# Inequality and Income Gaps

*Ian Preston*

THE INSTITUTE FOR FISCAL STUDIES
WP06/25

# Inequality and Income Gaps [*]

Ian Preston [†]
University College London and Institute for Fiscal Studies

November 2006

**Abstract**

This paper discusses inequality orderings based explicitly on closing up of income gaps, demonstrating the links between these and other orderings, the classes of functions preserving the orderings and applications showing their usefulness in comparison of economic policies.

JEL: D31, D63
Keywords: Inequality, income distribution

[*]I am grateful for helpful comments from Tim Besley, Chris Gilbert, Terence Gorman, Chris Harris, Peter Lambert, James Mirrlees, Stephen Nickell, Hyun Shin, seminar participants and an anonymous referee. The paper draws in large part upon my doctoral thesis, for the funding of which I am grateful to the Economic and Social Research Council.

[†]Address: Ian Preston, Department of Economics, University College London, Gower Street, London WC1E 6BT, UK. Email: i.preston@ucl.ac.uk

**Executive Summary**

It is a truism to say that inequality is about gaps between incomes and that reducing inequality is about closing these gaps up.

Common means of comparison between income distributions all use criteria which do show inequality as falling when gaps close. However explicitly asking whether the gaps reduce throughout the whole distribution in concertina-like fashion is a rare criterion to apply. This paper seeks to investigate the related orderings. The most common criteria for inequality comparison are those based on Lorenz curves, made plausible most persuasively as indicators of inequality by their link to progressive transfers of income. Progressive transfers are often seen, since the arguments of Pigou and Dalton, as uncontentiously inequality reducing but this view could be challenged if there are more than two people. A transfer from the top to the middle of the income distribution reduces inequality between the top and the middle but increases it between the middle and the bottom. Regarding inequality as having fallen overall involves giving priority to the former effect - the effect on the gap between incomes of those involved directly in the transfer - for which there may be good reason, but it is not obvious that it would not be sensible to say inequality simply could not be compared. A minor function of the current paper is to bring some overlooked but highly germane mathematical literature to the attention of inequality theorists. The major function, though, is to tell a rounded story about the ratio and difference dominance concepts, and associated orderings and welfare properties.

# 1 Introduction

It is a truism to say that inequality is about gaps between incomes and that reducing inequality is about closing these gaps up. Common means of comparison between income distributions all use criteria which do show inequality as falling when gaps close. However explicitly asking whether the gaps reduce throughout the whole distribution in concertina-like fashion is a rare criterion to apply. This paper seeks to investigate the related orderings.

The most common criteria for inequality comparison are those based on Lorenz curves, made plausible most persuasively as indicators of inequality by their link to progressive transfers of income. Progressive transfers are often seen, since the arguments of Pigou (1912) and Dalton (1920), as uncontentiously inequality reducing but this view could be challenged if there are more than two people. A transfer from the top to the middle of the income distribution[1] reduces inequality between the top and the middle but increases it between the middle and the bottom[2]. Regarding inequality as having fallen overall involves giving priority to the former effect - the effect on the gap between incomes of those involved directly in the transfer - for which there may be good reason, but it is not obvious that it would not be sensible to say inequality simply could not be compared[3].

What convinces Dalton (1920) is the link to welfare - he is "primarily interested, not in the distribution as such, but in the effects of the distribution of income upon the distribution and total amount of economic welfare, which may be derived from income (p.348)". That such transfers raise welfare is well known to be true in a typical utilitarian setting if individual welfare depends only upon own income but if income gaps matter to individual welfare then this need not be so. This issue is taken up below and links between economic welfare and the orderings based explicitly on gaps are considered.

The earliest discussions of these orderings can be found outside the economic context (for example in Marshall, Olkin and Proschan (1967) and Bar-

[1] Progressive transfers are sometimes called Robin Hood transfers. The Robin Hood of legend stole from the rich to give to the poor. Noone would disagree that that reduces inequality. He never, however, stole from the rich to give to the middle or stole from the middle to give the poor.

[2] Blum and Kalven (1953), for example, discuss income redistribution in this sort of way.

[3] One might raise a similar objection to Sen's (1976, 1978) strengthened criterion for comparisons of ordinal inequality for cases of more than two persons - STOIC. Recognising this makes it clearer why he finds a link with Lorenz dominance.

low and Proschan (1975)). There are some useful papers from this period which are less well known than perhaps they should be and a minor function of the current paper is to bring some overlooked but highly germane mathematical literature to the attention of inequality theorists. I think particularly here of Marshall, Walkup and Wets (1967) which anticipates several results of this paper[4].

The major function, though, is to tell a rounded story about the ratio and difference dominance concepts, and associated orderings and welfare properties, which extend the well-known Lorenz ordering in different ways. In this, I am in fact taking up again some work which I engaged in some time ago (Preston 1989, 1990a, 1990b) and ideas which have been developed by Moyes (1994) on the "dominance in relative differentials" and "dominance in absolute differentials" concepts, for which he coined those terminologies, and by Zheng (2006). The style and manner of development, in the sequel, is intended to be somewhat in similar fashion to the way in which Rothschild and Stiglitz's (1973) paper developed a rounded story for the Lorenz ordering, which had been begun by Kolm (1969) and Atkinson (1970) (see also Dasgupta, Sen and Starrett 1973).

Section 2 defines the orderings and considers relations between them. Section 3 outlines classes of functions which preserve the orderings. Section 4 considers how policies map underlying variation into distributions which may be related according to the orderings. Section 5 concludes.

# 2 Dominance orderings

## 2.1 Inequality

### 2.1.1 Orderings on $\mathbb{R}^n$

It is convenient to define inequality orderings on income vectors which have been placed in order from poorest to richest. To that end let $\mathbb{D}^n = \{x \in \mathbb{R}^n \mid x_{i+1} \geq x_i, \text{ for } i = 1, \ldots, n-1\}$ and $\mathbb{D}^n_+ = \{x \in \mathbb{R}^n \mid x_{i+1} \geq x_i, \ , x_i > 0 \text{ for } i = 1, \ldots, n-1\} = \mathbb{D}^n \cap \mathbb{R}^n_+$ denote spaces of ordered vectors.

The two crucial orderings of interest for this paper are defined by closing up of all gaps in relative or absolute terms.

[4] I am myself grateful to an anonymous referee for bringing the pertinence of this paper and its precedence in proving certain results to my attention.

**Definition 1** *(a) Say that $x$ difference dominates $y$, written $x \succsim_A y$, iff $x_{i+1} - x_i \leq y_{i+1} - y_i$ for $i = 1, \ldots, n-1$ and $x,\ y \in \mathbb{D}^n$.*

*(b) Say that $x$ ratio dominates $y$, written $x \succsim_R y$, iff $\ln(x_{i+1}) - \ln(x_i) \leq \ln(y_{i+1}) - \ln(y_i)$ for $i = 1, \ldots, n-1$ and $x,\ y \in \mathbb{D}^n_+$.*

If $x$ difference dominates $y$ then all absolute gaps are smaller and if $x$ ratio dominates $y$ then all relative gaps are smaller. Both orderings are discussed in Marshall, Walkup and Wets (1967) where they are treated as special cases of cone orderings[5].

These orderings go by different names. Moyes (1994) refers to dominance in absolute and relative differentials. Zheng (2006) refers to absolute and ratio differential conditions. In the absence of unanimity on any alternative terminology, I keep to that used in Preston (1990a).

Ratio and difference dominance can obviously be nested as special cases within a more general class of orderings requiring the closing up of gaps in the value of any increasing function of incomes, say $U$. Zheng (2006) makes this generalisation, defining a more general class of utility gap orderings. If we choose $U(x) = \ln(x + \mu)$ with $\mu \in \mathbb{R}_+$ then we get a class of orderings which will give ratio and difference dominance as extreme cases, in line with the treatment of intermediate orderings in Kolm (1976a, b)[6].

It is also useful to have definitions of transformations of vectors which do not change inequality. In absolute terms all gaps are maintained by a parallel shift in a vector, called a translation, and all relative gaps by multiplying all incomes by a positive constant, referred to here as a rescaling. Let $e_n \in \mathbb{R}^n_+$ denote the vector all elements of which are unity.

[5] $x$ and $y$ are cone ordered if the difference between them lies in a specified convex cone. If $x \succsim_A y$ or $x \succsim_R y$, for example, then the differences between the absolute or relative gaps in the two vectors lie in the particular cone defined by the nonnegative orthant. Majorisation, discussed below, is also a cone ordering. This is a framework which yields useful insights but we do not adopt it here.

[6] Kolm regards the view that equal absolute increases in income preserve inequality as "leftist" and the view that equal proportional increases preserve inequality as "rightist". This categorisation could be questioned, particularly if we consider the implications for views on decreases rather than increases in income. Is it more leftist to think that equal absolute cuts in income, as through a poll tax, preserve inequality? While it seems certainly true that the leftist would prefer a given positive sum to be distributed through equal absolute increases than through equal proportional ones it is not obvious that this reflects a view about how inequality should be measured.

**Definition 2** *(a) Say that $x$ is a translation of $y$ iff $x_i = y_i + \lambda$, $i = 1, \ldots, n$ or more simply $x = y + \lambda e_n$ for some $\lambda \in \mathbb{R}$ and $x,\ y \in \mathbb{D}^n$.*

*(b) Say that $x$ is a rescaling of $y$ iff $x_i = \lambda y_i$, $i = 1, \ldots, n$, or more simply $x = \lambda y$ for some $\lambda \in \mathbb{R}_+$ and $x,\ y \in \mathbb{D}^n_+$.*

We can link the inequality orderings defined above to changes in income vectors which do unambiguously close up gaps.

**Definition 3** *(a) Say that $x$ is an absolute lower end elevation of $y$ iff, for some $k$ and some $\lambda > 0$, $x_i = y_i + \lambda$ for $i = 1, \ldots, k$ and $x_i = y_i$ for $i = k+1, \ldots, n$ with $x,\ y \in \mathbb{D}^n$.*

*(b) Say that $x$ is a relative lower end elevation of $y$ iff, for some $k$ and some $\lambda > 1$, $x_i = \lambda y_i$ for $i = 1, \ldots, k$ and $x_i = y_i$ for $i = k+1, \ldots, n$ with $x,\ y \in \mathbb{D}^n_+$*

Combining absolute or relative lower end elevations with translations or rescalings are the only ways to secure difference or ratio dominance. We state this formally.

**Theorem 1** *(a) $x \succsim_A y$ iff $x$ can be obtained from $y$ by a finite series of absolute lower end elevations and a translation.*

*(b) $x \succsim_R y$ iff $x$ can be obtained from $y$ by a finite series of relative lower end elevations and a rescaling.*

**Proof of Theorem 1.**

(a) Sufficiency follows from the facts that any lower end elevation reduces absolute gaps for $i = 1, \ldots, k$ and leaves them unchanged for $i = k+1, \ldots, n$ whereas any translation leaves absolute gaps unchanged. To see necessity, suppose $x \succsim_A y$. Then $x$ can be obtained from $y$ by a series of $n-1$ absolute lower end elevations, where the $k$th lower end elevation raises $y_i$ by $y_{k+1} - y_k - x_{k+1} + x_k \geq 0$ for $i = 1, \ldots, k$, and a translation by $x_n - y_n$.

(b) The result follows from the above given that $x \succsim_R y$ iff $\ln(x) \succsim_A \ln(y)$. ■

Difference and ratio dominance are stronger inequality concepts than those prevalent in the literature. Since Pigou (1912) and, especially, Dalton (1920) it has been widely accepted that inequality is reduced by a sort of change which can not necessarily be reduced to changes of the sort discussed above.

**Definition 4** *Say that $x$ can be obtained from $y$ by an (elementary) progressive transfer*[7] *if, for some $k$ and some $\lambda > 0$, $x_k = y_k + \lambda$, $x_{k+1} = y_{k+1} - \lambda$ and $x_i = y_i$, $i = 1, \ldots, k-1, k+2, \ldots, n$ and $x,\ y \in \mathbb{D}^n$ (Muirhead 1903, Pigou 1912, Dalton 1920).*

Well known results link progressive transfers to the most widely accepted criterion for inequality comparison, that of Lorenz dominance.

**Definition 5** *Say that $x$ Lorenz dominates $y$, written $x \succsim^L y$, iff $\sum_{i=1}^{k}[x_i - y_i] \geq 0$, $\sum_{i=1}^{n}[x_i - y_i] = 0$ and $x,\ y \in \mathbb{D}^n$.*

The Lorenz ordering is equivalent to a relation known as majorisation (Marshall and Olkin 1979) which is widely used outside of economic contexts. A famous result establishes that $x \succsim^L y$ iff $x$ can be obtained from $y$ by a finite series of progressive transfers (Hardy, Littlewood and Pólya 1934, Atkinson 1970). The Lorenz ordering can be extended to comparisons which do not involve equal means by allowing progressive transfers to be combined with translations and rescalings.

**Definition 6** *(a) Say that $x$ absolute Lorenz dominates $y$, written $x \succsim^L_A y$, iff $x \succsim^L y + \lambda e_n$ for some $\lambda$ (Shorrocks 1983, Moyes 1987).*

*(b) Say that $x$ relative Lorenz dominates $y$, written $x \succsim^L_R y$, iff $x \succsim^L \lambda y$ for some $\lambda > 0$ (Lorenz 1905).*

[7]Progressive transfers are sometimes defined as any transfers from richer to poorer individuals, not necessarily next to each other in the ordering of incomes. The term "elementary progressive transfer" is from Arnold (1987).

### 2.1.2 Orderings on distributions

We can also define analogous orderings more generally on spaces of distribution functions or, equivalently, quantile functions. Such a setting clearly subsumes that of the earlier section, allowing for comparison of income vectors of different dimensions but also of inequality in continuous distributions.

Let $\mathcal{D}$ denote the space of nondecreasing functions from $[0,1]$ to $\mathbb{R}$ and $\mathcal{D}_+$ denote the space of nondecreasing functions from $[0,1]$ to $\mathbb{R}_+$. Let $\mathcal{D}^{\mathcal{C}}$ denote the space of differentiable, nondecreasing functions from $[0,1]$ to $\mathbb{R}$ and $\mathcal{D}^{\mathcal{C}}_+$ denote the space of differentiable, nondecreasing functions from $[0,1]$ to $\mathbb{R}_+$.

If $x$ is distributed according to distribution function $F_x : \mathbb{R} \to [0,1]$, let $\xi_x : [0,1] \to \mathbb{R}$ be the corresponding quantile function, or inverse distribution function, defined by $\xi_x(p) = \sup\{x \mid F_x(x) \leq p\}$.

We can now define the corresponding orderings defined on distributions.

**Definition 7** *(a) Say that $F_x \succsim_A F_y$ iff $\xi_x(p) - \xi_y(p)$ is nonincreasing for $p \in [0,1]$ and $\xi_x,\ \xi_y \in \mathcal{D}$*

*(b) Say that $F_x \succsim_R F_y$ iff $\ln(\xi_x(p)) - \ln(\xi_y(p))$ is nonincreasing for $p \in [0,1]$ and $\xi_x,\ \xi_y \in \mathcal{D}_+$.*

Marshall, Olkin and Proschan (1967) and Barlow and Proschan (1975) say that $F_x$ is star-shaped with respect to $F_y$ if $\xi_y(F_x(x))/x$ is increasing. Arnold defines an ordering identical to $\succsim_R$ which he calls star-shaped ordering, star ordering or simply $*$-ordering.

Suppose that the distributions under comparison and their inverses are differentiable with associated densities $f_x$ and $f_y$. Then $\xi_x'(p) = 1/f_x(F_x^{-1}(p))$ so that $F_x \succsim_A F_y$ iff $f_x(\xi_x(p)) \geq f_y(\xi_y(p))$ for $p \in [0,1]$.

We can also define absolute and relative Lorenz curves using the quantile functions (see Gastwirth 1971)

$$
\begin{aligned}
L_x^A(q) &= \int_0^q \xi_x(p)\mathrm{d}p - q\int_0^1 \xi_x(p)\mathrm{d}p \\
L_x^R(q) &= \int_0^q \xi_x(p)\mathrm{d}p \Big/ \int_0^1 \xi_x(p)\mathrm{d}p
\end{aligned}
$$

and thus define Lorenz orderings

**Definition 8** *(a) Say that* $F_x \succsim_A^L F_y$ *iff* $L_x^A(q) \geq L_y^A(q)$ *for all* $q \in [0,1]$ *and* $\xi_x,\ \xi_y \in \mathcal{D}$

*(b) Say that* $F_x \succsim_R^L F_y$ *iff* $L_x^R(q) \geq L_y^R(q)$ *for all* $q \in [0,1]$ *and* $\xi_x,\ \xi_y \in \mathcal{D}_+$.

### 2.1.3 Relations between orderings

The fact that difference and ratio dominance imply but are not implied by absolute and relative Lorenz dominance is long established.

**Theorem 2** *(Marshall, Olkin and Proschan 1967; Marshall, Walkup and Wets 1967; Jakobsson 1979; Thon 1987; Arnold 1987)*

*(a) If* $x \succsim_A y$ *then* $x \succsim_A^L y$.

*(b) If* $x \succsim_R y$ *then* $x \succsim_R^L y$.

*(c) If* $F_x \succsim_A F_y$ *then* $F_x \succsim_A^L F_y$.

*(d) If* $F_x \succsim_R F_y$ *then* $F_x \succsim_R^L F_y$.

The point is that an absolute or relative lower end elevation can always be implemented by a translation or rescaling followed by a finite series of progressive transfers. A lower end elevation obviously raises mean income. Consider a translation or rescaling which leads to the same increase in mean income and following this by redistributing the income increase to those at the lower end by transferring the income gains at the top end down the distribution to those at the bottom.

It is not possible on the other hand to implement a progressive transfer by a series of lower end elevations and translations or rescalings. A progressive transfer anywhere other than at the extremes of the distribution raises some income gaps at the same time as it reduces others and we have shown that lower end elevations cannot raise income gaps.

We can illustrate the relation between these orderings by showing areas of dominance in comparisons within the standard simplex as in Sen (1973) and many later papers. Imagine looking down at the origin from a point along the ray of equality in the positive orthant of income space with $n = 3$. Any allocation of a given total income, which we normalise to unity, can be represented as a point in the simplex illustrated in Figure 1. If we take an arbitrary initial income vector then the six possible permutations give the vertices of the irregular hexagonal shape shown in the Figure. As is well known, the convex hull of these six points represents the set of income vectors which, permuted into appropriate order, Lorenz dominate the initial income vector.

Now consider which set of points, appropriately permuted, ratio dominate the initial vector. Ratios between incomes for any two individuals are constant throughout planes containing the third axis, and these planes intersect the simplex along rays passing through its vertices. Points of greater equality between these two are those lying between such rays and the bisecting ray through the same vertex. Hence the area in which all ratios are nearer to unity is the star-shaped[8] area outlined in bold within the Lorenz hexagon. Since this fits inside the hexagon, coinciding only at its vertices, it is diagrammatically plain that ratio dominance implies without being implied by Lorenz dominance in comparisons of vectors with equal means.

Differences between any two incomes are maintained in planes which cut the simplex along lines perpendicular to its sides. Hence, constructing, as in the ratio based case, an area in which all differences are diminished gives the inverted Y-shape outlined with dots inside the hexagon. Again this fits entirely inside the Lorenz hexagon except at its vertices. It contains however some points inside and some outside the ratio-based star shape (the areas shaded on the diagram), demonstrating that neither ratio nor difference dominance implies the other.

## 2.2 Welfare

Dalton's conviction that progressive transfers reduce welfare was motivated by the recognition that if social welfare was the sum of individual utilities which depend only on own income and do so in a concave fashion then such transfers raise social welfare. In comparisons between vectors with equal total income Lorenz dominance can be identified with improvement in utilitarian social welfare (Hardy, Littlewood and Pólya 1934, Atkinson 1970). We can extend this observation to comparisons involving vectors with different means by defining a generalisation of Lorenz dominance.

**Definition 9** *(a) Say that $x$ generalised Lorenz dominates $y$, written $x \succsim^{GL}$*

[8] The coincidence between the name of the star-shaped ordering and the shape of the figure is purely fortuitous. The origin of the name lies in the connection with star-shaped functions, as explained below.

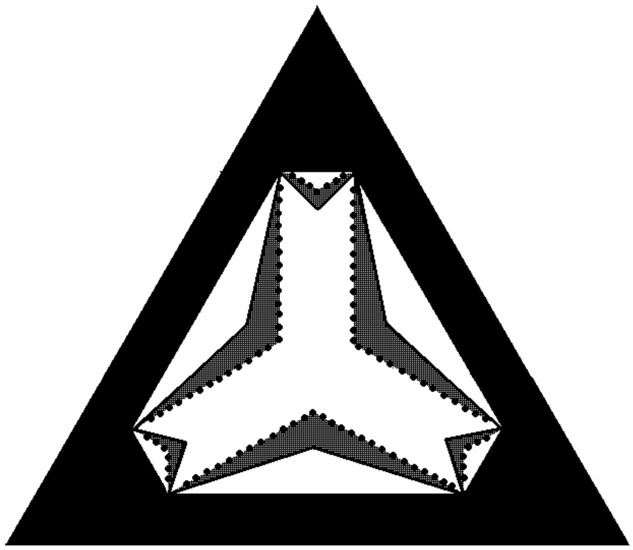

Figure 1: Dominance orderings illustrated in the standard simplex

*$y$, iff $\sum_{i=1}^{k}[x_i - y_i] \geq 0$ and $x,\ y \in \mathbb{D}^n$ (Kolm 1969; Shorrocks 1983).*

*(b) Say that $F_x \succsim^{GL} F_y$, iff $\int_0^q [\xi_x(p) - \xi_y(p)]dp \geq 0$ for all $q \in [0,1]$ and $\xi_x,\ \xi_y \in \mathcal{D}$*

This ordering is called supermajorisation in the noneconomic context (Marshall and Olkin 1979). Kolm does not use the term generalised Lorenz dominance but refers to a social preference for distributions which generalised Lorenz dominate others as isophily. Generalised Lorenz dominance is also strongly linked to utilitarian social welfare. Suppose social welfare $W : \mathbb{D}^n \rightarrow \mathbb{R}$ is the sum of individual utilities which are continuous, increasing, concave functions of individual incomes, $u_i : \mathbb{R} \rightarrow \mathbb{R}$. Clearly $x \succsim^{GL} y$ implies $W(x) \geq W(y)$. In fact $W(x) \geq W(y)$ for all social welfare functions with these properties iff $x \succsim^{GL} y$ (Kolm 1969, Marshall and Olkin 1979, Shorrocks 1983).

What, though, if utilities are not formed in a purely self regarding way? Suppose individual utilities $u_i(x_i, x - ex_i)$ depend not only on own income but also on the gaps between own income and the incomes of others[9]. For example, suppose individual utilities have the form $u_i(x_i, x) = \upsilon_i(x_i)+\psi(x_{i+1}-x_i)$ where $\upsilon_i$ has the usual continuous, increasing and concave properties but $\psi$ is strongly enough decreasing. It is easy to construct an example where a progressive transfer in the middle of the income distribution does not increase social welfare because of the harm done to the utilities of individuals with incomes below the recipient of the transfer. In such a context an absolute lower end elevation, however, would always still increases social welfare because no income would fall and no gap increase.

Lower end elevations are the natural basis for welfare comparisons corresponding to the difference and ratio dominance orderings.

Let us define two orderings which correspond to income changes which should unambiguously increase social welfare even if income gaps matter to individual well being.

**Definition 10** *(a) Say that* $x \succsim_A^* y$ *iff* $x_{i+1} - x_i \leq y_{i+1} - y_i,\ i = 1, \ldots, n-1$ *and* $x_n \geq y_n$ *with* $x,\ y \in \mathbb{D}^n$*.*

*(b) Say that* $x \succsim_R^* y$ *iff* $\ln(x_{i+1}) - \ln(x_i) \leq \ln(y_{i+1}) - \ln(y_i),\ i = 1, \ldots, n-1$ *and* $x_n \geq y_n$ *for* $x,\ y \in \mathbb{D}_+^n$*.*

These say that no element in the vector is reduced and either all absolute or all relative gaps are reduced. The link to lower end elevations is obvious.

**Theorem 3** *(a)* $x \succsim_A^* y$ *iff* $x$ *can be obtained from* $y$ *by a finite series of absolute lower end elevations.*

*(b)* $x \succsim_R^* y$ *iff* $x$ *can be obtained from* $y$ *by a finite series of relative lower end elevations.*

These criteria for welfare comparison are plainly stronger than generalised Lorenz dominance since all partial sums of income are obviously increased by changes which raise all incomes.

[9]This sort of dependence is sometimes referred to as envy but that is a somewhat tendentious term, envy, one of the seven deadly sins, being condemned in most ethical codes. The term envy is suggestive of a wish to bring down the incomes of those better off. It need not be supposed that such sentiments are necessarily implied by demoralisation arising from accentuated feelings of social inferiority - indeed the perception that those on higher incomes deserve the high social position that one can not attain may be precisely the source of deterioration in psychological well being.

**Theorem 4** *If $x \succsim_A^* y$ or $x \succsim_R^* y$ then $x \succsim^{GL} y$*

We can also define similar orderings over distribution functions for which similar links will hold.

**Definition 11** *(a) Say that $F_x \succsim_A^* F_y$ iff $\xi_x(p) - \xi_y(p)$ is nondecreasing for $p \in [0, 1]$, $\xi_x(1) \geq \xi_y(1)$ and $\xi_x,\ \xi_y \in \mathcal{D}$*

*(b) Say that $F_x \succsim_A^* F_y$ iff $\ln(\xi_x(p)) - \ln(\xi_y(p))$ is nondecreasing for $p \in [0, 1]$, $\xi_x(1) \geq \xi_y(1)$ and $\xi_x,\ \xi_y \in \mathcal{D}_+$*

# 3 Inequality and welfare indices

We have seen that progressive transfers increase the sum of concave functions of the individual elements in the vector. The class of functions defined on the vector which are such that they rise with progressive transfers is a wider class than this, first studied by Schur (1923) and known as Schur convex functions. Schur convex functions are those which are said to preserve the Lorenz order. A decreasing function of a Schur convex function is said to be Schur concave. If Lorenz dominance is felt to be a convincing criterion for judging inequality then this is the natural class of functions to use for measuring inequality and most measures proposed for this purpose do fall into this class (although there are notable exceptions such as the variance of logarithms which is infamously not Schur concave, as discussed in Foster and Ok 1999). Similarly if generalised Lorenz dominance is felt to be a suitable criterion for judging social welfare improvement then measures of social welfare ought to rise with progressive transfers and therefore to be Schur convex, besides having other properties such as being increasing in all incomes.

In this section of the paper we discuss the classes of functions of income vectors and distributions which preserve the orderings defined above based on income gaps. We start by defining formally what it means for a function to preserve an ordering.

**Definition 12**

*Say that a function $\phi : X \to \mathbb{R}$ preserves an ordering $\succsim$ iff $\phi(x) \geq \phi(y)$ whenever $x \succsim y$, $x,\ y \in X$.*

We consider functions which are differentiable. For this purpose we need a notion of derivative for a mapping defined on a space of functions. Specifically, suppose $\phi : \mathcal{D}^{\mathcal{C}} \to \mathbb{R}$ is a functional defined on the space of quantile functions. Then we assume the existence of a functional derivative[10] $\delta\phi : [0,1] \to \mathbb{R}$ with the property that $\int \delta\phi \ \psi dp = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\phi(\xi + \epsilon\eta) \mid_{\epsilon=0}$ for $\xi \in \mathcal{D}^{\mathcal{C}}$ and differentiable functions $\eta : [0,1] \to \mathbb{R}$.

## 3.1 Inequality

Functions which preserve the ratio and difference dominance orderings are characterised in the following result.

**Theorem 5** *(a) A differentiable function $\phi : \mathbb{D}^n \to \mathbb{R}$ preserves $\succsim_A$ iff $\sum_{i=1}^k \frac{\partial}{\partial x_i}\phi(x) \geq 0$ for $k = 1, \ldots, n-1$ and $\sum_{i=1}^n \frac{\partial}{\partial x_i}\phi(x) = 0$ (Marshall, Walkup and Wets 1967).*

*(b) A differentiable function $\phi : \mathbb{D}^n_+ \to \mathbb{R}$ preserves $\succsim_R$ iff $\sum_{i=1}^k x_i \frac{\partial}{\partial x_i}\phi(x) \geq 0$ for $k = 1, \ldots, n-1$ and $\sum_{i=1}^n x_i \frac{\partial}{\partial x_i}\phi(x) = 0$ (Marshall, Walkup and Wets 1967).*

*(c) A differentiable function[11] $\phi : \mathcal{D}^{\mathcal{C}} \to \mathbb{R}$ preserves $\succsim_A$ iff $\int_0^q \delta\phi \ \mathrm{d}p \geq 0$ for $q < 1$ and $\int_0^1 \delta\phi \ \mathrm{d}p = 0$*

*(d) A differentiable function $\phi : \mathcal{D}^{\mathcal{C}}_+ \to \mathbb{R}$ preserves $\succsim_R$ iff $\int_0^q \xi \ \delta\phi \ \mathrm{d}p \geq 0$ for $q < 1$ and $\int_0^1 \xi \ \delta\phi \ \mathrm{d}p = 0$*

**Proof of Theorem 5.**

[10]Suppose that $\phi$ extends continuously to a function on the space of all continuous functions on the unit interval. Since this is a Banach space under suitable norm $\|\cdot\|$ we could then, for example, take $\delta\phi$ to be the Fréchet derivative defined by

$$\lim_{\eta \to 0} \frac{\|\phi(\xi + \eta) - \phi(\xi) - \delta\phi(\xi)\|}{\|\eta\|} = 0.$$

[11]The assumption that the quantile function is differentiable or even continuous, implicit in the stated domain for $\phi$, is probably stronger than needed. However, given the use made of integration by parts in the proof, it would be necessary to introduce a generalisation of the notion of derivative for $\xi$, and a careful treatment of the issues involved is beyond the scope of this paper.

(a) If $x \succsim_A y$ then we can get from $y$ to $x$ by a finite series of lower end elevations and a translation. Given $\sum_{i=1}^{k} \frac{\partial}{\partial x_i}\phi(x) \geq 0$ for $k = 1, \ldots, n-1$ the lower end elevations all increase $\phi$ and given $\sum_{i=1}^{n} \frac{\partial}{\partial x_i}\phi(x) = 0$ the translation leaves it unaffected.

It is necessary for $\sum_{i=1}^{n} \frac{\partial}{\partial x_i}\phi(x) = 0$ since, for any $\lambda$, $x \succsim_A x + \lambda e_n$ and $x + \lambda e_n \succsim_A x$. It is necessary for $\sum_{i=1}^{k} \frac{\partial}{\partial x_i}\phi(x) \geq 0$ as $\phi$ must increase with any lower end elevation.

(b) Obvious, noting that ratio dominance is just difference dominance in logarithms.

(c) If $F_x \succsim_A F_y$ then $\xi_x = \xi_y - \eta$ where $\eta \in \mathcal{D}^{\mathcal{C}}$. For any $\eta \in \mathcal{D}^{\mathcal{C}}$, we have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\epsilon}\phi(\xi - \epsilon\eta) \mid_{\epsilon=0} &= -\int_0^1 \delta\phi \; \eta \; \mathrm{d}p \\ &= -\eta(1)\left[\int_0^1 \delta\phi \; \mathrm{d}p\right] + \int_0^q \eta' \; \left[\int_0^q \delta\phi \; \mathrm{d}p\right] \mathrm{d}q \end{aligned}$$

Since, for any $\lambda \in \mathbb{R}$, $F_x \succsim_A F_y$ and $F_y \succsim_A F_x$ if $\xi_x = \xi_y - \lambda e$, we must have $\int_0^1 \delta\phi \; \mathrm{d}p = 0$. Then

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\phi(\xi - \epsilon\eta) \mid_{\epsilon=0} = \int_0^q \eta' \; \left[\int_0^q \delta\phi \; \mathrm{d}p\right] \mathrm{d}q$$

and for this to be positive for all increasing $\eta$ requires $\int_0^q \delta\phi \; \mathrm{d}p \geq 0$ for all $q$.

(d) Obvious, noting that ratio dominance is just difference dominance in logarithms.

■

The results for orderings on $\mathbb{R}^n$ were derived by Marshall, Walkup and Wets (1967), although the proof and the interpretation here are quite different.

The classes of indices derived here are obviously broader than that of Schur convex functions, which are characterised by

$$\frac{\partial}{\partial x_i}\phi(x) \geq \frac{\partial}{\partial x_{i+1}}\phi(x).$$

Functions appropriate as indices of inequality are, of course, those which preserve the reverse orderings, $\precsim_A$ and $\precsim_R$, and indices which preserve $\precsim_R$ but not $\precsim_R^L$ include, for example, the variance of logarithms.

## 3.2 Welfare

Just as progressive transfers raise welfare as well as reducing inequality, we consider welfare functions which are increased by lower end elevations.

**Theorem 6** *(a) A differentiable function $\phi : \mathbb{D}^n \to \mathbb{R}$ preserves $\succsim_A^*$ iff $\sum_{i=1}^k \frac{\partial}{\partial x_i}\phi(x) \geq 0$ for $k = 1, \ldots, n-1$.*

*(b) A differentiable function $\phi : \mathbb{D}_+^n \to \mathbb{R}$ preserves $\succsim_R^*$ iff $\sum_{i=1}^k x_i \frac{\partial}{\partial x_i}\phi(x) \geq 0$ for $k = 1, \ldots, n-1$ (Marshall, Walkup and Wets 1967).*

*(c) A differentiable function $\phi : \mathcal{D}^{\mathcal{C}} \to \mathbb{R}$ preserves $\succsim_A^*$ iff $\int_0^q \delta\phi \, \mathrm{d}p \geq 0$ for $q < 1$*

*(d) A differentiable function $\phi : \mathcal{D}_+^{\mathcal{C}} \to \mathbb{R}$ preserves $\succsim_R^*$ iff $\int_0^q \xi \, \delta\phi \, \mathrm{d}p \geq 0$ for $q < 1$*

**Proof of Theorem 6.**

(a) If $x \succsim_A^* y$ then we can get from $y$ to $x$ by a finite series of lower end elevations. Arbitrary lower end elevations increase $\phi$ iff $\sum_{i=1}^k \frac{\partial}{\partial x_i}\phi(x) \geq 0$ for $k = 1, \ldots, n-1$.

(b) Obvious, noting that ratio dominance is just difference dominance in logarithms.

(c) If $F_x \succsim_A^* F_y$ then $\xi_x = \xi_y + e\eta(1) - \eta$ where $\eta \in \mathcal{D}^{\mathcal{C}}$. For any $\eta \in \mathcal{D}^{\mathcal{C}}$, we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\epsilon}\phi(\xi - \epsilon\,[e\eta(1) - \eta])\,|_{\epsilon=0} &= \eta(1)\left[\int_0^1 \delta\phi \, \mathrm{d}p\right] - \int_0^1 \delta\phi \, \eta \, \mathrm{d}p \\ &= \int_0^q \eta' \left[\int_0^q \delta\phi \, \mathrm{d}p\right] \mathrm{d}q\end{aligned}$$

For this to be positive for all increasing $\eta$ requires $\int_0^q \delta\phi \, \mathrm{d}p \geq 0$ for all $q$.

(d) Obvious, noting that ratio dominance is just difference dominance in logarithms.

■

Again these results are partly anticipated by Marshall, Walkup and Wets (1967) though the treatment is quite different.

The conditions for functions to preserve the welfare orderings differ only from those to preserve the equality orderings in that the requirements for invariance to translation and rescaling are dropped. The requirement for nonnegativity of partial sums of derivatives remains. Functions which preserve $\succsim_A^*$ and $\succsim_R^*$ need not be increasing in all incomes but must increase with positive translations and scalings up of incomes, respectively.

There is a tradition of linking inequality indices with measurement of social welfare going back at least to Dalton (1920) and exemplified by Kolm (1969), Atkinson (1970) and Blackorby and Donaldson (1978, 1980, 1984).

To expound the key results here we need to define properties of functions implying certain sorts of response to translations and rescalings. Let $e \in \mathcal{D}_+^{\mathcal{C}}$ denote the quantile function which is constant at unity.

**Definition 13** *(a) A differentiable function $\phi : \mathbb{D}^n \to \mathbb{R}$ is translatable iff there exists an increasing function $g : \mathbb{R} \to \mathbb{R}$ and a function $\psi : \mathbb{D}^n \to \mathbb{R}$ such that $\phi = g(\psi(x))$ and $\psi(x + \lambda e_n) = \psi(x) + \lambda$.*

*(b) A differentiable function $\phi : \mathbb{D}_+^n \to \mathbb{R}$ is homothetic iff there exists an increasing function $g : \mathbb{R} \to \mathbb{R}$ and a function $\psi : \mathbb{D}_+^n \to \mathbb{R}$ such that $\phi = g(\psi(x))$ and $\psi(\lambda x) = \lambda\psi(x)$.*

*(c) A differentiable function $\phi : \mathcal{D}^{\mathcal{C}} \to \mathbb{R}$ is translatable iff there exists an increasing function $g : \mathbb{R} \to \mathbb{R}$ and a function $\psi : \mathbb{D}^n \to \mathbb{R}$ such that $\phi = g(\psi(\xi))$ and $\psi(\xi + \lambda e) = \psi(\xi) + \lambda$.*

*(d) A differentiable function $\phi : \mathcal{D}_+^{\mathcal{C}} \to \mathbb{R}$ is homothetic iff there exists an increasing function $g : \mathbb{R} \to \mathbb{R}$ and a function $\psi : \mathbb{D}_+^n \to \mathbb{R}$ such that $\phi = g(\psi(\xi))$ and $\psi(\lambda\xi) = \lambda\psi(\xi)$.*

The notion of the equally distributed equivalent income function due to Kolm (1969) and Atkinson (1970) is a crucial one in this literature. For comparisons of vectors define this as $\chi : \mathbb{D}^n \to \mathbb{R}$ by $\phi(\chi(x)e_n) = \phi(x)$ and for comparisons of distribution functions define it as $\chi : \mathcal{D}^{\mathcal{C}} \to \mathbb{R}$ by $\phi(\chi(\xi_x)e) = \phi(\xi_x)$. Clearly $\chi$ preserves the same orderings as does $\phi$. If $\phi$ is Schur convex and translatable then subtracting $\chi$ from mean income gives an index which is Schur concave and invariant to translation. If $\phi$ is Schur convex and homothetic then the proportional difference between mean income and $\chi$ gives an index which is Schur concave and invariant to rescaling. In the first case we have an absolute inequality index linked to measurement of welfare and in the latter a relative inequality index - both ideas are fund in Kolm (1969) and developed by later authors (Atkinson 1970, Blackorby and Donaldson 1978, 1980, 1984).

Given that progressive transfers leave mean incomes unchanged it is natural to construct inequality indices intended to preserve the Lorenz ordering by comparison to mean income but this will not work if the index is intended to preserve the orderings based on gaps. Lower end elevations do not leave mean income unchanged. However we can construct suitable indices by comparison to the maximum income.

**Theorem 7** *(a) If a differentiable function $\phi : \mathbb{D}^n \to \mathbb{R}$ is translatable and preserves $\succsim_A^*$ then $\theta(x) = x_n - \chi(x)$ preserves $\precsim_A$.*

*(b) If a differentiable function $\phi : \mathbb{D}_+^n \to \mathbb{R}$ is homothetic and preserves $\succsim_R^*$ then $\theta(x) = 1 - \chi(x)/x_n$ preserves $\precsim_R$.*

*(c) If a differentiable function $\phi : \mathcal{D}^{\mathcal{C}} \to \mathbb{R}$ is translatable and preserves $\succsim_A^*$ then $\theta(\xi) = \xi(1) - \chi(\xi)$ preserves $\precsim_A$.*

*(d) If a differentiable function $\phi : \mathcal{D}_+^{\mathcal{C}} \to \mathbb{R}$ is homothetic and preserves $\succsim_R^*$ then $\theta(\xi) = 1 - \chi(\xi)/\xi(1)$ preserves $\precsim_R$.*

**Proof of Theorem 7.**

(a) If $\phi$ is translatable then $\chi(x+\lambda e_n) = \chi(x)+\lambda$ so that $\theta(x+\lambda e_n) = \theta(x)$ and $\sum_{i=1}^n \frac{\partial}{\partial x_i}\theta(x) = 0$. For $k < n$, $\sum_{i=1}^k \frac{\partial}{\partial x_i}\theta(x) = -\sum_{i=1}^k \frac{\partial}{\partial x_i}\phi(x) \leq 0$.

(b) If $\phi$ is homothetic then $\chi(\lambda x) = \lambda\chi(x)$ so that $\theta(\lambda x) = \theta(x)$ and $\sum_{i=1}^n \frac{\partial}{\partial x_i} x_i\theta(x) = 0$. For $k < n$, $\sum_{i=1}^k \frac{\partial}{\partial x_i} x_i\theta(x) = -\sum_{i=1}^k \frac{\partial}{\partial x_i} x_i\phi(x) \leq 0$.

(c) If $\phi$ is translatable then $\chi(\xi + \lambda e) = \chi(\xi) + \lambda$ so that $\theta(\xi + \lambda e) = \theta(\xi)$ and $\int_0^1 \delta\theta \, \mathrm{d}p = 0$. For $q < 1$, $\int_0^q \delta\theta \, \mathrm{d}p = -\int_0^q \delta\phi \, \mathrm{d}p \leq 0$.

(d) If $\phi$ is homothetic then $\chi(\lambda\xi) = \lambda\chi(\xi)$ so that $\theta(\lambda\xi) = \theta(\xi)$ and $\int_0^1 \xi \, \delta\theta \, \mathrm{d}p = 0$. For $q < 1$, $\int_0^q \xi \, \delta\theta \, \mathrm{d}p = -\int_0^q \xi \, \delta\phi \, \mathrm{d}p \leq 0$.

■

Thus we have a way of constructing inequality indices preserving ratio and difference dominance orderings from social welfare measures with appropriate properties.

To take the simplest example, if social welfare is measured by the homothetic and translatable function giving mean income $\phi(x) = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, which is invariant to progressive transfers and therefore yields no interesting Schur concave inequality measure, then the construction just outlined gives measures $x_n - \bar{x} = \frac{1}{n}\sum_{i=1}^{n}[x_n - x_i]$ and $1 - \bar{x}/x_n = 1 - \frac{1}{n}\sum_{i=1}^{n}[x_i/x_n]$, the mean absolute and relative shortfall from the top income, which do preserve $\precsim_A$ and $\precsim_R$ respectively.

A social welfare function which is translatable, homothetic and strictly Schur convex is $\phi(x) = \frac{2}{n(n+1)}\sum_{i=1}^{n}(n-i)x_i$. The Kolm (1969) procedure gives a relative inequality index $1 - \chi(x)/\bar{x} = G = \frac{2}{n(n+1)\bar{x}}\sum_{i=1}^{n} i[x_i - \bar{x}]$ equal to the well-known Gini coefficient. The procedure outlined above gives the alternative $1 - \chi(x)/x_n = \frac{2}{n(n+1)x_n}\sum_{i=1}^{n}(i-n)[x_i - x_n]$ which approaches $G\bar{x}/x_n - 3[1 - \bar{x}/x_n]$ for large $n$. There exist, though, functions which are Schur convex yet preserve the gap based inequality orderings. Take the social welfare function $\phi(x) = \frac{2}{n(n+1)}\sum_{i=1}^{n} ix_i$. This is Schur concave yet $1 - \chi(x)/x_n = \frac{2}{n(n+1)x_n}\sum_{i=1}^{n} i[x_n - x_i]$ is still an inequality measure consistent with $\precsim_R$.

# 4 Inequality-reducing policies

Often the purpose to which the inequality orderings are to be put is to compare outcomes under two policy regimes in which incomes are determined by different mappings from some underlying source of variation. For such cases, general results can be derived relating the properties of these mappings to the resulting gaps. Let us assume the underlying variation is in a single dimension and denote the variable in question by $z \in \mathbb{R}$, distributed according to $F_z$, drawn from a class of distribution functions $\mathcal{F}$. The outcomes of interest in the two regimes are denoted $x$ and $y$, where $x = \phi_x(z, F_z)$ and $y = \phi_y(z, F_z)$ with $\phi_x : \mathbb{R} \times \mathcal{F} \to \mathbb{R}$ and $\phi_y : \mathbb{R} \times \mathcal{F} \to \mathbb{R}$ continuous and increasing in their first arguments. Then $\xi_x(p) = \phi_x(\xi_z(p))$ and $\xi_y(p) = \phi_y(\xi_z(p))$.

**Theorem 8** *(a) If $x = \phi_x(z, F_z)$, $y = \phi_y(z, F_z)$ with $\phi_x : \mathbb{R} \times \mathcal{F} \to \mathbb{R}$ and $\phi_y : \mathbb{R} \times \mathcal{F} \to \mathbb{R}$ continuous and increasing in their first arguments then the following are equivalent* *(i)* $F_x \succsim_A F_y$ *for all* $F_z \in \mathcal{F}$

*(ii)* $\phi_x(z, F_z) = g(\phi_y(z, F_z), F_z)$ *for some function* $g : \mathbb{R} \times \mathcal{F} \to \mathbb{R}$ *such that, for all* $F_z \in \mathcal{F}$, $g(\phi_y(\xi_z(p), F_z), F_z) - \phi_y(\xi_z(p), F_z)$ *is nonincreasing in* $p$ *for all* $p \in [0, 1]$

*(b) If* $x = \phi_x(z, F_z)$, $y = \phi_y(z, F_z)$ *with* $\phi_x : \mathbb{R} \times \mathcal{F} \to \mathbb{R}_+$ *and* $\phi_y : \mathbb{R} \times \mathcal{F} \to \mathbb{R}_+$ *continuous and increasing in their first arguments the following are equivalent*

*(i)* $F_x \succsim_R F_y$ *for all* $F_z \in \mathcal{F}$

*(ii)* $\phi_x(z, F_z) = g(\phi_y(z, F_z), F_z)$ *for some function* $g : \mathbb{R}_+ \times \mathcal{F} \to \mathbb{R}_+$ *such that, for all* $F_z \in \mathcal{F}$, $g(\phi_y(\xi_z(p), F_z), F_z)/\phi_y(\xi_z(p), F_z)$ *is nonincreasing in* $p$ *for all* $p \in [0, 1]$

**Proof of Theorem 8.**

(a) Sufficiency of the condition in *(ii)* for $F_x \succsim_A F_y$ for all $F_z \in \mathcal{F}$ is obvious.

If, for some $F_z \in \mathcal{F}$, $\phi_x(z, F_z)$ is not a function of $\phi_y(z, F_z)$ then there exist $p_0, p_1 \in [0, 1]$ with $p_1 > p_0$, such that $\phi_y(\xi_x(p_0), F_z) < \phi_y(\xi_z(p_1), F_z)$ but $\phi_x(\xi_z(p_0), F_z) \neq \phi_x(\xi_z(p_1), F_z)$. Then it cannot be that, for $z$ distributed as $F_z$, $\xi_x(p) - \xi_y(p)$ is nonincreasing in $p$ for $p \in [p_0, p_1]$.

If $\phi_x(z, F_z)$ is a function of $\phi_y(z, F_z)$ but there is an $F_z$ and $p_1 > p_0$ such that $g(\phi_y(\xi_z(p_1), F_z), F_z) - \phi_y(\xi_z(p_1), F_z) > g(\phi_y(\xi_z(p_0), F_z), F_z) - \phi_y(\xi_z(p_0), F_z)$ then again it cannot be that, for $z$ distributed as $F_z$, $\xi_x(p) - \xi_y(p)$ is nonincreasing in $p$ for $p \in [p_0, p_1]$.

(b) Obvious given ratio dominance is difference dominance in logarithms. ■

Functions $g : \mathbb{R} \to \mathbb{R}$ such that $g(x)/x$ is increasing are called star-shaped (Bruckner and Ostrow 1962). If the functions $\phi_x$ and $\phi_y$ are differentiable in $z$ then the conditions in the theorem reduce to a comparison of derivatives or elasticities - specifically, $g(x, F_z) - x$ is decreasing in $x$ iff $\partial g/\partial x$ is less than unity and $g(x, F_z)/x$ is decreasing in $x$ iff $\partial \ln g/\partial \ln x$ is less than unity.

If the functions $\phi_x$ and $\phi_y$ are additively or multiplicatively separated then the role of $F_z$ becomes irrelevant. Finding distributions $F_z$ such that dominance fails in either direction is less restricted and the theorem can be extended to cover not only the inequality orderings based on gaps but also Lorenz dominance. In particular, suppose $\mathcal{F}$ includes all distributions with support within $Z$, $x = \phi_x(z)\psi_x(F_z)$, $y = \phi_y(z)\psi_y(F_z)$ and $\phi_x(z) = g(\phi_y(z))$ with $g(z)/z$ increasing over an interval $\bar{Z} \subseteq Z$ then we can find a counterexample in which there is ratio dominance in the reverse direction by choosing $F_z$ with support wholly within $\bar{Z}$. Since ratio dominance implies relative Lorenz dominance then this means $g(z)/z$ falling everywhere is a necessary condition not only for difference dominance but also for Lorenz dominance to hold for all $F_z \in \mathcal{F}$.

**Theorem 9** *(a) If $x = \phi_x(z)+\psi_x(F_z)$, $y = \phi_y(z)+\psi_y(F_z)$ with $\phi_x : \mathbb{R} \to \mathbb{R}$ and $\phi_y : \mathbb{R} \to \mathbb{R}$ continuous and increasing and $\psi_x : \mathcal{F} \to \mathbb{R}$ and $\psi_y : \mathcal{F} \to \mathbb{R}$ then the following are equivalent*

- *(i) $F_x \succsim_A F_y$ for all $F_z$ with support in $Z$*
- *(ii) $F_x \succsim_A^L F_y$ for all $F_z$ with support in $Z$*
- *(iii) $\phi_x(z) = g(\phi_y(z))$ for some function $g : \mathbb{R} \to \mathbb{R}$ such that $g(x) - x$ is nonincreasing in $x$ for all $x \in g(Z)$*

*(b) If $x = \phi_x(z)\psi_x(F_z)$, $y = \phi_y(z)\psi_y(F_z)$ with $\phi_x : \mathbb{R} \to \mathbb{R}_+$ and $\phi_y : \mathbb{R} \to \mathbb{R}_+$ continuous and increasing and $\psi_x : \mathcal{F} \to \mathbb{R}_+$ and $\psi_y : \mathcal{F} \to \mathbb{R}_+$ then the following are equivalent*

- *(i) $F_x \succsim_R F_y$ for all $F_z$ with support in $Z$*
- *(ii) $F_x \succsim_R^L F_y$ for all $F_z$ with support in $Z$*
- *(iii) $\phi_x(z) = g(\phi_y(z))$ for some function $g : \mathbb{R} \to \mathbb{R}$ such that $g(x)/x$ is nonincreasing in $x$ for all $x \in g(Z)$*

For $\psi_x = \psi_y = 1$, the second part of this theorem is the result of Jakobsson (1976). Sufficiency of the condition on $g$ for $F_x \succsim_R^L F_y$ to hold for all $F_z$ was recognised by Fellman (1976) and Kakwani (1977). Eichhorn, Funke and Richter (1984) and Arnold (1987) extend this result to drop the assumption that $\phi_x$ and $\phi_y$ are continuous and increasing.

## 4.1 Applications

### 4.1.1 Progressive taxation and inequality

**Fixed pretax incomes** Suppose $z$ denotes pretax incomes, assumed distributed in a way unaffected by taxation. Let the tax function be $\tau(z)$ so that posttax incomes are $z-\tau(z)$. By Theorem 9 the posttax distribution absolute Lorenz dominates and difference dominates the pretax distribution whatever $F_z$ iff $\tau(z)$ is increasing in $z$ which is to say that marginal tax rates are everywhere positive (Moyes 1988). This is a property that Fei (1981) calls minimal progression.

The posttax distribution relative Lorenz dominates and ratio dominates the pretax distribution whatever $F_z$ iff $\tau(z)/z$ is increasing in $z$ which is to say that marginal tax rates are everywhere above average tax rates. This is the property usually characterised as progression. The posttax distribution dominates the pretax distribution iff the tax is progressive in the sense of taking a greater share of the incomes of the rich all the way along the distribution. This result, proved for example in Jakobsson (1976), captures an old idea. Seligman (1894) dates the first recorded occurrence of progressive taxation to Solonic Athens and the first written recognition that it "will lessen the disparity of fortunes" to Guiccardini's sixteenth century discussion of Florentine taxation, reprinted in Guicciardini (1932)[12].

One tax system will reduce inequality further than another, in the sense of relative Lorenz dominance and ratio dominance, iff the elasticity of posttax income to pretax income - known as residual income progression (Musgrave and Thin 1948) - is everywhere lower. This is among the results due to Jakobsson (1976)[13]. Suppose we have a linear tax on pretax income, $\tau(z) = tz - G$ so that posttax income is $z(1-t)+G$. Then the elasticity is $\frac{z}{z+G/(1-t)}$ so an increase in $t$ reduces inequality.

[12]This is not to say, of course, that one can date this far back recognition of formal criteria for inequality comparison and their relation to properties of the tax system but Guiccardini does say, for example, in discussing the advantages of a particular progressive tax structure that "so doing, not only shall such benefits follow as I have said, but also the ranks of each shall be equally preserved, since we are all citizens of the same rank, and thus all shall become truly equal as we reasonably should be" (*"E cosí faccendo, non sole ne seguiranno tante utilitá e tanti beni che ho io detto, ma ancora si conserverá equalmente el grado di ognuno, perché tutti siamo cittadini e di uno medesimo grado, e cosí diventereno tutti veramente pari, come ragionevolmente dobbiamo essere." (ibid. p.206).*

[13]Keen, Papapanagos and Shorrocks (2000) extend Jakobsson's results to cover the case where taxes on certain ranges of income are zero.

**Fixed pretax wages** Changes in taxation change work incentives and therefore the distribution of earnings given a fixed distribution of wages. The inequality-reducing effect of progressive taxation is less obvious in such a context. However the same theorems still provide the tools for assessing necessary and sufficient conditions for reduction in the inequality of incomes[14].

Let $z$ now denote wages, distributed again according to $F_z \in \mathcal{F}$. Let individual hours of work be $h$ and taxes be $\tau(zh)$ so that posttax income is $zh - \tau(zh)$ assuming no other source of income. Chosen hours under the given tax system, $h = \eta(z)$, may decline with wage but we assume at least that posttax income does not, $z\eta'(z)(1-\tau'(z\eta(z)) > 0$ - the so-called Mirrlees condition.

Then, by Theorem 9, the posttax distribution under one tax system Lorenz dominates and ratio dominates that under another iff the elasticity of posttax income to the wage is lower at each wage. This elasticity is the product of residual income progression at $z\eta(z)$ and one plus the elasticity of hours $\eta(z)$ to $z$ (wherever $\eta(z) > 0$). Two new considerations emerge when considering the impact of a progressive tax change. Firstly, labour supply responses may make the distribution of earnings less equal. Secondly, even if residual income progression falls at each level of earnings, labour supply responses could move individuals into less progressive parts of the tax system so that residual income progression need not fall at each wage rate.

As an example, consider the case of CES preferences with a linear tax, $\tau(zh) = tzh - G$. Posttax incomes are

$$\frac{\beta z^\sigma (1-t)^\sigma}{z(1-t) + \beta z^\sigma (1-t)^\sigma}(z(1-t) + G)$$

The elasticity is

$$(\sigma - 1)\frac{z(1-t)}{z(1-t) + \beta z^\sigma (1-t)^\sigma} + \frac{z}{z + G/(1-t)}$$

An increase in $t$ reduces the second term but increases the first and there is no guarantee that the expression as a whole falls. Preston (1990b) shows that this possible for parameter values which are not unreasonable. In principle, there exist distributions of pretax wages such that inequality is not reduced.

[14]We should be wary of linking changes in income inequality to welfare in this context since individual wellbeing depends on both income and hours.

### 4.1.2 Immigration and inequality

To take a slightly different example, consider the effect of immigration on income inequality in the preexisting resident population of a country. Suppose what is now fixed is the distribution of productive abilities, $z \in [0, 1]$, distributed in the resident population of size $N$ according to $F_z$.

The economy employs $n(z)$ workers of type $z$ to produce a single type of output according to a CES technology whereby output is $[\int_0^1 zn(z)^\sigma \mathrm{d}z]^{1/\sigma}$ with $\sigma < 1$. This output is traded internationally at a fixed world price which we normalise to 1. Demand for labour of type $z$ is determined by equating its marginal value product to the wage $w(z)$

$$zn(z)^{\sigma-1}\left[\int_0^1 \zeta n(\zeta)^\sigma \mathrm{d}\zeta\right]^{(1/\sigma)-1} = w(z).$$

Prior to immigration, this defines an equilibrium wage distribution equating demand for labour of type $z$ to its supply, $n(z) = Nf_z(z)$. For this to involve wages increasing in $z$ requires that we restrict attention to distributions $F_z$ such that $1 + (\sigma - 1)\partial \ln f_z(z)/\partial \ln z > 0$ everywhere[15].

Now assume that there is immigration of $M = mN$ workers with ability distributed according to distribution function $I_z$ with density $i_z$. The economy reaches a new equilibrium at which the ratio of new to old wages at labour type $z$ is $(1 + mi_z(z)/f_z(z))^{\sigma-1}$. Theorem 8 can be applied. The distribution of wages across preexisting resident workers is made more equal, in the sense of ratio dominance, whatever $F_z$ iff immigration policy is such as to guarantee $i_z(z)/f_z(z)$ is increasing at all $z$ so that immigrants are more concentrated in higher earning groups than the in the population already resident.

# 5 Conclusion

This paper has drawn attention to and argued a case for the interest of inequality orderings based explicitly on closing up of income gaps. Drawing

[15]If this condition were to fail then it would be appropriate to alter the definition of equilibrium rather than to assume wage might be decreasing in $z$. Assuming workers of higher ability can do the jobs of less able workers then a sensible definition of equilibrium would equate demand for labour of type $z$ and above to supply of such labour and the equilibrium wage distribution over and around ranges of ability where the condition fails would have flat sections. We assume away this complication here.

where necessary on earlier papers, the links between these and other orderings have been outlined, the classes of functions preserving the orderings has been characterised and applications have been presented showing their useful in comparison of economic policies.

# References

[1] Arnold, B. C., 1987, Majorization and the Lorenz Order: A Brief Introduction, Berlin: Springer-Verlag.

[2] Atkinson, A. B., 1970, On the measurement of inequality, Journal of Economic Theory, 2, 244-263.

[3] Barlow, R. E. and F. Proschan, 1975, Statistical Theory of Reliability and Life Testing: Probability Models, New York: Holt, Rinehart and Winston.

[4] Blackorby, C. and D. Donaldson, 1978, Measures of relative inequality and their meaning in terms of social welfare, Journal of Economic Theory, 18, 59-80.

[5] Blackorby, C. and D. Donaldson, 1980, A theoretical treatment of indices of absolute inequality, International Economic Review, 21, 107-136.

[6] Blackorby and Donaldson, 1984, Ethically significant ordinal indexes of relative inequality, in: R.L. Basmann and G.F. Rhodes (eds), Advances in Econometrics, Vol. 3, Economic Inequality: Measurement and Policy, Greenwich: JAI Press, 131-147.

[7] Blum, W. J. and H. Kalven, 1953, The Uneasy Case for Progressive Taxation, Chicago: University of Chicago Press.

[8] Bruckner, A. M. and E. Ostrow 1962, Some function classes related to the class of convex functions, Pacific Journal of Mathematics, 12, 1203-1215.

[9] Dalton, H., 1920, The measurement of the inequality of incomes, Economic Journal, 30, 348-361.

[10] Dasgupta, P., A. K. Sen and D. Starrett, 1973, Notes on the measurement of inequality, Journal of Economic Theory, 6, 180-187.

[11] Eichhorn, W., H. Funke and W. F. Richter, 1984, Tax progression and inequality of income distribution, Journal of Mathematical Economics, 13, 127-131.

[12] Fei, J. C. H., 1981, Equity oriented fiscal programs, Econometrica, 49, 869-881.

[13] Fellman, J., 1976, The effect of transformations on Lorenz curves, Econometrica, 44, 823-824.

[14] Foster, J.E. and E.A. Ok, 1999, Lorenz dominance and the variance of logarithms. Econometrica, 67, 901-907.

[15] Gastwirth, J. L., 1971, A general definition of the Lorenz curve, Econometrica, 37, 1037-1039.

[16] Guicciardini F., 1932, La decima scalata, in: F. Guiccardini, Opere, Vol VII, Dialogi e discorsi del regimmento di Firenze, ed. by R. Palmarocchi, Bari: G. Laterza.

[17] Hardy, G., J. Littlewood and G. Pólya, 1934, Inequalities, Cambridge: Cambridge Universiy Press.

[18] Jakobsson, U., 1976, On the measurement of the degree of progression, Journal of Public Economics, 5, 161-168.

[19] Kakwani, N. C., 1977, Applications of Lorenz curves in economic analysis, Econometrica, 45, 719-727.

[20] Keen, M., H. Papapanagos and A. Shorrocks, 2000, Tax reform and progressivity, Economic Journal, 110, 50-68.

[21] Kolm, S.-C., 1969, The optimal production of social justice, in: H. Guitton and J. Margolis, (eds) Public Economics, London: Macmillan, 145-200.

[22] Kolm, S.-C., 1976a, Unequal inequalities I, Journal of Economic Theory, 12, 416-442.

[23] Kolm, S.-C., 1976b, Unequal inequalities II, Journal of Economic Theory, 13, 82-111.

[24] Lorenz, M. O., 1905, Methods of measuring the concentration of wealth, Publications of the American Economic Association, 9, 209-219.

[25] Marshall, A. W., D. W. Walkup and R. J.-B. Wets, 1967, Order-preserving functions: applications to majorization and order statistics, Pacific Journal of Mathematics, 23, 569-584.

[26] Marshall, A. W. and I. Olkin, 1979, Inequalities: Theory of Majorization and its Applications, New York: Academic Press.

[27] Marshall, A. W., I. Olkin and F. Proschan, 1967, Monotonicity of ratios of means and other applications of majorization, in: O. Shisha (ed) Inequalities: Proceedings of a Symposium, New York: Academic Press, 177-190.

[28] Moyes, P., 1987, A new concept of Lorenz domination, Economics Letters, 23, 203-207.

[29] Moyes, P., 1988, A note on minimally progressive taxation and absolute income inequality, Social Choice and Welfare, 5, 227-234.

[30] Moyes, P. , 1994, Inequality reducing and inequality preserving transformations of income: symmetric and individualistic transformations, Journal of Economic Theory, 63, 271-298.

[31] Muirhead, R., 1903, Some methods applicable to identities and inequalities of symmetric algebraic functions of $n$ letters, Proceedings of the Edinburgh Mathematical Society, 21, 144-157.

[32] Musgrave, R. A. and T. Thin, 1948, Income tax progression 1929-1948, Journal of Political Economy, 56, 498-514.

[33] Pigou, A. C., 1912, Wealth and Welfare, London: Macmillan.

[34] Preston, I., 1989, The Redistributive Effect of Progressive Taxation, DPhil thesis, University of Oxford.

[35] Preston, I., 1990a, Ratios, differences and inequality indices, IFS Working Paper W90/9.

[36] Preston, I., 1990b, Income redistribution and labour supply specification, IFS Working Paper W90/14.

[37] Rothschild, M. and J. Stiglitz, 1973, Some further results on the measurement of inequality, Journal of Economic Theory, 6, 188-204.

[38] Schur, I., 1923, Über eine Klasse von Mittelbildungen mit Andwungen auf die Determinanttheorie, Sitzungsber Berlin Mathematische Gesellschaft, 22, 9-20.

[39] Seligman, E. R. A., 1894, Progressive Taxation in Theory and Practice, Publications of the American Economic Association, 9, 7-222.

[40] Sen, A. K., 1973, On Economic Inequality, Oxford: Clarendon Press.

[41] Sen, A. K., 1976, Welfare inequalities and Rawlsian axiomatics, Theory and Decision, 7, 243-262.

[42] Sen, A. K., 1978, Ethical measurement of inequality: some difficulties, in: W. Krelle and A. F. Shorrocks (eds), Personal Income Distribution, Amsterdam: North-Holland, 81-94.

[43] Shorrocks, A. F., 1983, Ranking income distributions, Econometrica, 50, 3-17.

[44] Thon, D., 1987, Redistributive properties of progressive taxation, Mathematical Social Sciences, 14, 185-91.

[45] Zheng, B., 2006, Utility gap dominances and inequality orderings, Social Choice and Welfare, forthcoming.