/X} < \frac{I(\dot{x}) - I}{I};$$

the growth rate of inequality is lower than the growth rate in global income times the relative excess of the inequality in growth tendency over the present one, except when all incomes grow at the same rate, in which case inequality does not change.

We can also write a relation about the effect of government welfare transfers on income inequality: the income inequality after transfers is lower than the before transfers one plus the inequality in transfers (for per person or total inequalities; for per pound ones the sum should be weighted by the respective volumes of incomes and transfers). The equality case in the relations is of course irrelevant. But for the same reason the result is not very informative in this case, since one of the usual reasons for transfers is to decrease income inequality. However, the relation becomes much more interesting if we consider the money equivalent of government services for the persons, so as to see how the inequality of benefits from government expenditures mixes with that of private incomes. Again, total inequality would generally be lower than the sum of these inequalities (weighted by private income and government expenditure levels for the per pound measures). But, for this problem, the equality case (proportionality of benefits to incomes) has a high degree of empirical plausibility.

All this sounds rather optimistic, after all; if we add incomes, inequality increases less (per person), or is lower than the highest (per pound). However, the main tool to affect the inequality of incomes in our society does not add to them but subtracts from them: it is the tax system. Now, when adding income distributions, we excluded the possibility of "negative" incomes, by the very definition of an income distribution. But if some variation is a decrease in all incomes (or at least no increase in any), we may consider it as a positive (or nonnegative) addition to the final distribution to obtain the initial one. Thus, if I_{at}, I_{bt}, I_t respectively are inequalities in after tax and before tax incomes and in the tax distribution, we have

$$I_{at} \geqslant I_{bt} - I_t$$

for inequality per person or total, and, calling Y the global income and T the tax revenue,

$$I_{at} \geqslant (Y/(Y-T))\, I_{bt} - (T/(Y-T))\, I_t$$

for inequality per pound. So, after tax inequality is not lower than before tax inequality less fiscal inequality (with weights equal to the respective amounts in the case of per pound inequality). It is equal to this difference for a proportional income tax, and only in this case (this of course does not imply that it is the structure which gives the lowest after tax inequality).

IX.e. *Diminishing Returns in Equality*

We have found that inequality measures I_l, I_c, I_r^a, σ are convex, whereas I_r, I_l^r, I_c^r, $\sigma/\bar{x}$ are constant-sum convex. This gives them some "diminishing returns" property, which they share with any other measures which are convex, or convex in some sub-spaces of the x space. This property is described by relations between the inequalities of several distributions. These distributions must have the same total sum or average for the constant-sum convex measures (but then the property holds for all such sums or averages), whereas no such restriction holds for the merely convex measures.

This property can be expressed as: if we move regularly along a line in x space, the increase in inequality becomes faster and faster, or the decrease in inequality slower and slower. There is, of course, a limiting exception for the fully convex measures (I_r^a, I_c, I_l, σ), obtained when this line is a projection on x space of a line located on the hypersurface $I(x)$, ray of the cone graph of I_r^a, σ or I_c, or generatrix of the cylinder graph of I_l (i.e., line parallel to the equality direction e or Δ, along which I_l is constant): inequality differences vary proportionally to distance differences for the former (for I_r^a and σ, this is an equiproportional variation in all x_i's), and inequality does not change for the latter (equal variation in all x_i's); these are the cases for which equality holds in the following relations. x^0 and x^1 being specific distributions, the property can thus be written either classically as

$$I[\lambda x^0 + (1 - \lambda)\, x^1] \leqslant \lambda \cdot I(x^0) + (1 - \lambda) \cdot I(x^1)$$

for $0 < \lambda < 1$, or differentially as

$$\sum (x_i - x_i^0)(\partial I(x)/\partial x_i) \geqslant I(x) - I(x_0).$$

It is particularly interesting to take a reference point with equality: call ζ its coordinates; we have $I(\zeta e) = 0$. Then, for all x's whose x_i's are not all equal ($I(x) \neq 0$), convexity gives:

$$I[\lambda \cdot (x - \zeta e) + \zeta e] \gtreqless \lambda \cdot I(x)$$

depending on $\lambda \gtreqless 1$, whatever ζ for I_l, for $\zeta > 0$ for I_r^a and σ, for $\zeta > -\xi$ for I_c. The inequalities are reversed if $\zeta < 0$ for I_r^a and σ, and for $\zeta < -\xi$ for I_c. In all cases, a possible ζ is $\bar{x}$ so that the transformation is mean preserving; this is a case of the former category for I_r^a, σ, and I_c. But $\zeta = \bar{x}$ is the only possible case for the constant-sum convexity case of I_r, I_l^r, I_c^r, $\sigma/\bar{x}$. The property then is

$$I[\lambda \cdot (x - \bar{x}e) + \bar{x}e] \gtreqless \lambda \cdot I(x)$$

depending on $\lambda \gtreqless 1$.

If, in particular, we take the origin as this point when possible, we find again some previous results: $I(\lambda x) \gtreqless I(x)$ depending on $\lambda \gtreqless 1$, for inequality measures I_l, I_c (for $\xi > 0$) and σ.

X. Union of Populations

X.a. *Required Properties*

If two countries which display the same degree of inequality unite to form a unique country, will we want the measure of inequality to indicate that inequality per person or per pound in the latter is the same one as in the constituting countries? No, because if each of the two initial countries has only one inhabitant, its income distribution will display no inequality, whereas the union country will be an unequal one if these two persons do not have the same income, and, more generally, if income is equally distributed in each of these two countries, but average incomes differ, inequality is inexistent in these initial countries but exists in the union. Thus, what we might want is total inequality in a union of populations to be larger, or not smaller, than the sum of the total inequalities in the constituting countries. That is, if k is an index representing a population, n_k and I_k^a the number of persons and the inequality per person in population k, $n = \sum n_k$ and I^a the total number of persons and inequality per person in the union,

$$nI^a \geqslant \sum_k n_k I_k^a.$$

This is equivalent to the relation between per person inequalities

$$I^a \geqslant \sum_k (n_k/n)\, I_k^a,$$

the right-hand side of which is an average of inequalities per person, appropriately weighted by the number of persons in each population. Calling $\bar{x}_k$, $X_k = n_k\bar{x}_k$ and I_k^r the average and total incomes and the inequality per pound in population k, and $\bar{x} = \sum (n_k/n)\, \bar{x}_k$, $X = n\bar{x} = \sum X_k$

and I^r the average and total incomes and the inequality per pound in the total population, these inequalities are also equivalent to the following one between inequalities per pound

$$I^r \geqslant \sum (X_k/X)\, I_k{}^r,$$

where the right-hand side is an average of per pound inequalities, appropriately weighted by the number of pounds (total incomes) in the populations.

The general result is that these relations hold for all the inequality measures we have considered until now, the equal sign holding if all the constituting populations have the same inequality and the same average income (but, for most measures, not only in this case).

Let us consider separately the measures with "welfare independence" property, and standard deviation with coefficient of variation.

X.b. *Independent Measures* [1, *Theorem* 21]

We have the following results, for inequality measures of the type $I^a = \bar{x} - \bar{\bar{x}}$, $I^r = I^a/\bar{x}$, nI^a.

With the independence property and the basic properties of inequality measures (zero at equality, positive out of it, impartiality-symmetry), the above relations hold for all populations and unions. They hold with equality (resp., strict inequality) if and only if all constituting populations have (resp., have not) the same equal equivalent income. We recall that equal equivalent incomes are the same if both average incomes and inequalities are (but this sufficient condition is not necessary).

And if, for an inequality measure which is independent, impartial and zero at equality, these relations strictly hold for unions of populations which do not all have the same equivalent income, this measure is positive out of equality (if these relations just hold for all unions of populations, the measure is nonnegative). And positivity out of equality is then equivalent to strict transfers principle and strict "isophily" (a small transfer from a richer person to a poorer one decreases inequality, a Lorenz curve nowhere lower and somewhere higher implies lower inequality for distributions of same total and average incomes) and even to constant-sum strict quasi-convexity of the measure in the x_i's (inequality of a distribution which is an average of several other ones of same total income is smaller than the largest of the latters' inequalities) (cf. Section XI below).[9]

[9] However, it is not equivalent to convexity, or even quasi-convexity, of the per person measure in the set of the x_i's, although both hold together in the special cases of independent measures studied above (I_l, I_c, $I_r{}^a$). These convexities imply nonnegativity of the measure, but the converse is not true (cf. Section XI below).

These results show that if all the constituting populations present the same degree of inequality, the global population will generally not have itself this inequality: it will generally be more unequal than each of its components. The exception, where global inequality is equal to the equal inequalities of the constituent populations, occurs if and only if the latter furthermore have equal average incomes (since this is then identical to saying that they have the same equal equivalent incomes). This neatly shows the double dependence of global inequality upon inequalities both within and between the constituting populations.

We note that H. Dalton's "principle of equiproportionate additions to persons" [8] is a special case of union of populations with the same equal equivalent income, since it comes to lumping together populations which duplicate all persons by the same numbers and thus have the same $\varphi(\bar{\bar{x}}) = (1/n) \sum \varphi(x_i)$. All these populations also have identical average incomes and inequalities.

The proof of these results is straightforward. We call J_k the set of indices i of persons in population k, $\bar{\bar{x}}^k$ the equal equivalent income of population k, $\bar{\bar{x}}$ the global equal equivalent income. We choose an increasing specification of function φ, which is always possible since a φ can be replaced by $a\varphi + b$ with a negative a; $I > 0$ out of equality is then equivalent to strict concavity of φ (cf. Section XI.d below). Then,

$$\varphi(\bar{\bar{x}}) = \frac{1}{n} \sum_i \varphi(x_i) = \frac{1}{n} \sum_k \sum_{i \in J_k} \varphi(x_i) = \sum_k \frac{n_k}{n} \varphi(\bar{\bar{x}}^k) \leqslant \varphi\left(\sum \frac{n_k}{n} \bar{\bar{x}}^k\right)$$

with equality if and only if all $\bar{\bar{x}}^k$'s are equal. This last inequality and precision is equivalent to the strict concavity of φ. And the comparison of the first and last terms is equivalent to $\bar{\bar{x}} \leqslant \sum (n_k/n) \bar{\bar{x}}^k$ with the same precision since φ is increasing. We also have $n\bar{x} = \sum n_k \bar{x}^k$. With $I_k{}^a = \bar{x}^k - \bar{\bar{x}}^k$ and $I^a = \bar{x} - \bar{\bar{x}}$, the above mentioned results follow.

X.c. *Standard Deviation and Coefficient of Variation*

As for the standard deviation σ, summing

$$\begin{aligned}(x_i - \bar{x})^2 &= (x_i - \bar{x}^k + \bar{x}^k - \bar{x})^2 \\ &= (x_i - \bar{x}^k)^2 + (\bar{x}^k - \bar{x})^2 + 2(x_i - \bar{x}^k)(\bar{x}^k - \bar{x})\end{aligned}$$

over $i \in J_k$, and then over k, gives

$$\sigma^2 = \sum_k (n_k/n)[\sigma_k{}^2 + (\bar{x}^k - \bar{x})^2]$$

which neatly shows the separate effects on total inequality of intra-population inequalities σ_k, and inter-populations inequality of average incomes $\sum (n_k/n)(\bar{x}^k - \bar{x})^2$. Thus,

$$\sigma^2 \geqslant \sum (n_k/n)\, \sigma_k^2 \geqslant \left(\sum (n_k/n)\, \sigma_k\right)^2$$

with equality if and only if $\bar{x}^k = \bar{x}$ for all k's for the first inequality and all σ_k's are equal for the second one. Therefore,

$$\sigma \geqslant \sum (n_k/n)\, \sigma_k$$

with equality if and only if all populations have both the same average income and the same standard deviation of incomes. This is also equivalent to the required relation between "total inequalities" $n\sigma$, and between the "per pound" coefficient of variation:

$$\sigma/\bar{x} \geqslant \sum (X_k/X)\, \sigma_k/\bar{x}^k.$$

These properties are thus exactly the same ones as for the other measures under consideration, the only difference being that it is now both necessary and sufficient that both average incomes and inequalities be the same for all populations in order that the equality sign holds in the relations.

XI. General Structures of Inequality Measures

XI.a. *The Problem*

We started from specific measures of inequality, then considered measures of more and more general form, and economic properties which belong to still much more general classes of measures (such as the economic meanings of intensivity, equal increase in all incomes, subadditivity, convexity, quasi-convexity, Schur-convexity, etc.). We consider now these more general structures, the economic consequences of which have already been discussed *a propos* the properties of more specific measures exhibiting them.

We shall consider properties pertaining to the distribution $x = \{x_i\}$ $(i = 1,..., n)$, its average income $\bar{x} = \sum (x_i/n)$, the "evaluation function" $V(x)$, the equal equivalent income $\bar{\bar{x}}\,(V(x) = V(\bar{\bar{x}}e))$, and the measures of per person and per pound inequality $I^a = \bar{x} - \bar{\bar{x}}$ and $I^r = I^a/\bar{x}$; when a property holds for both I^a and I^r, we shall mention it for "I." When it holds only at given $\bar{x}$, we shall add the adjective "constant-sum."

The subject matter will be properties of the functions of x V, $\bar{\bar{x}}$, I^a, and I^r; the topic will be both to relate the corresponding properties of these functions, and more interestingly to establish the general relations between these different properties. V can have only ordinal properties. The properties of $\bar{\bar{x}}$ will be used, and they are quite obvious: as a function of x it is increasing, symmetrical, intensive, increased by some amount if all x_i's are increased by this amount, at the limit of the "principle of diminishing transfers," and altogether weakly concave, convex, quasi-concave, quasi-convex, Schur-concave, Schur-convex.

$I = 0$ if all x_i's are equal by definition of $\bar{\bar{x}}$ and I. Symmetry of V, $\bar{\bar{x}}$, and I go together ("impartiality"). We assume V is an increasing function of the x_i's ("benevolence"), so that $\bar{\bar{x}}$ is well defined and has the same property.

XI.b. *The Various Convexities and Concavities*

We first point out the general relations between the various kinds of convexities and concavities. They will be applied to the functions of x V, $\bar{\bar{x}}$, I^a, I^r. The following sentence implies four sentences: we can replace "concavity" by "convexity," and in each case this word can mean either the "strict" or the "weak" property. *Concavity implies constant-sum concavity and quasi-concavity; either of the latter two implies constant-sum quasi-concavity; constant-sum quasi-concavity plus symmetry*[10] *imply Schur-concavity which implies rectifiance* ("transfers principle"). These relations apart from the last two ones, result from the fact that the intersection of two convex sets is convex.

To prove the last but one, we call F a constant-sum quasi-concave and symmetrical function of x, B a bistochastic matrix (an $n \times n$ square matrix whose entries b_{ij} are nonnegative and satisfy $\sum_i b_{ij} = \sum_j b_{ij} = 1$ for all i and j's) P^π the permutation matrix of index π (P^π is a B whose entries are all 0 or 1), λ_π weights ($\lambda_\pi \geqslant 0$ for all π's, $\sum \lambda_\pi = 1$). All $P^\pi x$ have the same average $\bar{x}$, and $F(P^\pi x) = F(x)$ is the definition of F's symmetry. Birkhoff's theorem [7] says that B can be written as $\sum \lambda_\pi P^\pi$. Ostrowski's theorem says that rectifiance ("transfers principle") and symmetry for a weakly Schur-concave F are equivalent to $F(Bx) \geqslant F(x)$ for all B's. And F's constant-sum weak quasi-concavity implies $F[\sum \lambda_\pi(P^\pi x)] \geqslant F(x)$. The result is then proved by

$$F(Bx) = F\left(\sum \lambda_\pi P^\pi x\right) \geqslant F(x)$$

for the weak properties. For the strict ones, we notice that strict Schur-concavity of x is $F(Bx) > F(x)$ for all x's and B's such that Bx is not a $P^\pi x$. And if $Bx = \sum \lambda_\pi P^\pi x$ is not a $P^\pi x$, the $P^\pi x$'s for the nonzero λ_π's

[10] Constant-sum symmetry would be enough. It is implied by symmetry.

are not all equal. Then, constant-sum strict quasi-concavity of F implies $F[\sum \lambda_\pi(P^\pi x)] > F(x)$. The result is then proved by

$$F(Bx) = F\left(\sum \lambda_\pi P^\pi x\right) > F(x)$$

for Bx not a $P^\pi x$. Finally, reversing the inequalities proves the relation for the convexities.

Now, we also have the following property for an inequality measure $I(x)$ however defined (not necessarily as I^a or I^r above). *If $I(x)$ is zero at equality and symmetrical ("impartial"), and if it is either rectifiant (thus, Schur-convex), or constant-sum quasi-convex, or constant-sum convex, or quasi-convex, it is positive out of equality* for the strict forms, and nonnegative for the weak ones (thus a symmetrical, zero at equality, convex $I(x)$ is nonnegative, strict convexity being excluded by the zero at equality condition). All these properties result from the one with Schur-convexity. Strict Schur-convexity means $I(Bx) < I(x)$ if B is a bistochastic matrix and Bx's coordinates are not a permutation of x's. If we take a B whose entries are all $1/n$, the latter condition just requires that all x_i's are not equal, we have $Bx = \bar{x}e$, and thus

$$I(x) > I(Bx) = I(\bar{x}e) = 0.$$

For the weak forms, we replace $>$ by $\geqslant$.

Coming back to I functions which are I^a or I^r defined as above from $V(x)$ and $\bar{\bar{x}}$, these definitions, any definition of Schur-convexity or -concavity and the above property show that *Schur-concavity of V, of $\bar{\bar{x}}$, and Schur-convexity of $I(I^a$ and $I^r)$ are equivalent, and they imply $\bar{\bar{x}} < \bar{x}$, i.e., $I > 0$, out of equality* (for the strict forms, the same holding for the weak ones with replacement of $>$ by $\geqslant$ in the properties and proofs).

But we also remark that *constant-sum quasi-concavity of V and $\bar{\bar{x}}$ are equivalent to constant-sum quasi-convexity of I (I^a and I^r)*, with correspondence between the strict and weak forms. This is so because

$$\frac{\bar{x}^1 + \bar{x}^2}{2} - \bar{\bar{x}}\left(\frac{x^1 + x^2}{2}\right) < \operatorname{Max}(\bar{x}^1 - \bar{\bar{x}}^1, \bar{x}^2 - \bar{\bar{x}}^2),$$

which is equivalent to

$$\bar{\bar{x}}\left(\frac{x^1 + x^2}{2}\right) > \operatorname{Min}\left(\frac{\bar{x}^2 - \bar{x}^1}{2} + \bar{\bar{x}}^1, \frac{\bar{x}^1 - \bar{x}^2}{2} + \bar{\bar{x}}^2\right),$$

is equivalent to

$$\bar{\bar{x}}\left(\frac{x^1 + x^2}{2}\right) > \operatorname{Min}(\bar{\bar{x}}^1, \bar{\bar{x}}^2)$$

if $\bar{x}^1 = \bar{x}^2$. The strict form then results by considering $x^1 \neq x^2$, and the weak form by replacing $>$ by $\geqslant$. Of course, the relation for $I^r = I^a/\bar{x}$ is the same as that for I^a.

XI.c. *Homotheticities and Translatednesses*

A function $F(x)$ is said to be *homothetic* when its hypersurfaces $F(x) =$ constant in x space are transformed into each other by an homothetic transformation the center of which is the origin. We shall say it is X-homothetic when this property holds with the only difference that the center is point X in x space. The former homotheticity thus is 0-*homotheticity*. When such a center goes to infinity, the relation between the hypersurfaces in x space $F(x) =$ constant becomes that they are translated from each other in a given direction. When this direction is that of a vector X, we shall say that $F(x)$ is a X-*translated* function. We now have the following properties.

V is homothetic, $\bar{\bar{x}}$ and I^a are linearly homogeneous, I^r is intensive, are equivalent properties. The relation between $\bar{\bar{x}}$, I^a, I^r just results from the definitions of I. On the other hand, we know that V is homothetic if and only if it has a linearly homogeneous specification, i.e., there exists increasing functions Φ and W such that $V(x) \equiv \Phi[W(x)]$ and W is linearly homogeneous. This property and the definition of $\bar{\bar{x}}$ then give $W(x) = W(\bar{\bar{x}}e) = \bar{\bar{x}} \cdot W(e)$ which shows both the relation between W and $\bar{\bar{x}}$ and the linear homogeneity of $\bar{\bar{x}}$.

V is e-translated, an equal variation in all incomes affects $\bar{\bar{x}}$ in the same way, and it does not change I^a, are equivalent properties. The relation between $\bar{\bar{x}}$ and I^a is obvious. The one between V and $\bar{\bar{x}}$ is deduced from the previous paragraphs by a change of variables which consists in replacing x_i by a^{x_i} and $\bar{\bar{x}}$ by $a^{\bar{\bar{x}}}$ where a is any positive constant.[11]

Intermediate cases between these two happen when V is $-e\xi$-homothetic (ξ is a scalar), which is equivalent to $\bar{\bar{x}} + \xi$ being multiplied by λ when all $x_i + \xi$'s are, as is seen by the change of variables from x_i to $x_i + \xi$ and $\bar{\bar{x}}$ to $\bar{\bar{x}} + \xi$ in the homothetic case.

Now, a marriage between these properties of homotheticity or translatedness on the one hand, and quasi-concavity on the other hand, gives an interesting offspring: convexity of I. More precisely: *if V is both X-homothetic or X-translated and quasi-concave, $\bar{\bar{x}}$ is concave and I^a is convex*, I^r therefore is constant-sum convex; and if V's quasi-concavity is strict, $\bar{\bar{x}}$'s concavity and I^a's convexity are strict out of lines issued from

[11] We similarly find the result that a function $F(x)$ is X-translated if and only if there exists functions Φ and Ψ such that $F(x) \equiv \Phi[\Psi(x)]$ with Ψ such that $\Psi(x + \lambda X) = \Psi(x) + k\lambda$ (where k is a constant).

point X (for the homotheticity case) or parallel to vector X (for the translatedness case). In particular, X can be 0 (V homothetic) or $-\xi e$ for the homotheticity cases, or e for the translatedness case. To prove these results, apart from obvious implications in them, we can change variables to bring all the cases back to the one of a homothetic V, and then prove the only nontrivial relation by recalling that homothetic quasi-concave V implies linearly homogeneous quasi-concave $\bar{\bar{x}}$, which in turn is concave from [9, Chap. VIII, Sect. 8, Theorem 3].

This specific case is of special interest since *V homothetic quasi-concave, i.e., $\bar{\bar{x}}$ linearly homogeneous and concave, and thus I^a linearly homogeneous and convex, occur if and only if I^a is subadditive.*

XI.d. *Independence and Convexities*

There still is another property, which we used at length in previous sections: "independence," or additive separability of V. Our starting point was that this structure, plus impartiality-symmetry and X-homotheticity or X-translatedness, give the specific forms of $\bar{\bar{x}}$ and I previously discussed (the symmetry imposes X to be parallel to e). But "independence" also interfers interestingly with the various kinds of convexity.

We consider again $I(x)$ defined as $I^a = \bar{x} - \bar{\bar{x}}$ or $I^r = I^a/\bar{x}$, with $\bar{\bar{x}}$ defined form $V(x) = V(\bar{\bar{x}}e)$ with V an increasing function ("benevolence"). "Independence" means that there exists $n + 1$ functions of one variable Φ and φ^i ($i = 1,..., n$) such that $V \equiv \Phi[\sum \varphi^i(x_i)]$. Since V is increasing in all the relevant domains, Φ and the φ^i's must be monotonic with the same sense of variation, and we can always assume they are increasing functions, which we do. V is, furthermore, "impartial" (symmetrical) if and only if we can take the same function φ for all the φ^i's. This is the case if $\varphi^i(y) \equiv \varphi(y) + c^i$ with constant c^i for all i's since, then, calling $\Psi(z) \equiv \Phi(z + \sum c^i)$ makes $\Phi[\sum \varphi^i(x_i)]$ identical to $\Psi[\sum \varphi(x_i)]$. Reciprocally, if an "independent" V is "impartial," $\varphi^i(y) \equiv \varphi(y) + c^i$ for all i's since this symmetry implies $\varphi^j(y_1) + \varphi^k(y_2) \equiv \varphi^j(y_2) + \varphi^k(y_1)$, i.e., $\varphi^i(y_1) - \varphi^i(y_2)$ is the same for all i's.

The general results are the following.

Assuming the respective differentiabilities when required, with "independence," the properties within each of the two following groups are equivalent to each other, with correspondence between weak and strict forms (and the second group implies the first one's weak form):

(1°) *"Rectifiance" ("transfers principle"); I is nonnegative (positive out of equality for the strict form); I is Schur-convex, or V or $\bar{\bar{x}}$ is Schur-concave; "impartiality" plus either of the following properties: I is constant-sum quasi-convex, V or $\bar{\bar{x}}$ is constant-sum quasi-concave or quasi-concave,*

$\sum \varphi(x_i)$ is Schur-concave or constant-sum quasi-concave or quasi-concave or constant-sum concave or concave, φ is concave.

(2°) *An impartial per person inequality measure I^a is convex; φ'/φ'' is convex (i.e., $\varphi'\varphi'''/\varphi''^2$ is decreasing).*

One of these results is that, with "independence," "rectifiance" implies "impartiality," and thus suffices by itself to define Schur-convexity of I or Schur-concavity of V. This is so because the rectifiance conditions with two "incomes" y_1 and y_2 such that $y_1 < y_2$ imply both $\varphi^{i\prime}(y_1) - \varphi^{j\prime}(y_2) > 0$ and $\varphi^{i\prime}(y_2) - \varphi^{j\prime}(y_1) < 0$ (resp. $\geqslant$ and $\leqslant$ for the weak forms), and thus, when y_1 and y_2 tend to the same value y, $\varphi^{i\prime}(y) = \varphi^{j\prime}(y)$, for all admissible y. Integrating shows that the φ^i's can all be written as $\varphi^i(y) = \varphi(y) + c_i$ where c_i is a constant, which proves the symmetry of $\sum \varphi^i$ and the "impartiality." Rectifiance then means that φ' is a decreasing (resp. nonincreasing) function, i.e., that φ is a concave function (with correspondence between strict and weak forms).

Now, concavity of φ is also equivalent to concavity of $\sum \varphi(x_i)$ with correspondence between weak and strict forms: x^1 and x^2 being two different x's and λ a scalar such that $0 < \lambda < 1$,

$$\lambda\varphi(x_i^1) + (1-\lambda)\,\varphi(x_i^2) \geqslant \varphi[\lambda x_i^1 + (1-\lambda)\,x_i^2]$$

for all i's implies

$$\lambda \sum \varphi(x_i^1) + (1-\lambda) \sum \varphi(x_i^2) \geqslant \sum \varphi[\lambda x_i^1 + (1-\lambda)\,x_i^2],$$

and if φ's concavity is strict, the first inequality is strict for at least one i (since $x_i^1 \neq x_i^2$ for at least one i), and the second inequality is strict; conversely, if $\sum \varphi(x_i)$ is concave, choosing x's each with equal x_i's shows that $\varphi(y)$ has the same concavity (strict or weak).

$\sum \varphi(x_i)$'s concavity in turn implies its quasi-concavity and its constant-sum concavity, either of which implies its constant-sum quasi-concavity, which, with its symmetry, implies its Schur-concavity, and thus its rectifiance, which is equivalent to φ's concavity and thus to $\sum \varphi(x_i)$'s, with correspondence between strict and weak forms. All these properties are thus equivalent to each other. Besides, the ordinal properties of quasi-concavity, constant-sum quasi-concavity and Schur-concavity are the same ones for $\sum \varphi(x_i)$, $V(x)$, and $\bar{x}(x)$. And V and $\bar{x}$'s constant-sum quasi-concavity and Schur-concavity are respectively equivalent to constant-sum quasi-convexity and Schur-convexity of I (I^a and I^r), with correspondence between strict and weak forms.

Schur-convexity of I—and thus, with "independence," mere rectifiance (transfers principle)—was seen to imply $I \geqslant 0$ for the weak forms and $I > 0$ out of equality for the strict ones. The converse is obviously also true with "independence" and "impartiality," since $\varphi(\bar{x}) \geqslant (1/n) \sum \varphi(x_i) \equiv \varphi(\bar{\bar{x}})$ for all admissible x's is equivalent both to $\bar{x} \geqslant \bar{\bar{x}}$, i.e., $I \geqslant 0$, for all admissible x's and to φ's weak concavity, and $\varphi(\bar{x}) > (1/n) \sum \varphi(x_i) \equiv \varphi(\bar{\bar{x}})$ for all admissible x's such that the x_i's are not all equal is equivalent both to $\bar{x} > \bar{\bar{x}}$, i.e., $I > 0$, out of equality and to φ's strict concavity.[12]

But, furthermore, with "independence," I's nonnegativity or positivity out of equality implies "impartiality" and is thus by itself equivalent to the other mentioned properties. To show this, let us consider a small transfer of ϵ from i to j and from j to i, starting from a situation of equality where $x_k = y$ for all k's. If k receives ϵ,

$$d\varphi^k(y) = \epsilon\varphi^{k\prime}(y) + (\epsilon^2/2)[\varphi^{k\prime\prime}(y) + 0_1{}^k(\epsilon)]$$

where $0_1{}^k(\epsilon)$ tends to zero with ϵ. If it is taken ϵ (i.e., if it receives $-\epsilon$),

$$d\varphi^k(y) = -\epsilon\varphi^{k\prime}(y) + (\epsilon^2/2)[\varphi^{k\prime\prime}(y) + 0_2{}^k(\epsilon)]$$

where $0_2{}^k(\epsilon)$ tends to zero with ϵ. Thus, the effect on $\sum \varphi^i(x_i)$ of a small transfer ϵ from i to j is

$$d\left[\sum \varphi^k(x_k)\right] = \epsilon \cdot [\varphi^{i\prime}(y) - \varphi^{j\prime}(y)] + (\epsilon^2/2)[\varphi^{i\prime\prime}(y) + \varphi^{j\prime\prime}(y) + 0_1(\epsilon)]$$

where $0_1(\epsilon)$ tends to zero with ϵ, and the effect of the reverse transfer is

$$d\left[\sum \varphi^k(x_k)\right] = -\epsilon \cdot [\varphi^{i\prime}(y) - \varphi^{j\prime}(y)] + (\epsilon^2/2)[\varphi^{i\prime\prime}(y) + \varphi^{j\prime\prime}(y) + 0_2(\epsilon)]$$

where $0_2(\epsilon)$ tends to zero with ϵ. These two operations do not change $\bar{x} = y$. If $I \geqslant 0$ (resp. >0 out of equality), in the new situations we must have $\bar{\bar{x}} \leqslant y$ (resp. $<$) and thus $\sum \varphi^k(x_k) \equiv \sum \varphi^k(\bar{\bar{x}}) \leqslant \sum \varphi^k(y)$ (resp. $<$) by definition of $\bar{\bar{x}}$ and since the φ^k's are increasing functions. That is, the two written variations of $\sum \varphi^k$ must be nonpositive (resp. negative). When ϵ is small enough, this requires both $\varphi^{i\prime}(y) \leqslant \varphi^{j\prime}(y)$ and $\varphi^{i\prime}(y) \geqslant \varphi^{j\prime}(y)$, i.e., $\varphi^{i\prime}(y) = \varphi^{j\prime}(y)$. Integrating shows that all the φ^k's are of the form $\varphi^k(y) = \varphi(y) + c_k$ where c_k is a constant, which means that $\sum \varphi^k$, V, $\bar{\bar{x}}$, I are symmetric functions. Given this result, the nonpositivity (resp. negativity) of the above differentials when ϵ is small enough becomes equivalent to $\varphi'' \leqslant 0$ (resp. $\varphi'' < 0$ almost everywhere) i.e., to concavity of φ (weak or strict).

[12] More generally, $\Sigma\varphi^i(x_i)$ is strictly (resp. weakly) concave if and only if each φ^i is strictly (resp. weakly) concave. The sufficiency is proved in the same way by replacing $\varphi(x_i)$ by $\varphi^i(x_i)$, and the necessity is proved by considering x's which differ by only one of their coordinates.

Concavity of $\sum \varphi(x_i)$ is one of these equivalent properties. If $\bar{\bar{x}}$ were concave, which is equivalent to $I^a = \bar{x} - \bar{\bar{x}}$ being convex (and implies a constant-sum convex I^r), $\bar{\bar{x}}$ and V would be quasi-concave and all the other mentioned equivalent properties would follow. Now, the specific $\varphi(y)$'s by which we begun the study, $y^{1-\epsilon}$, $-e^{-\alpha y}$, $(y + \xi)^{1-\epsilon}$, were shown to give convex I^a's for ϵ and α which give nonnegative (resp. positive out of equality) or Schur-convex I^a's. But this is not true for all φ's. Rather, by imitating the proof of [5, Theorem 106][13] we can show that an "independent impartial" inequality measure I^a with $\varphi'' < 0$ is convex if and only if φ'/φ'' is convex (i.e., $\varphi'\varphi'''/\varphi''^2$ is decreasing). Recalling that I_c and its special and limit cases $I_r{}^a$ and I_l constitute the class of independent inequality measures with linear φ'/φ'', we see that they fall in this category and we find again that they are convex.

XI.e. *Generalizations of the "Transfers Principles," "Rectifiance," Schur-Convexity, "Isophily"*

In this section, we relate very briefly and without proof nor precision the relation between some of the above results and other meaningful properties of inequality measures (cf. [11–13]).

In Section VII, we have seen that $I^a = \bar{\bar{x}} - \varphi^{-1}[(1/n) \sum \varphi(x_i)]$ with concave φ, or $I^r = I^a/\bar{x}$, always satisfies the "principle of diminishing transfers" if and only if φ', φ'', and φ''' alternate in sign, if these derivatives exist. The satisfaction of this "principle" for all admissible x_i, x_j, x_k, x_l, for a Schur-convex $I(x)$, constitutes a generalization of Schur-convexity which we may call 2-rectifiance (or rectifiance of order 2). The precedent property is the form it takes for "independent" measures and existence of these derivatives. If we call $z_i = \sum_{j=1}^{i} y_j$, and $V(x)$ a social evaluation function (increasing function of $\bar{\bar{x}} = \bar{x} - I$ or $\bar{x} \cdot (1 - I)$), I is 2-rectifiant if and only if $V(x^1) > V(x^2)$ for all x^1, x^2 such that $z^1 \geqslant z^2$; and $z^1 \geqslant z^2$ if and only if $V(x^1) > V(x^2)$ for all 2-rectifiant measures. This latter proposition is also true if we restrict ourselves to 2-rectifiant and independent measures. We can define, with similar relations and theorems, further degrees of principles of transfers or rectifiance (corresponding to further derivatives of φ alternating in sign in the "independent" case when these derivatives exist), and of "isophilies" or dominance of integrals. Each next degree represents one step more in egalitarianism. The ultimate degree is $\bar{\bar{x}} = \mathrm{Min}_i\, x_i$ with "impartiality," degree one is Pigou and Dalton's

[13] Similar properties have been used in justice theory in [1, Theorem 19], and generalizations of this theorem were used in the theory of choice under uncertainty in [10, Chap. VIII], to study the convexity properties of risk- and insurance-premia, which are the risk-theory meaning of I^a. We recall that we require here the property to hold for all n's (cf. Section II.c.).

principle of transfers and [1]'s "rectifiance" of functions and "isophily" on distributions, and degree zero just is "benevolence," i.e., V is an increasing function, and [3]'s "fundamental dominance" for distributions.

Another generalization is to consider all these properties for x_i's which are no more unidimensional magnitudes but multidimensional vectors [12, 13].

References

1. 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, ("Economie Publique," CNRS, Paris, 1968, and "Public Economics," Macmillan, New York).
2. A. B. Atkinson, On the measurement of inequality, *J. Economic Theory* **2** (1970).
3. S.-Ch. Kolm, "Justice et Équité," CNRS, Paris, 1972.
4. A. Ostrowski, Sur quelques applications des fonctions convexes et concaves au sens de I. Schur, *J. Math. Pures Appl.* **31** (1952).
5. G. H. Hardy, J.E. Littlewood, and G. Pólya, "Inequalities," 2nded. Cambridge Univ. Press, Cambridge, 1952–1964.
6. J. Karamata, Sur une inégalité relative aux fonctions convexes, Publications Mathématiques de l'Université de Belgrade **1** (1932).
7. G. Birkhoff, "Tres observaciones sobre el algebra lineal, *Rev. Universidad Nacional de Tucuman, Serie A*, **5** (1946), 147–151.
8. C. Berge, "Espaces topologiques et fonctions multivoques," Dunod, Paris, 1959.
9. H. Dalton, The measurement of the inequality of incomes, *Economic Journal* **30** (1920), 348–361.
10. S.-Ch. Kolm, "Les choix financiers et monétaires (théorie et technique modernes)," Dunod, Paris, 1966.
11. S.-Ch. Kolm, Rectifiances et dominances intégrales de tous degrés, mimeographed CEPREMAP, Paris, 1974.
12. S.-Ch. Kolm, More equal distributions of bundles of commodities, mimeographed, CEPREMAP, Paris, 1973.
13. S.-Ch. Kolm, Multidimensional egalitarianisms, mimeographed, CEPREMAP, Paris, 1975.