



Transfer Sensitive Inequality Measures

Author(s): Anthony F. Shorrocks and James E. Foster

Source: *The Review of Economic Studies*, Vol. 54, No. 3, (Jul., 1987), pp. 485-497

Published by: The Review of Economic Studies Ltd.

Stable URL: <http://www.jstor.org/stable/2297571>

Accessed: 11/08/2008 12:03

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=resl>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact support@jstor.org.

Transfer Sensitive Inequality Measures

ANTHONY F. SHORROCKS

Essex University

and

JAMES E. FOSTER

Purdue University

First version received May 1985; final version accepted November 1986 (Eds.)

Transfer sensitivity has been seen as a means of strengthening the Pigou-Dalton "principle of transfers", by ensuring that more weight in the inequality assessment is attached to transfers taking place lower down in the distribution. This paper examines the concept of transfer sensitivity in detail and proposes a new definition that can be usefully applied in general contexts. The definition is based on the notion of "favourable composite transfers" which involve a regressive transfer combined with a simultaneous progressive transfer at a lower income level. The paper proceeds to identify when one distribution can be obtained from another using a sequence of progressive transfers and favourable composite transfers, and hence when all transfer sensitive Pigou-Dalton indices agree on their pairwise inequality ranking. Since agreement occurs in some situations when Pigou-Dalton indices are not unanimous, transfer sensitivity adds power to the "unambiguous" inequality judgements based on the Pigou-Dalton condition and, in particular, enables distributions whose Lorenz curves intersect to be conclusively ranked.

1. INTRODUCTION AND SUMMARY

The Pigou-Dalton "principle of transfers" requires that any mean-preserving progressive transfers lowers the value of an inequality index. Since it captures an essential element of our concept of inequality, the Pigou-Dalton condition assumes a central role in the theory of inequality measurement and is at the heart of several well known results. As a basis for inequality comparisons, however, its scope is severely limited. On its own, it does not allow us to pass judgement when distributions are defined over populations of different sizes, or when they have different means. Nor does it enable us to rank a pair of distributions if both progressive and regressive transfers are needed to convert one distribution into the other.

The significance of this last point becomes clear if we consider a situation in which a millionaire makes a small (regressive) transfer to a slightly more affluent millionaire and a simultaneous large (progressive) transfer to the poorest person in society. It is difficult to believe that anyone would seriously wish to argue that inequality had risen as a result of this combined transfer. Yet such an eccentric conclusion is permitted by the Pigou-Dalton condition. For while the regressive transfer increases inequality and the progressive transfer reduces inequality, no constraint is placed on the relative magnitudes of these two effects. It is therefore possible for a Pigou-Dalton index to attach greater importance to the small transfer between millionaires and thereby conclude that the combined transfer leads to a rise in inequality.

The above example illustrates a fundamental weakness of the Pigou-Dalton criterion and directs attention towards a means of prohibiting eccentric inequality judgements of the kind described. One solution is to demand that the inequality index is more sensitive

to transfers taking place lower down in the distribution. A “transfer sensitivity” requirement of this type has been discussed on a number of previous occasions.¹ However, there has never been a detailed discussion of this property or its implications. Nor is there even a satisfactory general definition.

This paper offers a novel and powerful definition of transfer sensitivity that agrees with earlier formulations in the context in which they were proposed. More specifically, the definition here is based on composite transfers which combine a progressive transfer with a regressive transfer at a higher income level. Intuition suggests that if the progressive and regressive transfers are, in some sense, comparable, and if more emphasis is placed on transfers occurring lower down in the distribution, then the net effect of such composite transfers will be “favourable” (i.e. inequality reducing). Identifying a suitable comparability condition leads us to the notion of a “favourable composite transfer” (abbreviated to FACT). An inequality measure is then said to be transfer sensitive if it decreases under the operation of any FACT.²

This definition of transfer sensitivity, discussed in detail in the next section, has obvious similarities with the conventional statement of the Pigou–Dalton condition, and the analogy is exploited fully later on in the paper. Section 3 restates the conditions that are known to be equivalent to an “unambiguous” inequality judgement based on Pigou–Dalton indices, and shows how an analogous result can be obtained for the conclusive inequality ranking based on transfer sensitive inequality measures. Since transfer sensitivity imposes an additional restriction, conclusive judgements can be obtained in situations where Pigou–Dalton indices fail to agree. In particular, it will often be possible to rank distributions whose Lorenz curves intersect. Section 4 provides simple necessary and sufficient conditions for a conclusive (transfer sensitive) inequality ranking of distributions whose Lorenz curves intersect just once.

2. THE DEFINITION OF TRANSFER SENSITIVITY

Consider a homogeneous population of n individuals whose incomes x_i are drawn from some real interval $D = (x_0, \infty)$, which may be open or closed. Let $\mu(x)$ and $\sigma^2(x)$ denote the mean and variance of the income distribution $x = (x_1, \dots, x_n)$, and let $X := \{x \mid x_i \in D; \mu(x) = \mu\}$ represent the set of feasible n -person income distributions with a common mean $\mu > 0$. Any two distributions $x, x' \in X$ differ by a *simple transfer* if $x' - x = \Delta(e_i - e_j)$ for some scalar $\Delta \neq 0$ and some $i \neq j$, where e_i denotes the n -tuple $(0, \dots, 0, 1, 0, \dots, 0)$ whose only non-zero element occurs in the i -th position. If, in addition, $x_j > x'_j > x_i$, then x' is said to be obtained from x by a *progressive transfer*. Conversely, x' is obtained from x by a *regressive transfer* iff x is obtained from x' by a progressive transfer.

A function $I: X \rightarrow \mathbb{R}$ will be called an *inequality index* iff $I(\cdot)$ is symmetric and strictly Schur-convex. It will also be convenient to refer to inequality measures as “Pigou–Dalton” indices, since any strictly S -convex $I(\cdot)$ satisfies the *Pigou–Dalton* “principle of transfers”:

$$I(x) > I(x') \quad \text{whenever } x' \text{ is obtained from } x \text{ by a progressive transfer} \quad (1)$$

or, equivalently,

$$I(x') > I(x) \quad \text{whenever } x' \text{ is obtained from } x \text{ by a regressive transfer.} \quad (2)$$

These minimal properties are sufficient to generate useful results.³ In particular, if one distribution x can be derived from another distribution $y \in X$ by a non-empty sequence

of progressive transfers, the Pigou–Dalton condition implies that each successive transfer reduces the inequality value, and hence that $I(x) < I(y)$. The pairwise inequality comparison is therefore determined “unambiguously”. However, if both progressive and regressive transfers are needed to convert y into x , the Pigou–Dalton criterion alone will not rank the two distributions. For (1) and (2) dictate the *direction* of the response to a progressive or a regressive transfer, but place no constraint on the *magnitude* of the changes, and hence no constraint on the net impact of a mixed sequence of transfers. This limitation of the Pigou–Dalton condition is overcome, at least in part, by the assumption of transfer sensitivity.

Transfer sensitivity has not been widely used in the past, and there is consequently no well established definition. But the basic idea is that the inequality assessment should give more emphasis to transfers taking place lower down in the distribution, other things being equal. The *ceteris paribus* clause has normally been interpreted as meaning that the transfers should be of equal magnitude and should occur between participants the same “income distance” apart.⁴ This is captured in the following “weak” definition of transfer sensitivity, which states that the inequality reduction resulting from a progressive transfer of a fixed size between participants the same distance apart should be inversely related to the income of the donor (and recipient).

Definition. An inequality measure $I(\cdot)$ defined on X is *weakly transfer sensitive* iff

$$I(x) - I(x + t) > I(x) - I(x + \tau) \tag{3}$$

for all $x \in X$ and all t, τ satisfying

$$t = \Delta(e_i - e_j); \quad \tau = \delta(e_k - e_l); \quad \Delta = \delta > 0; \tag{4a}$$

$$x_j - x_i = x_l - x_k > \Delta \tag{4b}$$

$$x_k > x_i. \tag{4c}$$

The implications of transfer sensitivity are easily derived for a variety of inequality measures whose functional forms are given explicitly. Consider, for instance, the large number of common indices that can be expressed as

$$I(x) = F(\sum_{i=1}^n f(x_i), \mu, n), \tag{5}$$

where F is strictly increasing in its first argument, and $f''(s) > 0$ for all $s \in D$. Substituting (5) and setting $\gamma = x_j - x_i = x_l - x_k$ transforms (3) and (4) into the statement

$$f(x_i) - f(x_i + \Delta) + f(x_i + \gamma) - f(x_i + \gamma - \Delta) > f(x_k) - f(x_k + \Delta) + f(x_k + \gamma) - f(x_k + \gamma - \Delta) \tag{6}$$

whenever $\gamma > \Delta > 0$ and $x_k > x_i$. This in turn is equivalent to the condition

$$f''(x_i) > f''(x_k) \quad \text{whenever } x_k > x_i. \tag{7}$$

Thus indices satisfying (5) are weakly transfer sensitive if and only if f'' is strictly decreasing.

This result establishes that all members of the Atkinson (1970) family of inequality measures are transfer sensitive,⁵ and that the “Generalized Entropy” indices, which correspond to $f(s) = s^c / c(c - 1)$, are transfer sensitive iff $c < 2$.⁶ Similar procedures can be used to determine whether other indices are transfer sensitive: the Gini coefficient for instance, does not have this feature. However, when the implications of weak transfer sensitivity are examined in the context of inequality measures whose functional form is

not restricted in advance, it becomes evident that no useful results are likely to be available. Conditions (4) place too many constraints on the progressive transfers whose impact is considered in (3). As a consequence, relatively few transfers satisfy the requirements of (4), and the assumption of weak transfer sensitivity has little power outside the structure imposed by a specific inequality index. There is therefore a need for a more general purpose definition of transfer sensitivity, which is equivalent to weak transfer sensitivity in the context of commonly used indices.

To accomplish this objective, we first observe that setting $y = x + \tau$ and $y' = x + t$ in (3), and noting that (4) implies

$$\sigma^2(y') - \sigma^2(y) = \sigma^2(x + t) - \sigma^2(x + \tau) = \frac{2\Delta}{n}(x_i - x_j - x_k + x_l) = 0,$$

enables conditions (3) and (4) to be restated in the equivalent form:

$$I(y) > I(y') \quad (8)$$

whenever $y, y' \in X$ satisfy

$$y' - y = \Delta(e_i - e_j) + \delta(e_l - e_k); \quad \Delta > 0; \quad \delta > 0; \quad (9a)$$

$$\sigma^2(y') = \sigma^2(y); \quad (9b)$$

$$y_i < y_j \cong y_k \cong y_l; \quad y'_i \cong y'_j \cong y'_k < y'_l; \quad (9c)$$

and

$$\Delta = \delta. \quad (10)$$

Here (9a) indicates that y' is obtained from y by a *composite transfer* which combines a transfer of the amount Δ from person j to person i with a simultaneous transfer of the amount δ from person k to person l . Given the additional restrictions contained in (9) and (10), weak transfer sensitivity stipulates that the operation converting y into y' has a *favourable* (i.e. inequality reducing) effect. It therefore seems appropriate to describe the operation as an example of a "favourable composite transfer".

A suitable general formulation of transfer sensitivity is now obtained by relaxing the restrictions imposed by (9) and (10), so that a wider variety of composite transfers are required to have a favourable impact. The precise form this relaxation should take is suggested by the observation that if y' is related to y by a composite transfer of the form (9a), and if y' is obtained from y by a *succession* of operations of the type described in (9) and (10), then $I(y')$ is necessarily less than $I(y)$, and y, y' necessarily satisfy (9). But (10) need not apply. We therefore drop condition (10) to arrive at:

Definition. An inequality measure $I(\cdot)$ defined on X is *transfer sensitive* iff (8) holds whenever $y, y' \in X$ satisfy (9).

Although transfer sensitivity is clearly stronger than weak transfer sensitivity, the additional content is minimal and it seems unlikely that an index will meet the "weaker" requirement without also fulfilling the stronger condition, except in pathological cases. The two versions are certainly equivalent in the context of all inequality measures in common use.⁷

Finally, we note that transfer sensitivity may be restated in an equivalent form by defining the complete set of favourable composite transfer operations as:

Definition. Let $y, y' \in X$. Then y' is obtained from y by a *favourable composite transfer* (or *FACT*) iff (9) holds.⁸

It then follows that an inequality measure $I(\cdot)$ is transfer sensitive iff

$$I(y) > I(y') \quad \text{whenever } y' \text{ is obtained from } y \text{ by a FACT.} \tag{11}$$

The similarity between this statement of transfer sensitivity and the traditional formulation (1) of the Pigou–Dalton condition is immediately apparent and helps us provide results for transfer sensitive inequality measures that are direct analogues of those available for Pigou–Dalton indices.⁹

3. UNAMBIGUOUS INEQUALITY COMPARISONS

It is well known that Pigou–Dalton indices will, in some circumstances, unanimously agree on their inequality ordering of a pair of distributions and thus produce a “conclusive” or “unambiguous” ranking. This happens when the two distributions have a common mean, are defined over the same sized population and have non-intersecting Lorenz curves. The central element in the characterisation of conclusive Pigou–Dalton verdicts is a result due to Hardy *et al.* (1934, p. 47), which established the circumstances under which one distribution can be transformed into another by a sequence of progressive transfers. In this section we derive a similar result for sequences of FACTs which allows us to characterise unambiguous transfer sensitive inequality comparisons. First we require some additional terminology and notation.

We let \hat{x} represent the *ordered version* of $x \in X$, defined by $\hat{x} = \Pi x$ for some permutation matrix Π such that $\hat{x}_1 \leq \hat{x}_2 \leq \dots \leq \hat{x}_n$, and say that x is *ordered* iff $x = \hat{x}$. An operation (such as a progressive transfer or FACT) which converts $x \in X$ into $y \in X$ is called *rank-preserving* iff there exists a common permutation matrix Π such that $\Pi x = \hat{x}$ and $\Pi y = \hat{y}$. The expression $X_k := \sum_{i=1}^k \hat{x}_i$ represents the partial sum of the k lowest incomes in x , and the Lorenz curve ordinate corresponding to the lowest 100*p*% of income recipients is denoted by $L(x; p)$ for $p \in [0, 1]$. A distribution x *Lorenz dominates* y iff

$$L(x; p) \geq L(y; p) \quad \forall p \in [0, 1] \quad \text{and} \quad > \quad \text{for some } p \tag{12}$$

or, in other words, iff the Lorenz curve of x lies nowhere below, and somewhere above, that of y . We also need to define

$$\phi(x; z) := \sum_{i=1}^{q(x; z)} (z - \hat{x}_i) = \sum_{i=1}^{q^-(x; z)} (z - \hat{x}_i) \tag{13}$$

$$\Phi(x; z) := \sum_{i=1}^{q(x; z)} (z - \hat{x}_i)^2 = \sum_{i=1}^{q^-(x; z)} (z - \hat{x}_i)^2 \tag{14}$$

where

$$q(x; z) := \max \{i | \hat{x}_i \leq z\} \tag{15}$$

$$q^-(x; z) := \max \{i | \hat{x}_i < z\}$$

represent the number of incomes in x that do not exceed z , or are strictly less than z , respectively. The expressions for $\phi(x; z)$ and $\Phi(x; z)$ are related to the standard characterisations of second-order stochastic dominance (SSD) and third-order stochastic dominance (TSD) applied to discrete distributions. In particular, if $x, y \in X$, then x SSD y iff

$$\phi(y; z) \geq \phi(x; z) \quad \forall z \in D \quad \text{and} \quad > \quad \text{for some } z; \tag{16}$$

and x TSD y iff

$$\Phi(y; z) \geq \Phi(x; z) \quad \forall z \in D \quad \text{and} \quad > \quad \text{for some } z.^{10} \tag{17}$$

Finally, it will be helpful to note that, for any given distributions $x, y \in X$, the function

$$g(z) := \phi(y; z) - \phi(x; z) \tag{18}$$

is continuous and piecewise linear, with right and left derivatives given respectively by

$$g^+(z) = q(y; z) - q(x; z) \tag{19}$$

and

$$g^-(z) = q^-(y; z) - q^-(x; z). \tag{20}$$

Furthermore

$$g(z) = 0 \quad \forall z \leq \zeta_* := \min(\hat{x}_1, \hat{y}_1) \tag{21}$$

$$g(z) = 0 \quad \forall z \geq \zeta^* := \max(\hat{x}_n, \hat{y}_n)$$

and

$$G(z) := \Phi(y; z) - \Phi(x; z) = 2 \int_{x_0}^z g(\zeta) d\zeta. \tag{22}$$

We are now in a position to state, without proof, the following well known result:

Theorem 1. *For $x, y \in X$, the following statements are equivalent:*

- (T1a) \hat{x} can be obtained from \hat{y} by a non-empty finite sequence of rank-preserving progressive transfers.
- (T1b) x SSD y .
- (T1c) $I(y) > I(x)$ for all Pigou-Dalton indices $I(\cdot)$
- (T1d) $X_k \geq Y_k \quad \forall k$ and $>$ for some k .
- (T1e) x Lorenz dominates y .

Statement (T1c) indicates that the set of Pigou-Dalton indices unanimously agree on their inequality ranking of the distributions x and y , thus producing an unambiguous verdict. The significance of Theorem 1 is that it provides the equivalent conditions given in (T1b), (T1d) or (T1e) which enable us to check whether an unambiguous ranking applies. A corresponding result for transfer sensitive inequality measures is provided by the following theorem:

Theorem 2. *For $x, y \in X$, the following statements are equivalent:*

- (T2a) \hat{x} can be obtained from \hat{y} by a non-empty finite sequence of rank-preserving progressive transfers and/or favourable composite transfers.
- (T2b) x TSD y .
- (T2c) $I(y) > I(x)$ for all transfer sensitive inequality measures $I(\cdot)$.

Proof. See the Appendix.

Although Theorem 2 is analogous to Theorem 1, there are several significant differences. First, the operations converting \hat{y} into \hat{x} in (T2a) may comprise both progressive transfers and FACTs, although the obvious analogue of (T1a) would suggest that only FACTs be used. It is a simple matter to determine the circumstances under which two

distributions are linked by a sequence of FACTs alone. For by noting that the variance of a distribution is preserved by a FACT and reduced by a progressive transfer, it becomes evident that the sequence in (T2a) contains only FACTs iff $\sigma^2(x) = \sigma^2(y)$. This observation also reveals that (T2c) holds only if $\sigma^2(y) \geq \sigma^2(x)$.

Another respect in which Theorem 2 differs from Theorem 1 concerns the absence of conditions corresponding to (T1c) or (T1d). No analogue of ‘‘Lorenz dominance’’ is known to be equivalent to the ‘‘third-order stochastic dominance’’ condition (T2b). This could be troublesome, since it is not immediately obvious that an algorithm exists for establishing the validity of (T2b). However a simple procedure can be constructed.¹¹

4. SINGLE CROSSING LORENZ CURVES

The overall impact of moving from the unambiguous inequality ordering generated by Pigou–Dalton indices to that based on transfer sensitive inequality indices is an increase in the power of the ranking criterion. The significance of this improvement in power has yet to be fully evaluated. However some indication may be obtained by examining situations in which Pigou–Dalton indices disagree in their assessment of two distributions $x, y \in X$. Theorem 1 tells us that this will occur when the Lorenz curves of x and y intersect. Although multiple crossings of Lorenz curves can occur, in practice they are less common than single intersections. In this section we examine the implications of the unambiguous transfer sensitive inequality ordering in the context of distributions whose Lorenz curves cross just once.

The Lorenz curve for a distribution x will be said to intersect that of y *once from above* iff there exists $p^* \in (0, 1)$ and intervals $P := [0, p^*]$ and $P' := [p^*, 1]$ such that

$$\begin{aligned} L(x; p) &\geq L(y; p) \quad \forall p \in P \quad \text{and} \quad > \quad \text{for some } p \in P \\ L(x; p) &\leq L(y; p) \quad \forall p \in P' \quad \text{and} \quad < \quad \text{for some } p \in P'. \end{aligned} \tag{23}$$

Theorem 3. *Suppose the Lorenz curve of $x \in X$ intersects that of $y \in X$ once from above. Then*

$$(T3a) \quad I(y) > I(x) \text{ for all transfer sensitive inequality measures } I(\cdot)$$

if and only if

$$(T3b) \quad \sigma^2(y) \geq \sigma^2(x).$$

Proof. If $x, y \in X$ and $\zeta \geq \max \{\hat{x}_n, \hat{y}_n\}$,

$$n\sigma^2(y) - n\sigma^2(x) = \sum_{i=1}^n (\zeta - y_i)^2 - \sum_{i=1}^n (\zeta - x_i)^2 = \Phi(y; \zeta) - \Phi(x; \zeta). \tag{24}$$

Hence (T3a) and Theorem 2 immediately yield (T3b).

For the converse, assume (T3b) is true and note that (23) ensures that $\hat{x} \neq \hat{y}$, and that there exists $j < n$ such that

$$X_i \geq Y_i \quad \text{for } i \leq j; \quad X_i \leq Y_i \quad \text{for } i > j. \tag{25}$$

Now suppose (T3a) is false. Then $\hat{x} \neq \hat{y}$ and Theorem 2 together imply $\Phi(y; z_0) < \Phi(x; z_0)$ for some z_0 . Using the properties of the functions $g(z)$ and $G(z)$ described in (18)–(22), we deduce that there exists $z_1 < z_2$ such that

$$g(z_1) = 0 > g^+(z_1); \quad g(z_2) = 0 < g^+(z_2),$$

since $G(z_0) < 0 \leq G(\zeta)$. Define

$$q := q(y; z_1) < r := q(x; z_1) \leq s := q(x; z_2) < t := q(y; z_2). \tag{26}$$

Then $z_1 < \hat{y}_i$ for $q < i \leq r$ and $z_2 < \hat{x}_i$ for $s < i \leq t$, so

$$\begin{aligned} X_r - Y_r &= \sum_{i=1}^r (z_1 - \hat{y}_i) - \sum_{i=1}^r (z_1 - \hat{x}_i) \\ &= g(z_1) + \sum_{i=q+1}^r (z_1 - \hat{y}_i) < 0. \end{aligned}$$

Similarly

$$X_t - Y_t = g(z_2) + \sum_{i=s+1}^t (\hat{x}_i - z_2) > 0.$$

From (25) it follows that $r > j \geq t$, which contradicts (26). Hence (T3a) must hold, and the proof is complete. \parallel

Theorem 3 will not resolve all the ambiguous rankings associated with single crossing Lorenz curves, since the distribution with the lower variance must also have the Lorenz curve that is initially higher if all transfer sensitive inequality measures are to agree on their verdict. It will, however, assist in a large number of pairwise comparisons for which the Pigou–Dalton judgement is inconclusive. Furthermore, although this paper has focussed on the set of distributions X with a fixed mean μ and population size n , it is easy to see how Theorem 3 can be extended to a comparison of distributions with different means and population sizes.

Let $\mathcal{X} := \{x \mid x \in D^n \text{ for some finite } n; \mu(x) > 0\}$ represent the set of all distributions with a positive mean and finite population size. Any inequality measure $I(\cdot)$ defined on X can be extended to the larger set \mathcal{X} by assuming $I(\cdot)$ is both *scale invariant* and *replication invariant*, so that the inequality value remains unchanged if all incomes are subjected to a (positive) proportional change, or if the distribution is replicated any number of times.¹² To obtain the analogue of Theorem 3, we note that the Lorenz curves are unaffected by replication or a scale change of the distribution. In addition, although the variance is affected by a change of scale, the value of the coefficient of variation, σ/μ , is invariant to both scale changes and replication of the distribution. It therefore follows from Theorem 3 that:

Corollary 1. *If the Lorenz curve of $x \in \mathcal{X}$ intersects that of $y \in \mathcal{X}$ once from above*

(C4a) *$I(y) > I(x)$ for all inequality measures $I(\cdot)$ satisfying transfer sensitivity, scale invariance and replication invariance*

if and only if

(C4b)
$$\frac{\sigma(y)}{\mu(y)} \geq \frac{\sigma(x)}{\mu(x)}.$$

Atkinson (1973) gives some indication of the degree to which Corollary 1 enhances the ability to rank distributions. Using the Kuznets' data employed in his 1970 article, Atkinson finds that 24% of a possible 66 pairwise country comparisons can be ranked on the basis of Lorenz-dominance, while a further 71% involve single-crossing Lorenz curves. Corollary 1 raises the ranking success rate substantially, from 24% to 44%. Davies and Hoy (1985) perform a similar exercise on a group of OECD countries, on time series data for the U.S.A., and on age-groups and provinces within Canada. Their results confirm the ability of transfer sensitivity to resolve a significant proportion of the ambiguous Pigou–Dalton comparisons. The rate of success in ranking income distributions is raised by an average of one third by Corollary 1 alone, and by around 50% when multiple Lorenz curve intersections are also included.

APPENDIX

Proof of Theorem 2. (a) (T2a)⇒(T2c). The value of any transfer sensitive inequality measure $I(\cdot)$ is preserved by any permutation of the distribution (by symmetry) and reduced by a progressive transfer (by the Pigou–Dalton condition) or FACT (by transfer sensitivity). Hence for any transfer sensitive inequality measure $I(\cdot)$, (T2a) implies $I(y) = I(\hat{y}) > I(\hat{x}) = I(x)$.

(b) (T2c)⇒(T2b). Suppose (T2c) holds and define

$$V := \{v(s) \mid v'(s) > 0, v''(s) < 0 \text{ and } v'''(s) > 0 \forall s \in D\}.$$

Then all indices of the form

$$I(x) = -\frac{1}{n} \sum_{i=1}^n v(x_i), \quad v \in V$$

are continuous, and satisfy the Pigou–Dalton condition (since $v'' < 0$) and transfer sensitivity (since $v''' > 0$). (T2c) therefore implies

$$\frac{1}{n} \sum_{i=1}^n v(x_i) > \frac{1}{n} \sum_{i=1}^n v(y_i) \quad \forall v \in V. \tag{27}$$

Condition (27) corresponds to the statement that x TSD y (Bawa (1975), Theorem 3), so (17) and hence (T2b) follow immediately.

(c) (T2b)⇒(T2a). The derivation of (T2a) revolves around the properties of the functions $g(z)$ and $G(z)$ described in (18)–(22). We begin by noting that (T2b) implies

$$\Phi(\hat{y}; z) \geq \Phi(\hat{x}; z) \quad \forall z \in D \text{ and } > \text{ for some } z,$$

so we may assume WLOG that x and y are ordered.

Now (T2b) implies $x \neq y$ and

$$G(z) := \Phi(y; z) - \Phi(x; z) \geq 0 \quad \forall z \in D. \tag{28}$$

When

$$g(z) := \phi(y; z) - \phi(x; z) \geq 0 \quad \forall z \in D, \tag{29}$$

Theorem 1 ensures the existence of a non-empty finite sequence of rank-preserving progressive transfers converting y into x as required. Suppose, therefore, that

$$g(z) < 0 \text{ for some } z \in D. \tag{30}$$

If (28) and (30) hold we can construct a distribution y' , obtained from y by a rank-preserving FACT, such that (28) holds when y' is substituted for y . Formally, we will establish:

Lemma 1. *Let $x, y \in X$ be any ordered distributions satisfying (28) and (30). Then $Z(x, y) := \{z \mid g(z) = 0; g^-(z) > 0\}$ and $N(x, y) := \{i \mid y_i < \min Z(x, y); y_i \neq x_i\}$ are both non-empty finite sets. Furthermore, there exists an ordered distribution y' , obtained from y by a rank-preserving FACT, such that*

$$\Phi(y'; z) - \Phi(x, z) \geq 0 \quad \forall z \in D, \tag{31}$$

and either

$$Z(x, y') \subset Z(x, y) \tag{32a}$$

or else

$$Z(x, y') = Z(x, y) \text{ and } N(x, y') \subset N(x, y). \tag{32b}$$

Successive application of Lemma 1 generates a sequence of ordered distributions $y = y^0, y^1, y^2, \dots, y^K$ where y^K is obtained from y^{K-1} by a rank-preserving FACT. Since $Z(x, y)$ and $N(x, y)$ are finite sets, (32) ensures that the process terminates in a finite number of steps $K \geq 1$, and this occurs when

$$\phi(y^K; z) - \phi(x; z) \geq 0 \quad \forall z \in D.$$

Then either $y^K = x$, in which case the proof is complete; or else by Theorem 1 there exists a further finite sequence of rank-preserving progressive transfers which converts y^K into x as required. \parallel

Proof of Lemma 1. Consider the properties of the functions $g(z)$ and $G(z)$ described in (18)–(22). We begin by noting that all local minima of $G(z)$ in the interval (ζ_*, ζ^*) are achieved at points in the set $Z(x, y)$. Conditions (21) and (22) therefore enable us to state

$$\Phi(y; z) - \Phi(x; z) \geq 0 \quad \forall z \in D \quad \text{iff} \quad \Phi(y; z) - \Phi(x; z) \geq 0 \quad \forall z \in Z(x, y). \tag{33}$$

Assumption (30) and continuity ensure that g is negative over a non-zero and finite number of bounded open intervals, so the set $Z(x, y)$ is non-empty, finite and bounded. Let $\zeta := \min Z(x, y)$; $s := q^-(x; \zeta)$ and $t := q^-(y; \zeta)$. Then $g^-(\zeta) = t - s > 0$, $y_t < \zeta \leq x_t$, and $g(y_t) < g(\zeta) = 0$. In addition, $t \in N(x, y) = \{p \mid y_p < \zeta; y_p \neq x_p\}$, so $N(x, y)$ is non-empty and finite.

Now define $j := \min \{p \mid y_p > x_p\}$, which exists because $X_n = Y_n$ and $x \neq y$ (to satisfy (30)). Let $i := \max \{p < j \mid y_p < x_p\}$, which exists since otherwise $G(z) < 0$ for $z \in (x_j, y_j)$, in contradiction to (28). Then $i < j$; $y_p \leq x_p$ if $p < j$; $y_j > x_j$; $g(z) \geq 0 \quad \forall z \leq x_j$; and hence, $g(y_p) \geq 0 \quad \forall p < j$. Since $y_t < x_t$ and $g(y_t) < 0$, it follows that $j < t$. Therefore $m := \min \{p > j \mid y_p < x_p\} \leq t$ and $l := \min \{p \geq m \mid y_{p+1} \geq x_{p+1}\} \cup \{t\}$ both exist, and we have

$$\begin{aligned} y_p \leq x_p & \quad \text{if } p < i; & y_t < x_t; & y_p = x_p & \quad \text{if } i < p < j; & y_j > x_j; \\ y_p \geq x_p & \quad \text{if } j < p < m; & y_p < x_p & \quad \text{if } m \leq p \leq l; & \text{and} & y_l \leq y_t < \zeta. \end{aligned} \tag{34}$$

In particular, there exist $i < j < l$ for which

$$y_i < x_i \leq x_{i+1} \leq y_{i+1}; \quad y_j > x_j \geq x_{j-1} \geq y_{j-1}; \quad y_l < \bar{\zeta} := \min \{x_l, \zeta\} \tag{35}$$

and also $\bar{\zeta} \leq y_{l+1}$, unless $l = t = n$.

This construction has identified three income levels y_i, y_j and y_l which can be moved towards x_i, x_j and $\bar{\zeta}$, respectively, without ‘‘overtaking’’ adjacent values in the distribution y . A composite transfer which transforms y into another ordered distribution y' will be rank-preserving, therefore, if

$$y'_i > y_i; \quad y'_j < y_j; \quad y'_l > y_l; \quad y'_p = y_p \quad \forall p \neq i, j, l; \tag{36}$$

$$y'_i \leq x_i; \quad y'_j \geq x_j; \quad y'_l \leq \bar{\zeta}; \tag{37}$$

and will satisfy the requirements (9) of a FACT if, in addition,

$$\bar{\mu} := \mu(y'_i, y'_j, y'_l) = \mu(y_i, y_j, y_l), \tag{38}$$

$$\bar{\sigma}^2 := \sigma^2(y'_i, y'_j, y'_l) = \sigma^2(y_i, y_j, y_l). \tag{39}$$

To obtain a suitable transformed distribution y' , define

$$\psi_1(\theta) := \bar{\mu} - \sqrt{2} \bar{\sigma} \cos \theta$$

$$\psi_2(\theta) := \bar{\mu} - \sqrt{2} \bar{\sigma} \cos \left(\frac{\pi}{3} + \theta \right)$$

$$\psi_3(\theta) := \bar{\mu} - \sqrt{2} \bar{\sigma} \cos \left(\frac{\pi}{3} - \theta \right)$$

where $\theta \in [0, 2\pi)$. Then

$$\psi_1(\theta) + \psi_2(\theta) + \psi_3(\theta) = y_i + y_j + y_l \tag{40a}$$

$$\psi_1(\theta)^2 + \psi_2(\theta)^2 + \psi_3(\theta)^2 = y_i^2 + y_j^2 + y_l^2 \tag{40b}$$

for all θ , and

$$\psi_1(\theta) < \psi_2(\theta) \leq \psi_3(\theta) \quad \text{iff } 0 \leq \theta < \pi/3 \tag{41a}$$

$$\psi_1(\theta) \leq \psi_2(\theta) < \psi_3(\theta) \quad \text{iff } 0 < \theta \leq \pi/3. \tag{41b}$$

It follows that $(y_i, y_j, y_l) = (\psi_1(\alpha), \psi_2(\alpha), \psi_3(\alpha))$ for some unique $\alpha \in [0, \pi/3)$. In addition, since $d\psi_1/d\theta > 0$, $d\psi_2/d\theta < 0$ and $d\psi_3/d\theta > 0$ when $\theta \in (0, \pi/3)$, there exists $\theta \in (\alpha, \pi/3]$ such that

$$y_i = \psi_1(\alpha) < \psi_1(\theta) \leq x_i; \quad y_j = \psi_2(\alpha) > \psi_2(\theta) \geq x_j; \quad y_l = \psi_3(\alpha) < \psi_3(\theta) \leq \bar{\zeta}.$$

Hence

$$\theta' := \max \left\{ \theta \leq \frac{\pi}{3} \mid \psi_1(\theta) \leq x_i; \psi_2(\theta) \geq x_j; \psi_3(\theta) \leq \bar{\zeta} \right\} \tag{42}$$

exists and $\theta' > \alpha$. Setting $y'_i = \psi_1(\theta')$, $y'_j = \psi_2(\theta')$, $y'_l = \psi_3(\theta')$ and $y'_p = y_p \forall p \neq i, j, l$, we obtain a distribution y' satisfying all the conditions (36)–(39). So $y' = \hat{y}'$ is obtained from y by a rank-preserving FACT. Furthermore, if $\theta' < \pi/3$, (42) implies either $y'_i = \psi_1(\theta') = x_i$, or $y'_j = \psi_2(\theta') = x_j$, or $y'_l = \psi_3(\theta') = \xi$; and if $\theta' = \pi/3$, we have $y'_i = y'_j = \bar{\mu} - \bar{\sigma}/\sqrt{2}$ and hence $y'_i = x_i = x_j = y'_l$, since $y'_i \leq x_i \leq x_j \leq y'_l$. Therefore, at least one of the inequality constraints in (37) must bind, and

$$N' := \{p \mid y'_p < \xi; y_p \neq x_p\} \subset N(x, y). \tag{43}$$

Next consider condition (32). From (34)–(36) we deduce

$$\begin{aligned} y'_p \leq x_p & \text{ if } p < j; & y'_p \geq x_p & \text{ if } j \leq p < m; & y_p = y'_p < x_p & \text{ if } m \leq p < l; \\ y_l < y'_l \leq \xi & := \min\{x_l, \xi\}; & \text{ and } & y'_p = y_p & \text{ if } p > l. \end{aligned} \tag{44}$$

Hence, employing h and h as the analogues of g^- and g ,

$$\begin{aligned} h(z) &:= \phi(y'; z) - \phi(x; z) \geq 0 \quad \forall z \leq x_j \\ h^-(z) &:= q^-(y'; z) - q^-(x; z) \leq 0 \quad \forall z \in (x_j, y_m] \\ h^-(z) &> 0 \quad \forall z \in (y_m, y'_l] \end{aligned} \tag{45}$$

and

$$h(z) - g(z) = \begin{cases} y_j - y'_j = y'_l - y_l > 0 & z \in [y_m, y_l] \\ y'_l - y_l - (z - y_l) > 0 & z \in [y_l, y'_l] \\ 0 & z \geq y'_l. \end{cases} \tag{46}$$

From (45) and (46) we obtain

$$g(y_m) < h(y_m) < h(y'_l) = g(y'_l),$$

and hence

$$h(z) \neq 0 \quad \forall z \in [y_m, y'_l], \tag{47}$$

otherwise $\xi := \min\{z \mid g(z) = 0; g^-(z) = 0\} < y'_l$, in contradiction to (44). Now if $Z(x, y') = \{z \mid h(z) = 0; h^-(z) > 0\} \neq \emptyset$, (32) holds trivially. So suppose $\xi' \in Z(x, y')$ and define $r := q^-(y'; \xi')$. Then

$$h(\xi') = 0; \quad h^-(\xi') > 0; \quad \text{and} \quad h(z) < 0 \quad \forall z \in (y'_l, \xi'). \tag{48}$$

Conditions (45), (47), and (48) imply $\xi' \geq y'_l$. Since, by (46), $h(z) = g(z) \forall z \geq y'_l$, it follows that $\xi' \in Z(x, y)$ and hence $Z(x, y') \subseteq Z(x, y)$. Furthermore either (32a) holds, or else $Z(x, y) = Z(x, y')$. In the latter case, $\xi = \min Z(x, y')$ and, from (44), $N(x, y') = N' \subset N(x, y)$, so (32b) is true.

Finally we turn to condition (31), which is trivially satisfied if $y' = x$. If $Z(x, y') = \emptyset$ and $y' \neq x$, (33) ensures that (31) holds. Alternatively suppose $Z(x, y') \neq \emptyset$ and $y' \neq x$. It has been shown that $z \in Z(x, y')$ implies $z \geq \xi \geq y'_l$. Hence, for all $z \in Z(x, y')$,

$$\Phi(y'; z) - \Phi(y; z) = (z - y'_l)^2 + (z - y'_j)^2 + (z - y'_i)^2 - (z - y_l)^2 - (z - y_j)^2 - (z - y_i)^2 = 0$$

using (38) and (39). Thus, using (28),

$$\begin{aligned} \Phi(y'; z) - \Phi(x; z) &= G(z) + \Phi(y'; z) - \Phi(y; z) \\ &= G(z) \geq 0 \quad \forall z \in Z(x, y'), \end{aligned}$$

and (31) follows from (33). This completes the proof of Lemma 1. \parallel

Acknowledgement We should like to thank A. B. Atkinson, J. B. Davies, M. Lebreton and A. K. Sen for helpful suggestions. An earlier version of the paper was presented at the Econometric Society Meetings in Madrid, 1984.

NOTES

1. See, for example, Atkinson (1970, p. 256), Sen (1973, p. 28), Love and Wolfson (1976, pp. 59–63), Kolm (1976, pp. 87–88), Cowell (1977, pp. 30–31), Kakwani (1980, pp. 438–439) and Foster *et al.* (1984, p. 763). “Transfer sensitivity” is the term employed by Kakwani and Foster *et al.* in the context of poverty measurement, while Kolm refers to the same property as the “Principle of Diminishing Transfers”.

2. Similar suggestions are contained in two other unpublished papers that have recently come to our attention. Atkinson (1973), building on Kolm's "Principle of Diminishing Transfers", considers the impact of a pair of progressive and regressive transfers which together imply movement of income away from the middle of the distribution. By making explicit the analogy between the Principle of Diminishing Transfers and the Pigou-Dalton condition, and by exploiting their respective connection with third-order and second-order stochastic dominance, Atkinson anticipates many of our arguments and results. In particular, the substance of our Theorem 3, concerned with single-crossing Lorenz curves, is examined in some detail.

Davies and Hoy (1984) discuss the notion of "aversion to downside risk" developed by Menezes *et al.* (1980), and propose an analogous concept of "aversion to downside inequality". This is defined in terms of composite transfers which preserve both the mean and the variance of the distribution, as is the case with our FACTs.

3. The results contained in this paper are preserved if the inequality indices are assumed to be continuous and/or are normalized to zero for distributions with equal incomes.

4. See, for instance, Sen (1973, p. 28), Kolm (1976, p. 87), Cowell (1977, pp. 30-31), Kakwani (1980, p. 439) and Foster *et al.* (1984, p. 763).

5. These may be expressed in the form given by (5) by setting $F(s, \mu, n) = 1 - [s(\varepsilon - 1)/n]^{1/(1-\varepsilon)}/\mu$ and $f(s) = s^{1-\varepsilon}/(\varepsilon - 1)$, for $\varepsilon > 0$.

6. For the general class of additively decomposable or "Generalized Entropy" indices, see Cowell and Kuga (1981) or Shorrocks (1980, 1984). The coefficient of variation (and its square) correspond to $c = 2$, and are therefore not transfer sensitive. They have the property that a transfer of £1 between millionaires whose incomes differ by £1 million has the same impact on inequality as a £1 transfer from an individual with an income of £1 million to someone with zero income. It is this unattractive feature of the coefficient of variation that has prompted much of the interest in a transfer sensitivity condition: see, for example, Atkinson (1970, pp. 255-256), Sen (1973, pp. 27-28), Love and Wolfson (1976, pp. 59-63), Kolm (1976, p. 87) and Cowell (1977, pp. 30-31).

7. For example, in the context of indices of the form given by (5), it is relatively easy to show that condition (7) implies transfer sensitivity, and hence that transfer sensitivity and weak transfer sensitivity are equivalent.

8. Notice that (9c) implies that $i \neq j, k, l$ and $l \neq i, j, k$, although it is possible that $j \neq k$. The composite transfer described in (9) evidently involves a *progressive* transfer between persons i and j and a *regressive* transfer between persons k and l . Furthermore, the progressive exchange occurs lower down in the distribution than the regressive exchange.

9. The assumption that $I(\cdot)$ respects the Pigou-Dalton condition has played no part in the discussion concerning the definition of transfer sensitivity, and as a consequence could have been omitted. However it makes little sense to assume an index is transfer sensitive if the Pigou-Dalton criterion is violated. Indeed there is a strong temptation to regard transfer sensitivity as a generalization of the Pigou-Dalton condition, and this can be formally accomplished by expanding the set (9) of FACTs to include progressive transfers.

10. These characterisations of stochastic dominance follow from Bawa (1975), Theorems 2 and 3, since if $F_x(s)$ is the distribution function associated with $x \in X$,

$$\phi(x; z) = n \int_0^z F_x(s) ds$$

and

$$\Phi(x; z) = 2n \int_0^z \int_0^t F_x(s) ds dt.$$

Similar expressions also arise in the discussion of certain poverty orderings: see Foster and Shorrocks (1985).

11. See, for example, Bawa (1975, pp. 106-109). Note that it is *not* the case that x TSD y holds iff

$$\int_0^p [L(x; q) - L(y; q)] dq \geq 0 \quad \forall p \in [0, 1] \quad \text{and} \quad > \quad \text{for some } p \in [0, 1],$$

as some authors have claimed: see, for instance, Levy and Kroll (1978, p. 569). A simple counterexample is provided by the distributions $x = (4, 4, 4, 10)$ and $y = (3, 3, 8, 8)$.

12. See, for example, Shorrocks (1984, p. 1372). Scale invariance is best considered in the context of the income range $D = [0, \infty)$ or $D = (0, \infty)$, so that the set of feasible distributions is closed under multiplication by a positive scalar.

REFERENCES

- ATKINSON, A. B. (1970), "On the Measurement of Inequality", *Journal of Economic Theory*, **2**, 244-263.
 ATKINSON, A. B. (1973), "More on the Measurement of Inequality" (mimeo).
 BAWA, V. S. (1975), "Optimal Rules for Ordering Uncertain Prospects", *Journal of Financial Economics*, **2**, 95-121.
 COWELL, F. A. (1977) *Measuring Inequality* (Oxford: Philip Allan).
 COWELL, F. A. and KUGA, K. (1981), "Additivity and the Entropy Concept: An Axiomatic Approach to Inequality Measurement", *Journal of Economic Theory*, **25**, 131-143.

- DAVIES, J. and HOY, M. (1985), "Comparing Income Distributions Under Aversion to Downside Inequality" (mimeo).
- FOSTER, J. E., GREER, J. and THORBECKE, E. (1984), "A Class of Decomposable Poverty Measures", *Econometrica*, **52**, 761-766.
- FOSTER, J. E. and SHORROCKS, A. F. (1985), "Poverty Orderings and Welfare Dominance" (Discussion Paper No. 878, Purdue University).
- HARDY, G. H., LITTLEWOOD, J. E. and POLYA, G. (1934) *Inequalities* (Cambridge: Cambridge University Press).
- KAKWANI, N. (1980), "On a Class of Poverty Measures", *Econometrica*, **48**, 437-446.
- KOLM, S-Ch. (1976), "Unequal Inequalities II", *Journal of Economic Theory*, **13**, 82-111.
- LEVY, H. and KROLL, Y. (1978), "Ordering Uncertain Options with Borrowing and Lending", *Journal of Finance*, **33**, 553-574.
- LOVE, R. and WOLFSON, M. C. (1976) *Income Inequality: Statistical Methodology and Canadian Illustrations* (Ottawa: Statistics Canada).
- MENEZES, C., GEISS, C. and TRESSLER, J. (1980), "Increasing Downside Risk", *American Economic Review*, **70**, 921-932.
- SEN, A. K. (1973) *On Economic Inequality* (Oxford: Oxford University Press).
- SHORROCKS, A. F. (1980), "The Class of Additively Decomposable Inequality Measures", *Econometrica*, **48**, 613-625.
- SHORROCKS, A. F. (1984), "Inequality Decomposition by Population Subgroups", *Econometrica*, **52**, 1369-1385.