Journal of Economic Theory **91**, 199–222 (2000)
doi:10.1006/jeth.1999.2565, available online at http://www.idealibrary.com on IDEAL®

Path Independent Inequality Measures[1]

James E. Foster

Department of Economics, Vanderbilt University, Nashville, Tennessee 37235
james.e.foster@vanderbilt.edu

and

Artyom A. Shneyerov

Department of Managerial Economics and Decision Sciences,
Kellogg Graduate School of Management, Northwestern University,
Evanston, Illinois 60208
a-shneyerov@nwu.edu

Received April 28, 1998; final version received June 30, 1999

This paper explores a natural decomposition property motivated by Shorrocks (1980, *Econometrica* **48**, 613–625) and Anand (1983, "Inequality and poverty in Malaysia," Oxford University Press) that we call *path independent decomposability*. Between-group inequality is found by applying the inequality measure to the *smoothed* distribution, which replaces each income in a subgroup with its representative income. Within-group inequality is the measure applied to the *standardized* distribution, which rescales subgroup distributions to a common representative income level. *Path independence* requires overall inequality to be the sum of these two terms. We derive the associated class of relative inequality measures—a single parameter family containing both the second Theil measure (the mean logarithmic deviation) and the variance of logarithms. *Journal of Economic Literature* Classification Numbers: C43, D31, D63, 015. © 2000 Academic Press

Key Words: inequality measures; decomposition; Theil measure; variance of logarithms.

1. INTRODUCTION

What share of overall income inequality can be attributed to differences *between* population subgroups (defined by age, race, gender, or some other

[1] We thank Sudhir Anand, Andy Daughety, Peter Lambert, Efe Ok, Amartya Sen, Tony Shorrocks, and John Weymark for helpful comments and encouragement and J. S. Butler for kindly providing the dataset used in an illustration. Support from the John D. and Catherine T. MacArthur Foundation through its Network on Inequality and Poverty in Broader Perspective is gratefully acknowledged.

0022-0531/00 $35.00

meaningful characteristic)? What share is due to inequality *within* these subgroups? Questions of this type require inequality measures that can be sensibly decomposed into terms representing between-group and within-group inequality.

Shorrocks' [28, 29] characterization results provide a strong justification for using the generalized entropy measures (including the Theil measures and the squared coefficient of variation) to analyze inequality by population subgroups.[2] The generalized entropy class has an intuitive decomposition based on a hypothetical *smoothed* distribution in which each person's income is replaced with the mean income of the subgroup to which he or she belongs. This smoothing process removes all inequality within the subgroups; consequently, between-group inequality can be defined as the inequality level of the smoothed distribution, while within-group inequality is the difference between the total and the between-group terms.[3] So, for example, if the population were partitioned by age groups, the between-group term would indicate the inequality arising from age variations in income, while the within-group term would indicate the portion of inequality attributable to other factors.

As noted by Shorrocks [28] and Anand [3], there is another, equally plausible decomposition based on an alternative hypothetical distribution. The *standardized* distribution is obtained by proportionally scaling each subgroup distribution until it has the same mean as the overall distribution.[4] Standardization suppresses between-group inequality while leaving group inequality levels unaltered. Thus, it is reasonable to define the within-group term as the inequality level of the standardized distribution and to take between-group inequality as the residual.

Shorrocks [28] has shown that the generalized entropy measures (apart from Theil's second measure) yield different breakdowns for the two approaches. So, for example, standardizing the distribution to remove age–income variations leads to a different evaluation of life cycle effects on inequality than would be obtained via the smoothed distribution. In response to this observation, Shorrocks concludes that the second Theil measure is the "most satisfactory of the decomposable measures" as it renders the same decomposition independent of the path chosen. Analogous arguments on behalf of this property (which we will call *path*

[2] See also Bourguignon [8], Cowell [9], and Cowell and Kuga [12] for discussions relating to this family, and Foster [14] and Shorrocks [30] for affiliated characterizations. For empirical applications of generalized entropy measures, see for example Shorrocks and Mookherjee [32], Cowell [10], Tsakloglou [33], Jenkins [20], and Cowell and Jenkins [11].

[3] The within-group term can further be expressed as a weighted sum of the subgroup inequality levels. See Section 3 below.

[4] This is the terminology of Love and Wolfson [24] who also explored this approach to decomposition.

independent decomposability) have been presented by Anand [3], who identifies yet another measure that delivers consistent inequality decompositions: the variance of logarithms. But rather than using the mean in its within and between-group terms, the variance of logarithms uses the *geometric mean*—an alternative *representative income* that places more weight on low incomes. Blackorby *et al.* [7] proceed even further down this line and argue in favor of using more general forms of representative incomes in decomposition analysis, although their ethical decomposition has a rather different motivation—and a different formula—than the present one.

In this paper we adopt a broadened form of path independence that employs general representative income functions in its decomposition terms. Between-group inequality is defined as the inequality level of the smoothed distribution in which each person's income is replaced with the representative income of his or her subgroup. Within-group inequality is the inequality level of the standardized distribution in which each subgroup distribution is rescaled to the overall representative income level. An inequality measure has a path independent decomposition if overall inequality is the sum of these within-group and between-group terms.

This paper offers a complete characterization of path independent decomposable inequality measures. In particular, our results show that: (i) path independent inequality measures have within-group terms that can be expressed as population-share-weighted sums of the subgroup inequalities; (ii) the only representative income functions that are consistent with path independent measures are the means of order q (defined in Section 2 below); and (iii) for each q, there is a unique (up to a scalar multiple) inequality measure that has a path independent decomposition. Consequently, path independence characterizes a new, single-parameter class I_q, containing the variance of logarithms, where $q = 0$, the second Theil measure, where $q = 1$, and a host of others. We offer interpretations of the measures and show that they are closely related to the Kolm–Pollak measures of absolute inequality. In addition, we explore the properties satisfied by the various members of I_q and provide a useful interpretation of the parameter q. We conclude with a discussion of possible empirical benefits of the new class of measures.

2. DEFINITIONS

We will be concerned with income distributions of the form $x = (x_1, ..., x_n)$, where x_i is the (positive) income of the ith person, and $n = n(x)$ is the population size; consequently, $\mathscr{D} = \bigcup_{n=1}^{\infty} R^n_{++}$ is the domain under consideration. We denote the unit vector of dimension n by $u_n = (1, ..., 1)$

and call $x \in \mathscr{D}$ *completely equal* if it is a scalar multiple of some u_n. The geometric and arithmetic means of a given distribution x will be denoted by $g = g(x)$ and $\mu = \mu(x)$, respectively. We will use $\mu_q = \mu_q(x)$ to represent the *mean of order q* (for $q \in R$), where in particular μ_1 is the arithmetic mean, μ_0 is the geometric mean, and μ_{-1} is the harmonic mean of x.[5] It will also prove useful to define a number of basic transformations of a given distribution. We say that $x \in \mathscr{D}$ is obtained from $y \in \mathscr{D}$ by a *permutation* if $x = Py$ for some permutation matrix P of appropriate dimension, a *replication* if $x = (y, ..., y)$, a *proportional change* if $x = \alpha y$ for some $\alpha > 0$, or a *simple increment* if $x_i > y_i$ for some i and $x_j = y_j$ for all $j \neq i$. Permutations reorder the incomes, replications have several copies of each original income, and proportional changes rescale all incomes by the same factor, while simple increments increase one income and leave the rest unchanged.

For the purposes of this paper, an *inequality measure* is a function $I: \mathscr{D} \to R$ satisfying the following properties.[6] (1) *Symmetry*: If x is obtained from y by a permutation of incomes, then $I(x) = I(y)$; (2) *Replication invariance*: If x is obtained from y by a replication of incomes, then $I(x) = I(y)$; (3) *Scale invariance*: If x is obtained from y by a proportional change in incomes, then $I(x) = I(y)$; (4) *Normalization*: If x is a completely equal distribution, then $I(x) = 0$ (and if x is not completely equal, then $I(x) > 0$); and (5) *Continuity*: I is continuous on each n-person slice of $\mathscr{D}$.

The first three of these properties ensure that I is a *relative* measure, in the sense that it can only depend on the relative frequencies (as required by symmetry and replication invariance) of the relative incomes (according to scale invariance). Normalization fixes the value at zero when the distribution is completely equal and ensures that the measure has the correct orientation when inequality is present. Continuity is a commonly used regularity assumption. Note that we have not assumed the transfer principle (or, what is equivalent in this context, Lorenz consistency) since it is not needed for our characterization results.[7] (See, however, Section 5 below.)

Our approach to inequality decomposition uses a *representative income function* $r: \mathscr{D} \to R_{++}$ to reflect the average prosperity level or income

[5] I.e., $\mu_q(x) = ((1/n) \sum_{i=1}^{n} x_i^q)^{1/q}$ for $q \neq 0$ and $\mu_0(x) = g(x) = \prod_{i=1}^{n} x_i^{1/n}$. See, for example, Hardy *et al.* [19], pp. 12–15.

[6] This is the same collection of properties used by Shorrocks [29].

[7] The transfer principle requires a (regressive) transfer from poor to rich to increase inequality. Lorenz consistency requires an inequality measure to follow the Lorenz criterion when it applies. In the presence of symmetry, replication invariance, and scale invariance, the two properties are equivalent. See Foster [15].

standard $r(x)$ associated with a given distribution x. While most existing decomposition formulas use the arithmetic mean in this role, our formulation broadens the set of possibilities to include the geometric mean used by the variance of logarithms and the *equally distributed equivalent* (or *ede*) income functions employed by Blackorby *et al.* [7], as well as many other specifications. The only restrictions placed on r are the following. (1) *Symmetry*: If x is obtained from y by a permutation of incomes, then $r(x) = r(y)$; (2) *Replication invariance*: If x is obtained from y by a replication of incomes, then $r(x) = r(y)$; (3) *Monotonicity*: If x is obtained from y by a simple increment, then $r(x) > r(y)$; (4) *Linear homogeneity*: If $x = \alpha y$ for $\alpha > 0$, then $r(x) = \alpha r(y)$; (5) *Normalization*: If x is completely equal then $r(x) = x_1$; and (6) *Continuity*: r is a continuous function on each n-person slice of $\mathscr{D}$.

Symmetry and replication invariance are quite natural in this context (and in fact can be dropped without loss in the theorems that follow). Monotonicity requires the income standard to respond positively to increases in an individual's income. Linear homogeneity implies that a doubling of all incomes doubles the representative income. This plays an important role in the standardization process which underlies our definition of within-group inequality. Normalization requires that when all incomes are the same, the representative income must be the commonly held level of income. This ensures that the smoothing process, used in the definition of between-group inequality, leaves completely equal distributions unchanged. Continuity is once again assumed as a regularity condition.

This definition clearly goes far beyond the arithmetic (or geometric) mean typically employed in inequality decompositions. Prominent examples of representative income functions can be found among the equally distributed equivalent income functions (mentioned above) that have been extensively used in welfare-based inequality measurement.[8] The separable ede's employed by Atkinson [4]—which turn out to be the q-order means (μ_q) for $q \leqslant 1$—and the nonseparable generalized Gini ede's given by Donaldson and Weymark [13] are two well-known specifications. We should emphasize, however, that our notion of a representative income function need not be linked to any concept of welfare, nor must it have the curvature properties usually required of welfare-based standards. The q-order means in the range $q > 1$ are good examples of representative income functions having no such welfare interpretation due to their

[8] This is the terminology of Atkinson [4]. See also Kolm [21], Blackorby and Donaldson [5, 6], and Sen [27].

curvature.[9] On the other hand, the limiting cases of the maximum income (obtained when q tends to ∞) and the minimum income (when q tends to $-\infty$) do not qualify as representative income functions since they violate the monotonicity requirement, and the same is true of the other quantile-based standards, including the commonly used median income.

3. PATH INDEPENDENT INEQUALITY MEASURES

We now turn to the motivating property of this paper. Consider a distribution $x \in \mathscr{D}$ partitioned into $m \geqslant 2$ subgroup distributions $x^j \in \mathscr{D}$ for $j = 1, ..., m$, so that x may be written as $(x^1, ..., x^m)$. Let r be a representative income function. The *smoothed distribution* associated with $(x^1, ..., x^m)$ and r is defined by

$$x^B = (r(x^1)\, u_{n^1}, ..., r(x^m)\, u_{n^m}),$$

where $n^j = n(x^j)$ is the population size of x^j. The smoothed distribution gives each person in a group its representative income. The *standardized distribution* associated with $(x^1, ..., x^m)$ and r is defined by

$$x^W = r(x)\left(\frac{x^1}{r(x^1)}, ..., \frac{x^m}{r(x^m)}\right).$$

The standardized distribution rescales each group distribution so that the representative incomes of all groups equal the overall representative income $r(x)$.[10]

Either of these distributions may be used as a basis for decomposing total inequality into within- and between-group terms. The first defines the between-group term as $B = I(x^B)$, or the inequality level of the smoothed distribution, and then takes within-group inequality to be the residual $W = I(x) - I(x^B)$. The second defines within-group inequality as $W' = I(x^W)$, or the inequality level of the standardized distribution, and then fixes $B' = I(x) - I(x^W)$ as the between-group term. Both offer plausible assessments of within- and between-group inequality; it is therefore natural to require them to yield the same conclusions, namely, $W = W'$ and $B = B'$. Expressed differently, each defines a pathway between distribution x and

[9] The q-order means are strictly Shur-convex over this range of q. See Marshall and Olkin [25, p. 54].

[10] As noted above, these definitions make implicit use of the normalization and linear homogeneity properties for r. Normalization ensures that the representative income of the group distribution $r(x^j)\, u_{n^j}$ in the smoothed distribution is $r(x^j)$, while linear homogeneity ensures that the representative income of the group distribution $[r(x)/r(x^j)]\, x^j$ in the standardized distribution is $r(x)$.

complete equality by way of an intermediate distribution (x^B or x^W), with the decomposition being determined by the inequality levels along the path. Our property requires the resulting breakdown to be independent of the pathway taken.

A measure of inequality has a *path independent decomposition* or, more succinctly, is *path independent* if there is a representative income function r such that

$$I(x^1, ..., x^m) = I(x^W) + I(x^B) \tag{1}$$

for all $x^1, ..., x^m \in \mathscr{D}$ and $m \geqslant 2$. The first term on the right hand side of (1) corresponds to within-group inequality obtained when all groups are standardized to the same representative income level. The second term measures between-group inequality, which is what remains after group distributions are smoothed. The property requires overall inequality to be the sum of the within- and between-group terms.

An immediate implication of path independence is that any rescaling of subgroup distributions affects only the between-group term $B = I(x^B)$ and leaves the within-group term $W = I(x) - I(x^B)$ unchanged. In other words, $I(y) - I(y^B) = I(x) - I(x^B)$ whenever y is obtained from x by $y^k = \alpha^k x^k$ with $\alpha^k > 0$ for all k. In fact, the reverse implication is true as well, which can be seen by setting $y = x^W$ and noting that $I((x^W)^B)$ is completely equal. Consequently, path independence is equivalent to a natural requirement that within-group inequality be independent of between-group inequality.

Now which inequality measures satisfy path independence? It can be shown that Theil's second measure (or the mean logarithmic deviation), $D(x) = \ln[\mu(x)/g(x)]$, has a path-independent decomposition using the arithmetic mean as the representative income, while the variance of logarithms, $V_L = \sum_{i=1}^n (\ln x_i - \ln g)^2/n$, is path independent relative to the *geometric* mean.[11] Consider the following single parameter class of measures containing both D and (half) V_L:

$$I_q(x) = \begin{cases} \dfrac{1}{2}\, V_L(x), & q = 0 \\[2ex] \dfrac{1}{q} \ln \dfrac{\mu_q(x)}{g(x)}, & q \neq 0, \end{cases} \tag{2}$$

where, as before, $\mu_q(x) = (\sum_{i=1}^n x_i^q/n)^{1/q}$ is a q-order mean and $g(x) = \prod_{i=1}^n x_i^{1/n}$ is the geometric mean. Each member of the I_q family satisfies the five properties required of an inequality measure. In addition, $I_q(x)$ is

[11] See, for example, Love and Wolfson [24] and Anand [3, pp. 329–331]. The latter result follows from the decomposition properties of the variance.

continuous in the parameter q at each $x \in \mathcal{D}$, which follows immediately for $q \neq 0$, but requires a short argument (given in Section 5) for $q = 0$. Moreover,

$$\begin{aligned} I_q(x^W) + I_q(x^B) &= \frac{1}{q} [\ln(\mu_q(x^W)\, \mu_q(x^B)) - \ln(g(x^W)\, g(x^B))] \\ &= \frac{1}{q} [\ln \mu_q(x) - \ln g(x)] \\ &= I_q(x) \end{aligned}$$

(for $q \neq 0$) since $\mu_q(x^W) = \mu_q(x^B) = \mu_q(x)$ and $g(x^W) = \mu_q(x)\, g(x)/g(x^B)$. Consequently, each I_q measure has a path independent decomposition relative to the representative income function μ_q.

As an illustration of this, consider the distribution $x = (2, 2, 2, 8)$ partitioned into (2, 2) and (2, 8). When $q = 1$, so that the representative income is the arithmetic mean, the resulting standardized distribution is $x^W = (3.5, 3.5, 1.4, 5.6)$ while the smoothed distribution is $x^B = (2, 2, 5, 5)$. Consequently, for the second Theil measure, $I_1 = D$, we have $I_1(x^W) + I_1(x^B) = 0.112 + 0.101 = 0.213 = I_1(x)$. Alternatively, when $q = 2$, we obtain what might be called the *Euclidean* mean μ_2 (which is proportional to the Euclidean norm for each n). The intermediate distributions are now $x^W = (4.4, 4.4, 1.5, 6.0)$ and $x^B = (2, 2, 5.8, 5.8)$, and the decomposition is given by $I_2(x^W) + I_2(x^B) = 0.094 + 0.122 = 0.216 = I_2(x)$. Note that while these path independent measures have similar levels of overall inequality, they offer rather different assessments of within- and between-group inequality.

Perhaps the best-known family of decomposable inequality measures is the class of *generalized entropy* measures: $I_c(x) = [(\mu_c(x)/\mu(x))^c - 1]/(c(c-1))$ for $c \neq 0, 1$, with $I_c(x) = D(x)$ for $c = 0$ and $I_c(x) = 1/n \sum_{j=1}^{n} x_j/\mu(x) \ln(x_j/\mu(x))$ for $c = 1$ (Theil's first measure). Each has a special form of weighted additive decomposition,

$$I_c(x^1, ..., x^m) = \sum_{k=1}^{m} \frac{n^k}{n} \left(\frac{\mu^k}{\mu}\right)^c I_c(x^k) + I_c(x^B), \tag{3}$$

where $n^k = n(x^k)$, $\mu^k = \mu(x^k)$, and the representative income in x^B is the arithmetic mean. The first term of this decomposition is the weighted sum of the subgroup inequality levels, which can be particularly useful in, say, gauging the contribution of a single subgroup to total inequality. However, as noted by Shorrocks [28], the interpretation of this term as within-group inequality is not above criticism. One obvious problem is that, apart from the limiting cases of $c = 0, 1$, the subgroup inequality weights typically do not sum to one. For example, if we apply (3) to $x = (2, 2, 2, 8)$ for $c = 2$, so that I_c is (half) the squared coefficient of variation, then the weight on subgroup (2, 8) *itself* exceeds one.

More importantly for the present paper, for $c \neq 0$ the within-group term $W = I_c(x) - I_c(x^B)$ in (3) is *not* independent of the between-group term $B = I_c(x^B)$, since a rescaling of subgroup distributions will typically alter W. Consequently, by our previous discussion, the generalized entropy measures (apart from the second Theil measure) are not path independent. For example, consider $x = (2, 2, 2, 8)$ partitioned into (2, 2) and (2, 8), and let $c = 2$. Using the arithmetic mean as the representative income, we obtain the standardized distribution $x^W = (3.5, 3.5, 1.4, 5.6)$ and the smoothed distribution $x^B = (2, 2, 5, 5)$, as before. Then overall inequality is $I_c(x) = 0.275$ and yet the components in decomposition (1) are $I_c(x^W) = 0.090$ and $I_c(x^B) = 0.092$, yielding a shortfall of about one third of total inequality.

4. CHARACTERIZATION RESULTS

We now explore the implications of path independent decomposability. Our first result shows that every path independent measure has a weighted additive decomposition, where the weights are the population shares of the respective subgroups.

PROPOSITION 1. *Let I be an inequality measure and let r be a representative income function. Then I satisfies path independence for r if and only if*

$$I(x^1, ..., x^m) = \sum_{k=1}^{m} \frac{n^k}{n} I(x^k) + I(r(x^1)\, u_{n^1}, ..., r(x^m)\, u_{n^m}) \tag{4}$$

for any $m \geqslant 2$ and all $x^1, ..., x^m \in \mathscr{D}$.

Proof. Suppose that I satisfies path independence given r, and let $x = (y, z^1, ..., z^{m'}) \in \mathscr{D}$ be any distribution whose final $m' \geqslant 1$ components are completely equal distributions, while $y \in \mathscr{D}$ is arbitrary. It follows from path independence and scale invariance of I, as well as normalization of r, that

$$I(y, z^1, ..., z^{m'}) = I\left(\frac{y}{r(y)}, u_{n-n(y)}\right) + I(r(y)\, u_{n(y)}, z^1, ..., z^{m'}), \tag{5}$$

where $n = n(y, z^1, ..., z^{m'})$. This can be used to establish the following result.

Claim 1. $I(x) = \sum_{k=1}^{m} I(x^k/(r(x^k)), u_{n-n^k}) + I(x^B)$ for any $x = (x^1, ..., x^m)$ with $x^k \in \mathscr{D}$. Set $y = x^1$, and rearrange the remaining incomes in $x^2, ..., x^m$ into distributions $z^1, ..., z^{m'}$ in such a way that each $z^k \in \mathscr{D}$ is completely equal (and perhaps singleton). Then invoking symmetry of I, we obtain $I(x)=I(x^1/(r(x^1)),\ u_{n-n^1})+I(r(x^1)\,u_{n^1},\ x^2, ..., x^m)$ from (5). Applying the same reasoning to $(r(x^1)\,u_{n^1}, x^2, ..., x^m)$ with $y=x^2$ yields $I(r(x^1)\,u_{n^1}, x^2, ..., x^m)=I(x^2/(r(x^2)),\ u_{n-n^2})+I(r(x^1)\,u_{n^1},\ r(x^2)\,u_{n^2},\ x^3, ..., x^m)$. Continuing with this approach, we eventually obtain $I(r(x^1)\,u_{n^1}, ..., r(x^{m-1})\,u_{n^{m-1}}, x^m) = I(x^m/(r(x^m)),\ u_{n-n^m})+I(x^B)$. Consequently $I(x)=\sum_{k=1}^m I(x^k/(r(x^k)),\ u_{n-n^k})+I(x^B)$ which verifies the claim.

The next result shows that the within-group term can be further simplified.

Claim 2. $I(y/(r(y)), u_{n-p})=(p/n)\,I(y)$ for any $y\in\mathscr{D}$ satisfying $n(y)=p<n$. For ease of notation, let $a=y/(r(y))$ and $b=(y/(r(y)), u_{n-p})$, while $a^{(p)}$ and $b^{(p)}$ denote their respective p-replications. By symmetry and replication invariance of I we obtain $I(b)=I(b^{(p)})=I(a^{(p)},\ u_{p(n-p)})$. Now applying Claim 1 to $(a^{(p)},\ u_{p(n-p)})$ divided into the p-many vectors a and the single vector $u_{p(n-p)}$ yields

$$\begin{aligned} I(b) &= pI(a, u_{p(n-1)})+I(u_p, ..., u_p, u_{p(n-p)})+I(u_p, ..., u_p, u_{p(n-1)}) \\ &= pI(a, u_{p(n-1)}), \end{aligned} \tag{6}$$

where use has been made of normalization and linear homogeneity of r and normalization for I. Applying Claim 1 to the n-replication $a^{(n)}$ similarly entails

$$I(a)=nI(a, u_{(n-1)\,p})+I(u_p, ..., u_p)=nI(a, u_{(n-1)\,p}), \tag{7}$$

where replication invariance and normalization of I have been used. Combining (6) and (7) then yields $I(b)=(p/n)\,I(a)$ and hence by scale invariance,

$$I\left(\frac{y}{r(y)}, u_{n-p}\right)=I(b)=\frac{p}{n}\,I(a)=\frac{p}{n}\,I(y),$$

as desired.

Decomposition formula (4) follows immediately from the two claims, which completes one direction of the proof. For the converse direction, suppose that I exhibits decomposition (4) for every $m\geqslant 2$ and all $x^1, ..., x^m\in\mathscr{D}$. Applying (4) to the standardized distribution yields

$$\begin{aligned} I\left(\frac{x^1}{r(x^1)}, ..., \frac{x^m}{r(x^m)}\right) &= \sum_{k=1}^m \frac{n^k}{n}\,I\left(\frac{x^k}{r(x^k)}\right)+I(u_{n^1}, ..., u_{n^m}) \\ &= \sum_{k=1}^m \frac{n^k}{n}\,I(x^k), \end{aligned}$$

by normalization and scale invariance of I, and hence I is path independent.

Proposition 1 offers a first clue as to the kinds of restrictions imposed by this property: path independent measures have weighted additive decompositions with population share weights. Theil's second measure and the variance of logs both exhibit this form of decomposition (albeit at different r's), which ensures that each is path independent. On the other hand, the decompositions for the other generalized entropy measures do not have this special form and thus violate path independence (for $r=\mu$).

We now turn to the following question: Which forms of representative income functions are consistent with path independence?

Proposition 2. *Suppose that an inequality measure I satisfies path independence given the representative income function r. Then for some real q we have $r(x)=\mu_q(x)$ for all $x\in\mathscr{D}$.*

Proof. Assume that I is path independent for r. We begin by applying decomposition formula (4) to obtain the following functional equation for r:

$$r(x, y)=r(r(x)\, u_{n(x)}, r(y)\, u_{n(y)}) \qquad \text{for all} \quad x, y\in\mathscr{D}. \tag{8}$$

To establish this, consider any $x, y\in\mathscr{D}$ and $\alpha\in R_{++}$, and denote $n=n(x, y, \alpha)$. By applying decomposition (4) twice we obtain

$$\begin{aligned} I(x, y, \alpha) &= \frac{n(x)}{n} I(x)+\frac{n(y)}{n} I(y)+I(r(x)\, u_{n(x)}, r(y)\, u_{n(y)}, \alpha) \\ &= \frac{n(x)}{n} I(x)+\frac{n(y)}{n} I(y)+\frac{n-1}{n} I(r(x)\, u_{n(x)}, r(y)\, u_{n(y)}) \\ &\quad + I(r(r(x)\, u_{n(x)}, r(y)\, u_{n(y)})\, u_{n-1}, \alpha), \end{aligned}$$

where the invariance properties of I and r have been used. By decomposing the same distribution over different subgroups, we find

$$\begin{aligned} I(x, y, \alpha) &= \frac{n-1}{n} I(x, y)+I(r(x, y)\, u_{n-1}, \alpha) \\ &= \frac{n(x)}{n} I(x)+\frac{n(y)}{n} I(y)+\frac{n-1}{n} I(r(x)\, u_{n(x)}, r(y)\, u_{n(y)}) \\ &\quad + I(r(x, y)\, u_{n-1}, \alpha) \end{aligned}$$

once again invoking invariance properties. Thus,

$$I(r(r(x)\, u_{n(x)}, r(y)\, u_{n(y)})\, u_{n-1}, \alpha)=I(r(x, y)\, u_{n-1}, \alpha) \tag{9}$$

for arbitrary $\alpha > 0$. Substituting $\alpha = r(x, y)$ into (9) yields

$$I(r(r(x)\, u_{n(x)}, r(y)\, u_{n(y)})\, u_{n-1}, r(x, y)) = 0$$

by normalization of I. But this same property ensures that the distribution $(r(r(x)\, u_{n(x)}, r(y)\, u_{n(y)})\, u_{n-1}, r(x, y))$ is completely equal, which then leads to (8).

Now consider the restriction of r to the two-income domain R^2_{++}. The assumed properties of r ensure that $r(s, t) = r(t, s)$ for $s, t > 0$; r is strictly increasing in each argument; $r(\lambda s, \lambda t) = \lambda r(s, t)$ for $s, t, \lambda > 0$; $r(s, s) = s$ for $s > 0$; and r is continuous. Moreover, if we apply (8) to the distributions $x = (s, t)$ and $y = (w, v)$, and then to $x' = (s, w)$ and $y' = (t, v)$, we find that

$$r(r(s, t), r(w, v)) = r(r(s, w), r(t, v)) \qquad \text{for all} \quad s, t, w, v > 0, \tag{10}$$

where symmetry and replication invariance of r have been invoked. Aczel and Dombres [2, pp. 249, 291] have shown that the general solution to (10) satisfying the given properties is the means of order q. Therefore, there exists some real q such that

$$r(x) = \mu_q(x) \tag{11}$$

for all x with $n(x) = 2$.

An induction argument extends characterization (11) to any x with $n(x) = 2^k$ for positive integer k. Indeed, suppose that the characterization holds for all distributions of size 2^j, where $j = 1, ..., k-1$. Then given any $x \in \mathscr{D}$ with $n(x) = 2^k$, we may split x into distributions y and z satisfying $n(y) = n(z) = 2^{k-1}$. By hypothesis, $r(y) = \mu_q(y)$ and $r(z) = \mu_q(z)$ and hence $r(r(y), r(z)) = \mu_q(\mu_q(y), \mu_q(z)) = \mu_q(y, z) = \mu_q(x)$. Equation (8) immediately gives us $r(x) = \mu_q(x)$ for $n(x) = 2^k$ as desired.

To complete the proof, we use another induction argument—this time in the opposite direction. Suppose that characterization (11) is valid for all distributions of size n. We show that it holds for all distributions of size $n-1$. Pick any $x \in \mathscr{D}$ satisfying $n(x) = n-1$, so that $n(x, r(x)) = n$. Equation (8) and normalization yield $r(x, r(x)) = r(r(x)\, u_{n(x)}) = r(x)$, so that by the hypothesis $\mu_q(x, r(x)) = r(x)$. It is easy to verify[12] that the only positive real α satisfying $\alpha = \mu_q(x, \alpha)$ is $\alpha = \mu_q(x)$. Therefore $r(x) = \mu_q(x)$, and hence characterization (11) applies for all distributions of size $n-1$. By induction, whenever (11) is true for a given population size, it must hold for every smaller sized distribution. But since (11) applies for every population size

[12] For $q = 0$ it is clear that $\alpha = (x_1 \cdots x_{n-1})^{1/(n-1)}$ follows from $\alpha = (x_1 \cdots x_{n-1} \alpha)^{1/n}$; for $q \neq 0$ it is clear that $\alpha = [(x_1^q + \cdots + x_{n-1}^q)/(n-1)]^{1/q}$ follows from $\alpha = [(x_1^q + \cdots + x_{n-1}^q + \alpha^q)/n]^{1/q}$.

2^k (where k is a positive integer), we conclude that $r(x)=\mu_q(x)$ for all $x \in \mathscr{D}$, which completes the proof.

Proposition 2 concludes that the only representative income functions consistent with path independence are the means of order q. This result is similar to the welfare-based characterizations of μ_q for $q \leqslant 1$ by Atkinson [4] and Blackorby *et al.* [7]. However, their approach begins by assuming separability (or additivity) of the welfare function, whereas Proposition 2 places no such assumption on r. Our proof shows that path independence on I implies (8) for r, which in turn leads to the conclusion that $r=\mu_q$.

Note that the monotonicity requirement for r rules out the minimum income and the maximum income as representative incomes. Interestingly, both extreme incomes satisfy (8) and, moreover, have associated measures satisfying (1),[13] so our conclusions obviously depend on this assumption. On the other hand, the median income—which has well-known difficulties aggregating over populations—violates (8) as well as monotonicity and would therefore not be a candidate for r even if monotonicity were relaxed.

Our first two propositions have shown the consequences of path independence for the decomposition formula and the representative income function. We now use these results as a basis for characterizing the entire class of path independent measures.[14]

PROPOSITION 3. *I is a path independent inequality measure if and only if I is a positive multiple of I_q for some $q \in R$.*

Proof. If I is path independent, then by Propositions 1 and 2 there is a $q \in R$ for which

$$I(x^1, ..., x^m) = \sum_{k=1}^{m} \frac{n^k}{n} I(x^k) + I(\mu_q(x^1)\, u_{n^1}, ..., \mu_q(x^m)\, u_{n^m}) \tag{12}$$

for any $m \geqslant 2$ and all $x^1, ..., x^m \in \mathscr{D}$. We will show that I must be I_q for this specific q.

Suppose that $q=0$ in (12). Define $J\colon \mathscr{D} \to R$ by

$$J(y) = n(y)\, I(e^{y_1}, ..., e^{y_{n(y)}}) \qquad \text{for all} \quad y \in \mathscr{D}. \tag{13}$$

Then J satisfies

$$J(y^1, y^2) = J(y^1) + J(y^2) + J(\mu(y^1)\, u_{n^1}, \mu(y^2)\, u_{n^2}) \qquad \text{for all} \quad y^1, y^2 \in \mathscr{D}, \tag{14}$$

[13] Indeed, $I_{-\infty}(x) = \ln g(x) - \ln(\min x_i)$ satisfies (1) for the minimum income while $I_{\infty}(x) = \ln(\max x_i) - \ln g$ satisfies (1) for the maximum.

[14] We are indebted to Tony Shorrocks for substantial improvements in the proof of Proposition 3.

as well as symmetry and continuity, so that by Theorem 3 in Shorrocks [29], there exists a continuous function ϕ: $R_{++} \to R$ such that

$$J(y) = \sum_{i=1}^{n(y)} [\phi(y_i) - \phi(\mu(y))] \qquad \text{for all} \quad y \in \mathscr{D}. \tag{15}$$

Scale invariance of I ensures that

$$J(s+\lambda, t+\lambda) = J(s, t) \qquad \text{for all} \quad s, t, \lambda > 0$$

and so from (15)

$$\phi(s+\lambda) + \phi(t+\lambda) - 2\phi\left(\frac{s+t}{2} + \lambda\right)$$
$$= \phi(s) + \phi(t) - 2\phi\left(\frac{s+t}{2}\right) \qquad \text{for all} \quad s, t, \lambda > 0.$$

Then $h(s, \lambda) = \phi(s+\lambda) - \phi(s)$ is a continuous function satisfying Jensen's equation in $s > 0$ for each $\lambda > 0$, and so by Aczel [1, p. 44] we have $h(s, \lambda) = a(\lambda)\, s + b(\lambda)$ for some functions a and b. Thus

$$\phi(s+\lambda) = \phi(s) + a(\lambda)\, s + b(\lambda) \qquad \text{for all} \quad s, \lambda > 0 \tag{16}$$

and hence

$$\phi(s+1+\lambda) = \phi(s+1) + a(\lambda)(s+1) + b(\lambda)$$
$$= \phi(1) + a(s) + b(s) + a(\lambda) + b(\lambda) + a(\lambda)\, s$$
$$\text{for all} \quad s, \lambda > 0.$$

Expanding $\phi(\lambda + 1 + s)$ in a similar manner yields

$$a(\lambda)\, s = a(s)\, \lambda \qquad \text{for all} \quad s, \lambda > 0,$$

and hence there is a constant $c \in R$ such that $a(s) = cs$ for all $s > 0$. Substituting this into (16) yields

$$\phi(s) + c\lambda s + b(\lambda) = \phi(\lambda) + cs\lambda + b(s) \qquad \text{for all} \quad s, \lambda > 0,$$

and so there is a constant $d \in R$ such that $b(s) = \phi(s) + d$ for all $s > 0$. Substituting again into (16) yields

$$\phi(s+\lambda) = \phi(s) + c\lambda s + \phi(\lambda) + d \qquad \text{for all} \quad s, \lambda > 0.$$

The continuous function $\gamma(s) = \phi(s) + d - cs^2/2$ satisfies the Cauchy equation

$$\gamma(s+\lambda) = \gamma(s) + \gamma(\lambda) \qquad \text{for all} \quad s, \lambda > 0,$$

and hence by Aczel [1, p. 34] we have $\gamma(s) = \alpha s$ for some constant α. This in turn means that

$$\phi(s) = \alpha s + cs^2/2 - d \qquad \text{for all} \quad s > 0$$

and therefore

$$J(y) = \frac{c}{2} \sum_{i=1}^{n(y)} [y_i^2 - \mu(y)^2] \qquad \text{for all} \quad y \in \mathscr{D},$$

using (16). Reversing the transformation in (13) then yields

$$I(x) = \frac{c}{2n(x)} \sum_{i=1}^{n(y)} [(\ln x_i)^2 - (\ln g(x))^2] = \frac{c}{2} V_L(x) \tag{17}$$

for all $x \in \mathscr{D}' = \{x \in \mathscr{D} : x_i > 1 \text{ for all } i\}$. Since for any $x \in \mathscr{D}$ there is some $\beta > 0$ for which $\beta x \in \mathscr{D}'$ it follows from scale invariance that (17) holds for all $x \in \mathscr{D}$.

Now suppose that $q \neq 0$ in (12). Define $J: \mathscr{D} \to R$ by

$$J(y) = n(y)\, I(y_1^{1/q}, ..., y_{n(y)}^{1/q}) \qquad \text{for all} \quad y \in \mathscr{D}. \tag{18}$$

Then J is symmetric, is continuous, and satisfies (14), so that once again by Theorem 3 in Shorrocks [29], there exists a continuous function $\phi: R_{++} \to R$ satisfying (15). Scale invariance of I yields

$$J(s\lambda, t\lambda) = J(s, t) \qquad \text{for all} \quad s, t, \lambda > 0$$

and so from (15) we obtain

$$\phi(s\lambda) + \phi(t\lambda) - 2\phi\left(\frac{s+t}{2}\lambda\right)$$
$$= \phi(s) + \phi(t) - 2\phi\left(\frac{s+t}{2}\right) \qquad \text{for all} \quad s, t, \lambda > 0.$$

Then the continuous function $h(s, \lambda) = \phi(s\lambda) - \phi(s)$ satisfies Jensen's equation in $s > 0$ for each $\lambda > 0$, which yields $h(s, \lambda) = a(\lambda)\, s + b(\lambda)$ for some functions a and b, according to Aczel [1, p. 44]. Thus

$$\phi(s\lambda) = \phi(s) + a(\lambda)\, s + b(\lambda)$$
$$= \phi(1) + a(s) + b(s) + a(\lambda) + b(\lambda) + a(\lambda)(s-1) \tag{19}$$

for all $s, \lambda > 0$. Expanding $\phi(\lambda s)$ in the same way yields

$$a(\lambda)(s-1) = a(s)(\lambda-1) \qquad \text{for all} \quad s, \lambda > 0,$$

and hence there is a constant $c \in R$ such that $a(s) = c(s-1)$ for all $s > 0$. Substituting this into (19) yields

$$\phi(s) + c(\lambda-1)\, s + b(\lambda) = \phi(\lambda) + c(s-1)\, \lambda + b(s) \qquad \text{for all} \quad s, \lambda > 0,$$

and so there is a constant $d \in R$ such that $b(s) = \phi(s) + d - c(s-1)$ for all $s > 0$. Substituting again into (19) yields

$$\phi(s\lambda) = \phi(s) + c(\lambda-1)(s-1) + \phi(\lambda) + d \qquad \text{for all} \quad s, \lambda > 0.$$

The continuous function $\gamma(s) = \phi(s) + d - c(s-1)$ satisfies the Cauchy equation

$$\gamma(s\lambda) = \gamma(s) + \gamma(\lambda) \qquad \text{for all} \quad s, \lambda > 0,$$

and hence by Aczel [1, p. 34], $\gamma(s)$ is a multiple of $\ln s$, say $\gamma(s) = -(\alpha/q^2) \ln s$ for some constant $\alpha \in R$. This in turn means that

$$\phi(s) = -(\alpha/q^2) \ln s + c(s-1) \qquad \text{for all} \quad s > 0$$

and therefore

$$J(y) = \frac{\alpha}{q^2} \sum_{i=1}^{n(y)} [\ln \mu(y) - \ln y_i] \qquad \text{for all} \quad y \in \mathcal{D}.$$

Reversing the transformation in (18) then yields

$$I(x) = \frac{\alpha}{qn(x)} \sum_{i=1}^{n(y)} [\ln \mu_q(x) - \ln x_i] = \alpha I_q(x) \qquad \text{for all} \quad x \in \mathcal{D},$$

as desired.

Consequently, $I(x) = \alpha I_q(x)$ for some $q, \alpha \in R$, where $\alpha > 0$ follows from normalization of I. Conversely, suppose that $I(x) = \alpha I_q(x)$ for some $q \in R$ and $\alpha > 0$. Then as noted above, I is an inequality measure that has a path independent decomposition for the representative income function $r(x) = \mu_q(x)$. This completes the proof.

We previously noted that the members of the I_q family have path independent decompositions. Proposition 3 shows that the I_q measures are essentially the *only* measures satisfying path independence and thus are completely characterized by this property.

5. DISCUSSION

Every I_q measure with nonzero q is based on a comparison of the representative income μ_q to the *geometric* mean g. This form is clearly reminiscent of the generalized entropy family (defined above) and the Atkinson family (i.e., $1-\mu_q/\mu$, for $q \leqslant 1$), both of which depend on a comparison of μ_q with the *arithmetic* mean μ. Indeed, I_q has a graphical representation analogous to that given by Atkinson [4]; but rather than viewing inequality as a departure from the line of equal mean distributions, I_q takes as its standard a hyperbola containing all distributions with the same geometric mean.

The interpretation of I_q becomes more transparent in log-income space (where the transformation $y_i = \ln x_i$ has been used), since then the distributions with the same geometric mean as x^* can be viewed as a line segment through the associated log-income distribution y^*. The level sets for μ_q are curves that are convex for $q<0$ and concave for $q>0$, and cross the 45° line at the log-income distribution whose entries are all equal to $\ln \mu_q$. (See Fig. 1.) Consequently, inequality as measured by I_q is the extent to which the μ_q curve departs from the geometric mean line in log-income space. For example, the distance $b-a=\ln g-\ln \mu_{-1}$ in Fig. 1 is the inequality level $I_{-1}(x)$, while $c-b=\ln \mu-\ln g$ is $I_1(x)=D(x)$ or inequality according to the second Theil measure.

This approach leads quite naturally to an interpretation of I_q in terms of the Kolm–Pollak measures of absolute inequality.[15] The *Kolm means of order q* [22, 23], which are defined as

$$K_q(x)=\begin{cases}\ln\left(\frac{1}{n}\sum_{i=1}^{n} e^{qx_i}\right)^{1/q}, & q\neq 0\\ \mu(x), & q=0,\end{cases}$$

share all the properties of the usual q-order means except linear homogeneity. Instead, they have the property that the *addition* of a constant to all incomes raises K_q by that same constant. Consequently, the *Kolm–Pollak* measures

$$A_q(x)=[K_q(x)-K_0(x)]/q=[K_q(x)-\mu(x)]/q, \qquad q\neq 0,$$

which are based on the absolute difference of two Kolm means, are *translation invariant* in that an addition of a constant to all incomes leaves A_q

[15] See Kolm [22, 23], Pollak [26], and Blackorby and Donaldson [6]. We have extended the class to $q>0$ by multiplying the standard definition by $1/q$ and by taking the limit as $q \to 0$ to be the measure at $q=0$.

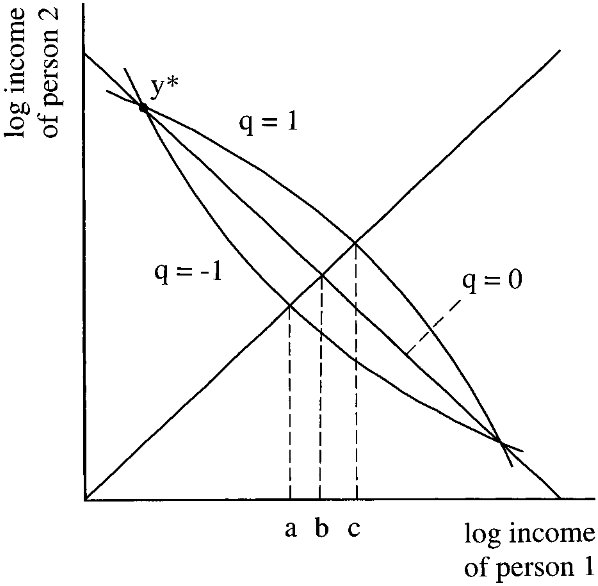

FIG. 1. Level curves of μ_q for three values of q.

unchanged. The same follows for the continuous limit of A_q as q tends to 0, which is

$$A_0(x) = V(x)/2,$$

or half the variance[16].

Now consider once again the log transformation $y_i = \ln x_i$. When a Kolm q-order mean is applied to the distribution of log incomes $y = (y_1, ..., y_n)$, we obtain the respective q-order mean μ_q. Consequently,

$$I_q(x) = A_q(y)$$

for all q, which shows that I_q is the Kolm–Pollak measure of absolute inequality applied to the distribution of log incomes. This, of course, is consistent with the definition of I_0 as (half) the variance of the logarithm of incomes and leads immediately to the result that I_0 is the limit of I_q. Moreover, it sheds light on the representation of I_q in Fig. 1 since the quantities depicted there are actually based on K_q and A_q over y and translate to μ_q and I_q when viewed as functions of the original distribution x.

[16] $\lim_{q \to 0} A_q = \lim_{q \to 0} \ln(\sum_{i=1}^n e^{q(x_i - \mu)}/n)/q^2 = \lim_{q \to 0} 1/(2qn) \sum_{i=1}^n (x_i - \mu)\, e^{q(x_i - \mu)} = 1/(2n) \sum_{i=1}^n (x_i - \mu)^2$, or half the variance, using L'Hopital's rule twice.

While the above interpretation relies on a given transformation of incomes and a class of measures, the proof of Proposition 3 provides a second way of viewing I_q for $q \neq 0$ that depends on a fixed inequality measure—Theil's second measure—and a range of transformations. Consider the *power distribution* $x^q = (x_1^q, ..., x_n^q)$, which transforms each income into the qth power of that income. Since $\mu_q(x) = [\mu(x^q)]^{1/q}$ for all $q \neq 0$, it follows that

$$I_q(x) = D(x^q)/q^2 \qquad \text{for all} \quad x \in \mathscr{D}, \tag{20}$$

so that I_q is proportional to the second Theil measure applied to the power distribution x^q for each nonzero q.[17] Equation (20) shows that there is a curious link between D and V_L as q tends to 0, namely, $\lim_{q \to 0} D(x^q)/q^2 = V_L/2$. Consequently, the way the variance of logarithms ranks distributions is approximately the way that D ranks qth power distributions for small q.

This expression also provides insight into the meaning of the parameter q. At lower q, the transformation x_i^q becomes more concave (or less convex) and hence the resulting measure places greater relative weight on differences at the lower end of the distribution; at higher q, the opposite is true.[18] The effect of q on I_q can be graphically illustrated with the help of the Kolm triangle, which depicts all distributions of a given amount of income among three persons.[19] Figure 2 presents level curves for several I_q measures. Note that the slope through x^* becomes steeper as q becomes larger.[20] At lower q's the level set is roughly parallel to the line from x^* to x', signifying that transfers between persons 2 and 3 leave I_q unaffected, while transfers involving the poorest person, 1, have the greatest impact on I_q. At higher q's, the slope rises, reflecting an increased sensitivity to transfers among the rich and less concern for the lowest income.

Figure 2 also shows that the slope of a level set through x^* can be *negative* for certain values of q, reflecting the fact that a small regressive transfer between persons 2 and 3 can actually lower inequality. Violations of the transfer principle by $I_0 = V_L/2$ have been well studied in the literature.[21] We now see which of the other I_q measures satisfy the transfer principle and which of these measures violate it.

[17] A similar approach can be used with other generalized entropy measures to obtain measures that have additive decompositions but are not path independent. See Foster and Shneyerov [18].

[18] In this respect, I_q is similar to the generalized entropy measures whose parameter likewise reflects sensitivity of transfers at different parts of the distribution (which is natural since both are based on μ_q).

[19] See Kolm [21], Sen [27], and Blackorby and Donaldson [5].

[20] This can be shown for any x^* with $x_1^* < x_2^* < x_3^*$, although the verification is tedious and will be omitted.

[21] See Foster and Ok [16] and the references cited therein.

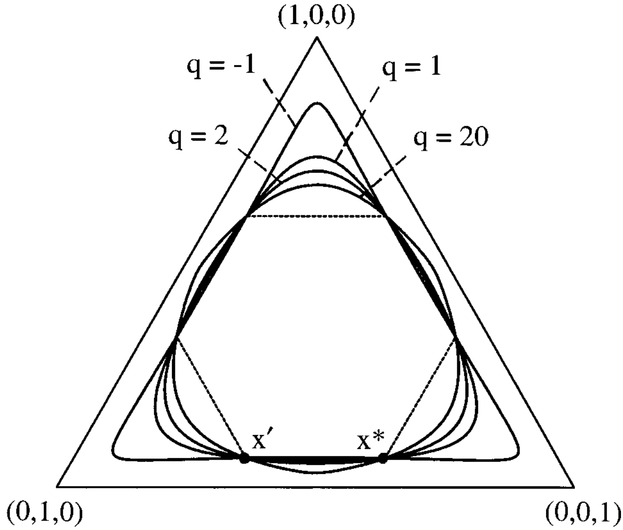

FIG. 2. Level curves of I_q in the Kolm triangle.

By a classic result (see, for example, Marshall and Olkin, [25, p. 57]), I_q satisfies the transfer principle whenever

$$\partial I_q/\partial x_i > \partial I_q/\partial x_j \tag{21}$$

for all $x_i > x_j > 0$. Now, $\partial I_q/\partial x_i = (x_i^{q-1}/\mu_q^q - x_i^{-1})/(qn)$, and so (21) holds whenever $q \geqslant 1$. On the other hand, if $q < 1$, then $\partial I_q/\partial x_i$ is not an increasing function of x_i; in fact, for x_i beyond $s = (1-q)^{-1/q}\mu_q$ the partial derivative is actually decreasing in x_i (for $q \neq 0$). Consequently, for any x_i and x_j above s, the inequality in (21) will be reversed and the transfer principle will be violated. Figure 3 depicts partial derivative functions $\partial I_q/\partial x_i$ for three values of q (namely, -1, 1.5, and 3), holding μ_q fixed at 1 and setting $n = 10$. Note that for $q = -1$, the function is negatively sloped after $s = 2$, which leads to I_q's violation of the transfer principle. In contrast, the other two curves are increasing throughout their domain and hence the transfer principle is satisfied by their associated measures.

One important distinction between the cases $q = 1.5$ and $q = 3$ is that for the former, $\partial I_q/\partial x_i$ is concave, while for the latter it is not. This means that the effect on $I_{1.5}$ of a marginal transfer between incomes x_i and x_j a fixed distance apart becomes larger as x_i falls, and this fact can be verified for all q in the range [1, 2]. Measures for which a given sized transfer has greater impact at lower incomes are called (*weakly*) *transfer sensitive*.[22] It can be

[22] See Shorrocks and Foster [31], who also formulate a stronger transfer sensitivity property.

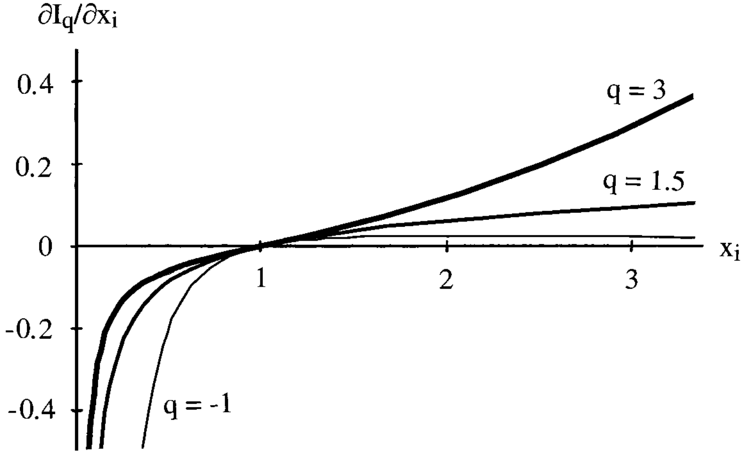

FIG. 3. The transfer principle and $\partial I_q/\partial x_i$.

verified that all I_q measures with $q \in [1, 2]$ satisfy this property. On the other hand, one can show that for incomes above $s = \mu_3$, a transfer between x_i and x_j a fixed distance apart has greater effect at higher incomes for the measure I_3. Indeed, a similar s can be found for any $q > 2$, which implies that all I_q with $q > 2$ violate transfer sensitivity. Therefore, I_q measures in the range $q \in [1, 2]$ have obvious advantages.

An additional property that has been suggested by Shorrocks [30] is *subgroup consistency*, which requires an inequality measure to increase whenever inequality in each subgroup increases, holding population size and mean income in each subgroup fixed.[23] If this property is regarded as essential, then as noted by Shorrocks, it limits consideration to the generalized entropy family (or monotonic transformations thereof). Consequently, the only path independent *and* subgroup consistent measure is Theil's second measure (or a multiple thereof). On the other hand, it should be noted that the subgroup consistency property, as normally defined, relies crucially on the use of the *arithmetic* mean as the representative income. If, as has been suggested in the paper, alternative notions are allowed, we might instead require inequality to increase with an increase in the subgroup inequalities, holding the *representative income* of each subgroup fixed. Under this broadened view of subgroup consistency, then, each of the I_q measures would exhibit the requisite relationship between subgroup and overall inequality, since the between-group term would be unchanged while the within-group term would increase.

One feature of the measures that should prove useful in empirical applications is its population-share weighted decomposition (4). In

[23] See Foster and Sen [17] for a critique of this property.

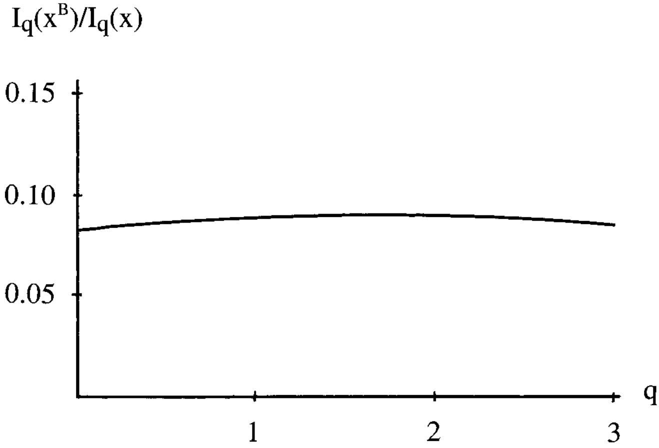

FIG. 4. Robustness of the between-group contribution.

particular, many analyses of inequality are concerned with the evolution of inequality over time and whether changes are due to (i) altered inequality within subgroups; (ii) demographic changes or population shifts among subgroups; or (iii) changes in the level of between-group inequality.[24] Population share weighting allows the researcher to separate out these effects, and so the I_q measures are particularly useful in these applications. In contrast, the weighted decompositions of the generalized entropy measures typically involve products of population shares and means and hence are not so amenable to this type of analysis.

Our characterization results have shown that there is an entire family of measures having the requisite decomposition, which offers the analyst a range of measures to evaluate within- and between-group inequality contributions. The parameter q can be interpreted as the extent to which the measure is sensitive to changes among the higher incomes. Even when we restrict attention to the Lorenz consistent and transfer sensitive subset, there is still a range (with q between 1 and 2) of measures available. Of course, as with any range of measures, there is a possibility that results obtained with one measure may be contradicted by another. The use of intersection quasiorderings, which indicate when all measures agree on a comparison, might be brought to bear on the problem of robust inequality comparisons.[25] Perhaps even more interesting is the question of robustness of the *between-group contribution* $I_q(x^B)/I_q(x)$ (and hence the within-group contribution $I_q(x^W)/I_q(x)$) to changes in the parameter q. One might, for

[24] See, for example, Shorrocks and Mookherjee [32] and Jenkins [20].

[25] See Atkinson [4], Sen [27], Shorrocks and Foster [31], and Foster and Sen [17] for discussions of this approach.

example, evaluate whether the contribution terms for Theil's second measure are robust to changes in q within some parameter range, say $0 \leqslant q \leqslant 3$. If the contributions obtained for D are greatly at variance with those of other I_q measures in this range, we might doubt the conclusions of D. On the other hand, if the contributions for D are mirrored closely by these measures, we may have greater faith in the original results. Figure 4 provides a simple illustration of this approach, based on a sample of 13,720 weekly earned incomes from the 1991 CPS, decomposed according to gender. According to the second Theil measure, the between-group contribution is 8.9%, and this is mirrored in the values found for I_2 (9.0%) and V_L (8.2%) and the remaining I_q measures in the range $0 \leqslant q \leqslant 3$.

REFERENCES

1. J. Aczel, "Lectures on Functional Equations and Their Applications," Academic Press, New York, 1966.
2. J. Aczel and J. Dombres, "Functional Equations Containing Several Variables," Cambridge University Press, Cambridge, UK, 1989.
3. S. Anand, "Inequality and Poverty in Malaysia," Oxford University Press, London, 1983.
4. A. B. Atkinson, On the measurement of inequality, *J. Econ. Theory* **2** (1970), 244–263.
5. C. Blackorby and D. Donaldson, Measure of relative equality and their meaning in terms of social welfare, *J. Econ. Theory* **18** (1978), 59–80.
6. C. Blackorby and D. Donaldson, A theoretical treatment of indices of absolute inequality, *Int. Econ. Rev.* **21** (1980), 107–136.
7. C. Blackorby, D. Donaldson, and M. Auersperg, A new procedure for the measurement of inequality within and among population subgroups, *Canadian J. Econ.* **14** (1981), 665–685.
8. F. Bourguignon, Decomposable income inequality measures, *Econometrica* **47** (1979), 901–920.
9. F. A. Cowell, On the structure of additive inequality measures, *Rev. Econ. Studies* **47** (1980), 521–531.
10. F. A. Cowell, "Measuring Inequality," Prentice–Hall/Harvester, London, 1995.
11. F. A. Cowell and S. P. Jenkins, How much inequality can we explain? A methodology and an application to the United States, *Econ. J.* **105** (1995), 421–430.
12. F. A. Cowell and K. Kuga, Inequality measurement: an axiomatic approach, *European Econ. Rev.* **15** (1981), 287–305.
13. D. Donaldson and J. A. Weymark, A single parameter generalization of the Gini indices of inequality, *J. Econ. Theory* **22** (1980), 67–86.
14. J. E. Foster, An axiomatic characterization of the Theil measure of income inequality, *J. Econ. Theory* **65** (1983), 105–121.
15. J. E. Foster, Inequality measurement, *in* "Fair Allocation" (H. P. Young, Ed.), Proceedings of Symposia in Applied Mathematics, American Mathematical Society, 1985.
16. J. E. Foster and E. Ok, Lorenz dominance and the variance of logarithms, *Econometrica* **67** (1999), 901–907.
17. J. E. Foster and A. Sen, On economic inequality after a quarter century, *in* "On Economic Inequality" (A. Sen), expanded ed., Oxford University Press, Oxford, 1997.

18 J. E. Foster and A. Shneyerov, A general class of additively decomposable inequality measures, *Econ. Theory* **14** (1999), 89–111.
19. G. H. Hardy, J. E. Littlewood, and G. Polya, “Inequalities,” Cambridge University Press, Cambridge, UK, 1934.
20. S. P. Jenkins, Accounting for inequality trends: Decomposition analyses for the UK, 1971–86, *Economica* **65** (1995), 29–63.
21. S. C. Kolm, The optimal production of social justice, *in* “Public Economics” (J. Margolis and H. Guitton, Eds.), McMillan, London, 1969.
22. S. C. Kolm, Unequal inequalities I, *J. Econ. Theory* **12** (1976), 416–442.
23. S. C. Kolm, Unequal inequalities II, *J. Econ. Theory* **13** (1976), 82–111.
24. R. Love and M. C. Wolfson, “Income Inequality: Statistical Methodology and Canadian Illustration,” Statistics Canada, 1976.
25. A. W. Marshall and I. Olkin, “Inequalities: Theory of Majorization and its Applications,” Academic Press, New York, 1979.
26. R. Pollak, Additive utility functions and linear engel curves, *Rev. Econ. Studies* **38** (1971), 401–414.
27. A. Sen, “On Economic Inequality,” expanded edition, Oxford University Press, Oxford, 1997.
28. A. F. Shorrocks, The class of additively decomposable inequality measures, *Econometrica* **48** (1980), 613–625.
29. A. F. Shorrocks, Inequality decomposition by population subgroups, *Econometrica* **52** (1984), 1369–1385.
30. A. F. Shorrocks, Aggregation issues in inequality measurement, *in* “Measurement in Economics” (W. Eichhorn, Ed.), Physica-Verlag, Heidelberg, 1988.
31. A. F. Shorrocks and J. E. Foster, Transfer sensitive inequality measures, *Rev. Econ. Studies* **54** (1987), 485–497.
32. A. F. Shorrocks and D. Mookherjee, A decomposition analysis of the trend in UK income inequality, *Econ. J.* **92** (1982), 886–902.
33. P. Tsakloglou, Aspects of inequality in Greece: Measurement, decomposition and intertemporal change: 1974, 1982, *J. Development Econ.* **40** (1993), 53–74.