Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality

by

Russell Davidson

Department of Economics
Queen's University
Kingston, Ontario, Canada
K7L 3N6

GREQAM
Centre de la Vieille Charité
2 Rue de la Charité
13002 Marseille, France

and

Jean-Yves Duclos

Department of Economics and CRÉFA
Université Laval
Ste-Foy, Québec, Canada
G1K 7P4

The paper is part of the research programme of the TMR network "Living Standards, Inequality and Taxation" [Contract No. ERBFMRXCT 980248] of the European Communities whose financial support is gratefully acknowleged. This research was also supported by grants from the Social Sciences and Humanities Research Council of Canada, from the Fonds FCAR of the Province of Québec, and from the MIMAP programme of the International Development Research Centre. We are grateful to Timothy Smeeding, Koen Vleminckx, and Stéphan Wagner for assistance in accessing and processing the data, to Paul Makdissi for helpful comments, and to Nicolas Beaulieu and Nicole Charest for research and secretarial assistance. We are grateful to numerous seminar participants for comments on earlier versions of this paper, and to two anonymous referees and a co-editor for many valuable suggestions that have much improved the paper. Duclos also acknowledges the hospitality of the School of Economics at the University of New South Wales, where he was visiting fellow during the revision of the paper.

November, 1999

Abstract

We derive the asymptotic sampling distribution of various estimators frequently used to order distributions in terms of poverty, welfare and inequality. This includes estimators of most of the poverty indices currently in use, as well as estimators of the curves used to infer stochastic dominance of any order. These curves can be used to determine whether poverty, inequality or social welfare is greater in one distribution than in another for general classes of indices and for ranges of possible poverty lines. We also derive the sampling distribution of the maximal poverty lines up to which we may confidently assert that poverty is greater in one distribution than in another. The sampling distribution of convenient dual estimators for the measurement of poverty is also established. The statistical results are established for deterministic or stochastic poverty lines as well as for paired or independent samples of incomes. Our results are briefly illustrated using data for 4 countries drawn from the Luxembourg Income Study data bases.

Keywords Stochastic dominance, Poverty, Inequality, Distribution-free statistical inference, Order-Restricted Inference.

1. Introduction

Since the influential work of Atkinson (1970), considerable effort has been devoted to making comparisons of welfare distributions more ethically robust, by making judgements only when all members of a wide class of inequality indices or social welfare functions lead to the same conclusion, rather than concentrating on some particular index. More recently, pleas have been made for similar robustness in poverty measurement, following up on the criticism in Sen (1976) of the headcount ratio and the poverty gap as not taking into account the intensity and the depth of poverty respectively. Such pleas are found, for instance, in Atkinson (1987), Foster and Shorrocks (1988a,b), and Howes (1993). Robustness is also needed to guard against the uncertainty and the frequent lack of agreement regarding the choice of a precise poverty line.

In this paper, we study estimation and inference in the context of inequality, welfare, and poverty orderings. Our main objective is to show how to estimate orderings which are robust over classes of indices and ranges of poverty lines, and how to perform statistical inference on them.[1]

In the next section, we review the definitions of the various indices in which we are interested for the distributions of entire populations, and we note some of the relations among them[2]. In Section 3, we study estimators of these indices, based on samples drawn from the populations, and we derive their asymptotic distributions. In particular, we discuss the statistical consequences of using *estimated* poverty lines. We also provide estimates of the thresholds up to which one population stochastically dominates another at a given order, and of cumulative poverty gap (CPG) curves. Our results apply equally to the case of observations drawn from independent distributions and to the case in which dependent observations are drawn from a joint distribution, as for instance when, with panel data, there are several observations of the same individual. We obtain as a corollary the distributions of the two most popular classes of poverty indices, both for deterministic and for sample-dependent poverty lines. The first is the class of additive poverty indices, which include the Foster *et al* (1984) indices, which themselves include the headcount and average poverty gap measures, the Clark *et al* (1981), Chakravarty (1983), and Watts (1968) indices. The second

[1] Work on these lines can be found in, for instance, Beach and Davidson, (1983), Beach and Richmond (1985), Bishop *et al* (1989), Howes (1993), Anderson (1996), Davidson and Duclos (1997).

[2] Foster (1984), Chakravarty (1990), Foster and Sen (1997) and Zheng (1997a), among others, can also be consulted for a review of different aspects of the social welfare, poverty, and inequality literatures.

is the class of linear poverty indices, which can be expressed as weighted areas underneath CPG curves. Members of that linear class include the poverty indices of Sen (1976), Takayama (1979), Thon (1979), Kakwani (1980), Hagenaars (1987), Shorrocks (1995), and Chakravarty (1997).

Statistical inference for inequality or poverty indices could be performed without recourse to the asymptotic theory of this paper by use of the bootstrap, with resampling of the observed data serving to provide estimates of the needed variances and covariances. However, it is well known that the bootstrap yields better results when applied to asymptotic pivots, and it is therefore a better idea to use our results in order to construct such asymptotic pivots before using the bootstrap – see Horowitz (1997) for an account of the relevant issues.

Finally, in Section 4, we provide a brief illustration of our techniques using cross-country data from the Luxembourg Income Study data bases. Most of the proofs are relegated to the appendix.

2. Stochastic Dominance and Poverty Indices

Consider two distributions of incomes, characterised by the cumulative distribution functions (CDFs) F_A and F_B, with support contained in the non-negative real line. We use the term "income" throughout the paper to signify a measure of individual welfare, which need not be money income. Let $D_A^1(x) = F_A(x)$ and

$$D_A^s(x) = \int_0^x D_A^{(s-1)}(y)\,dy, \tag{1}$$

for any integer $s \geq 2$, and let $D_B^s(x)$ be defined analogously. It is easy to check inductively that we can express $D^s(x)$ for any order s as

$$D^s(x) = \frac{1}{(s-1)!}\int_0^x (x-y)^{s-1}\,dF(y). \tag{2}$$

Distribution B is said to dominate distribution A stochastically at order s if $D_A^s(x) \geq D_B^s(x)$ for all $x \in \mathbb{R}$. For strict[3] dominance, the inequality must hold strictly over some interval of positive measure. Suppose that a poverty line is established at an income level $z > 0$. Then we will say that B (stochastically) dominates A at order s up to the poverty line if $D_A^s(x) \geq D_B^s(x)$ for all $x \leq z$.

[3] Since the main focus of this paper is statistical, we will not distinguish strict and non-strict dominance, since near the margin no statistical test can do so.

First-order stochastic dominance of A by B up to a poverty line z implies that $F_A(x) \geq F_B(x)$ for all $x \leq z$. This is equivalent to the statement that the proportion of individuals below the poverty line (the headcount ratio) is always (weakly) greater in A than in B, for any poverty line not exceeding z.

Second-order dominance of A by B up to a poverty line z implies that $D_A^2(x) \geq D_B^2(x)$, that is, that

$$\int_0^x (x-y)\,dF_A(y) \geq \int_0^x (x-y)\,dF_B(y) \tag{3}$$

for all $x \leq z$. When the poverty line is z, the *poverty gap* for an individual with income y is defined as

$$g(z,y) = (z-y)_+ = \max(z-y,0) = z - y^* \tag{4}$$

The notation x_+ will be used throughout the paper to signify $\max(x,0)$. In addition, *censored income* y^* is defined for a given poverty line z as $\min(y,z)$. We can see from (3) that stochastic dominance at order 2 up to z implies that, for all poverty lines $x \leq z$, the average poverty gap in A, $D_A^2(x)$, is greater than that in B, $D_B^2(x)$. The approach is easily generalised to any desired order s.

Ravallion (1994) and others have called the graph of $D^1(x)$ a poverty incidence curve, that of $D^2(x)$ a poverty deficit curve (see also Atkinson (1987)), and that of $D_A^3(x)$ a poverty severity curve. $D^1(x)$ is shown in Figure 1 for two distributions A and B. Distribution B dominates A for all common poverty lines below z. The area underneath $D^1(x)$ for x between 0 and z equals the average poverty gap $D^2(z)$, which is clearly greater for A than for B.

Following Atkinson (1987), we consider the class of poverty indices, defined over poverty gaps, that take the form

$$\Pi(z) = \int_0^z \pi\big(g(z,y)\big)\,dF(y). \tag{5}$$

These can be regarded as absolute indices since equal additions to both z and y do not affect them. In Atkinson (1987), Foster and Shorrocks (1988a) and McFadden (1989), it is shown, in the context of risk aversion, that, for all indices (5) for which π is differentiable and increasing with $\pi(0) = 0$, $\Pi_A(x) \geq \Pi_B(x)$ for all $x \leq z$ if and only if B stochastically dominates A up to z at first order. This class of indices, along with the headcount ratio, for which it is easy to see that the result holds as well, will be denoted P^1. Similarly, the class P^2 is defined by convex increasing functions π with $\pi(0) = 0$. The use of indices in P^2 is analogous to using social evaluation functions that obey the Dalton principle of transfers (see the discussion of this in Atkinson (1987)). It is easy to show that all indices in P^2 are greater for A than for B for all $x \leq z$ if and only if B stochastically dominates A up to z at second order. In general, for any desired order s, we can define the class P^s to contain those indices (5) for which $\pi^{(s)}(x) \geq 0$ for $x > 0$, $\pi^{(s-1)}(0) \geq 0$, and $\pi^{(i)}(0) = 0$ for $i = 0, \ldots, s-2$. Then it is easy to show that $\Pi_A(x) \geq \Pi_B(x)$ for all $x \leq z$ for all $\Pi \in P^s$ if and only if B dominates A up to z at order s. The classes P^s of poverty indices can be interpreted using the generalised transfer principles of Kolm (1976), Fishburn and Willig (1984) and Shorrocks (1987).[4] Note that, for class P^2, we need not require that $\pi'(0) = 0$, but, for $s > 2$, all the derivatives of π up to order $s-2$ must vanish at 0.[5]

A useful concept for the analysis of poverty is the maximum common poverty line z_s up to which B stochastically dominates A at order s. All indices in P^s will then indicate greater poverty in A than in B for any poverty line no greater than z_s. If B stochastically dominates A (at first order) for low thresholds z, then either B dominates A everywhere (in which case we have first-order welfare dominance in the sense of Foster and Shorrocks (1988b)), or else there is a reversal at the value z_1 defined by

$$z_1 = \inf \{x > 0 \,|\, F_A(x) < F_B(x)\}. \tag{6}$$

z_1 is illustrated in Figure 1. If z_1 is below the maximum possible income, we can repeat the exercise at order 2. Either B dominates A at second order everywhere,[6] or there exists z_2 defined by

$$z_2 = \inf \{x > 0 \,|\, D_A^2(x) < D_B^2(x)\}. \tag{7}$$

This procedure can be continued either until stochastic dominance at some order s is achieved everywhere, or until z_s has become greater than what is seen as a reasonable maximum possible value for the poverty line (or welfare censoring threshold) z. It is shown in Lemma 1 in the Appendix

4 For $s = 1, 2$, Foster and Shorrocks (1988b) show how some of these dominance relationships can be extended to poverty indices (or censored social welfare functions) that may be non-additive.

5 For $\alpha \geq s$, the FGT indices $\Delta^\alpha(z)$ defined below obey this continuity condition and thus belong to the classes P^s. Besides, if an additive index of the type $A(z) \equiv \int_0^\infty \delta(y, z)\, dF(y)$ defined below belongs to P^s, then, for $\gamma > 1$, all additive indices of the form $(\delta(y, z))^\gamma$ will belong to P^{s+1}.

6 This is equivalent to Generalised Lorenz dominance of the distribution of incomes in B over that in A, and to second-order welfare dominance.

that stochastic dominance of A by B up to any finite z *will* be achieved for some suitably large value of s. This result confirms the interpretation of stochastic dominance for general s given by Fishburn and Willig (1984) in terms of principles that give increasing weights to transfers occurring at the bottom of the distribution. The limit as $s \to \infty$ has a Rawlsian flavour, since, in that limit, only the very bottom of the two distributions determines which dominates the other for large s.

In comparing poverty across time, societies, or economic environments, it can be desirable to use different poverty lines for different income distributions.[7] This is particularly common in studies of poverty in developed economies where a proportion of median or average incomes is often used as a "poverty line" to make cross-country comparisons.[8] We may continue to use the classes P^s as above, but now, in order to compare A and B, we use two different poverty lines z_A and z_B. Then it is easily shown that there is at least as much poverty in A as in B, according to all indices in P^s, if and only if $D_A^s(z_A - x) - D_B^s(z_B - x) \geq 0$ for all $x \geq 0$.[9] This just involves checking whether B dominates A for all pairs of poverty lines of the form $(z_A - x, z_B - x)$ with $x \geq 0$. As x varies, the absolute difference between the two poverty lines remains constant. Of course, this relation no longer constitutes stochastic dominance at order s.

The popular FGT (see Foster *et al* (1984)) class of additive poverty indices is defined by[10]

$$\varDelta^\alpha(z) = \int_0^z (z - y)^{\alpha-1}\, dF(y) = \int_0^\infty g(z, y)^{\alpha-1}\, dF(y). \tag{8}$$

These indices are clearly related to the criteria for stochastic dominance, as was noted by Foster and Shorrocks (1988a,b). In fact, if α is an integer, it follows from (2) that $\varDelta^\alpha(x) = (\alpha - 1)!\, D^\alpha(x)$.

7 See, for instance, Greer and Thorbecke (1986) and Ravallion and Bidani (1994), where poverty lines are estimated for different socio-economic groups, and Sen (1981, p.21) on the issue of comparing poverty of two societies with either common or different "standards of minimum necessities".

8 On this, see, for instance, Smeeding *et al* (1990), Van den Bosch *et al* (1993), Gustafsson and Nivorozhkina (1996), or Atkinson (1995).

9 Well-known arguments of the type found in Foster and Shorrocks (1988b) can be used to show that this extends to non-additive indices for $s = 1, 2$. In the terminology of Jenkins and Lambert (1997), dominance for $s = 2$ implies an ordering for all generalised (additive or non-additive) poverty gap indices.

10 The original FGT indices are normalised by $z^{\alpha-1}$. We return to the interpretation of this normalisation below.

For any one member of the FGT class of indices, there may be a range of common poverty lines for which poverty in A is greater than in B. For any such line z, the index Δ^s shows more poverty in A than in B if $D_A^s(x) - D_B^s(x) \geq 0$ for $x = z$, but not necessarily for all $x < z$. Hence, it could be that, for a given range of z, we find dominance of A by B according to Δ^1 and Δ^3, but also find dominance of B by A according to Δ^2, a reversal which would not be possible with stochastic dominance relations. We could then define the thresholds z_s^- and z_s^+, such that B dominates A according to Δ^s only for $z \in [z_s^-, z_s^+]$. More generally, we may only wish to check whether $D_A^s(x) \geq D_B^s(x)$ for x in our range of interest. For $s = 1$ or $s = 2$, this leads to the concept of restricted stochastic dominance defined in Atkinson (1987) for the headcount ratio and the mean poverty gap respectively (see his Conditions 1 and 2). It is clear that such restricted dominance conditions can be applied and generalised to any order s of the FGT index.

Other poverty indices can also be expressed in the additive form of (2), that is, as

$$A(z) \equiv \int_0^\infty \delta(y, z)\, dF(y) \tag{9}$$

for suitable choices of $\delta(y, z)$. This is the case for the Clark *et al* (1981) second family of indices, for the Chakravarty (1983) index, for which $\delta(y, z) = 1 - (y^*/z)^e$ for $0 < e < 1$, and for the Watts (1968) index, where $\delta(y, z) = \log(z/y^*)$. Bourguignon and Fields (1997) also propose an additive index that allows for discontinuities at the poverty line, with $\delta(y, z) = g(z, y)^{\alpha_1 - 1} + \alpha_2 I(y \leq z)$, where $I(y \leq z)$ is an indicator function equal to 1 when $y \leq z$, and 0 otherwise.

Stochastic dominance at first and second order can also be expressed in terms of quantiles. This is called the p–approach to dominance. The indices in P^1 indicate at least as much poverty in A as in B if and only if, for all $0 \leq p \leq 1$,

$$(z_A - Q_A(p))_+ - (z_B - Q_B(p))_+ \geq 0, \tag{10}$$

where $Q_A(p)$ and $Q_B(p)$ are the p–quantiles of the distributions A and B respectively. If $z_A = z_B$, condition (10) simplifies to checking if the quantiles of B's censored distribution are never smaller than those of A. As can be seen from Figure 1, the condition (10) need only be checked for values of p less than the greatest value, p^* in the figure, for which $z_B - Q_B(p^*) \geq 0$. It is also clear from the figure that, where the curves for A and B cross, at $p = p_1$, the common value of $Q_A(p_1)$ and $Q_B(p_1)$ is the z_1 defined in (6).

There also exists a p–approach to second-order dominance. To see this, define the *cumulative poverty gap* (CPG) curve (also called TIP curve by

Jenkins and Lambert (1997), and poverty gap profile by Shorrocks (1998); see also Spencer and Fisher (1992)) by

$$G(p; z) = \int_0^{Q(p)} g(z, y)\, dF(y). \tag{11}$$

It is clear that $G(p; z)/p$ is the average poverty gap of the $100p\%$ poorest individuals. Typical CPG curves are shown in the upper panel of Figure 2. For values of p greater than $F(z)$, the CPG curve saturates and becomes horizontal. Since $F(z) = D^1(z)$, the abscissa at which the curve becomes horizontal is the headcount ratio. The ordinate for values of p such that $F(z) \le p \le 1$ is readily seen to be $D^2(z)$, the average poverty gap.

To make the link with second-order stochastic dominance, we quote a result of Jenkins and Lambert (1997) and Shorrocks (1998). They show that, for two distributions A and B and a common poverty line z, it is necessary and sufficient for the stochastic dominance of A by B at second order up to z that $G_A(p; z) \ge G_B(p; z)$ for all $p \in [0, 1]$. The more general case with different poverty lines can be easily derived from Theorem 2 in Shorrocks (1983). Using Shorrocks' result, we find that poverty is greater in A than in B according to all indices in the class P^2 if and only if the CPG curve for A (using z_A) everywhere dominates the CPG curve for B (using z_B)[11].

CPG curves can be related to generalised Lorenz curves $GL(p)$, defined by (see Shorrocks (1983)):

$$GL(p) = \int_0^{Q(p)} y\, dF(y).$$

It is clear from this definition and (11) that $G(p; z) = zp - GL(p)$ for $p \le D^1(z)$. Thus, as shown in Figure 2, for $p \le D^1(z)$, $G(p; z)$ is the vertical distance between the straight line zp with slope z and $GL(p)$. When $G(p; z)$ saturates at $p = D^1(z)$, its derivative with respect to p vanishes, and so we see that, at $p = D^1(z)$, $GL'(p) = z$. Thus, for $p \ge D^1(z)$, $G(p; z)$ is the vertical distance between the line zp and the tangent to $GL(p)$ at $p = D^1(z)$. When we compare two distributions, A and B, this link between $GL(p)$ and $G(p; z)$ shows that the critical second-order poverty line z_2 defined in (7) is given by the slope of the line that is simultaneously tangent to both of the generalised Lorenz curves, at points a and b in Figure 2. This follows because, as can be seen in the figure, the vertical distances between the line z_2p and the two generalised Lorenz curves at the points at which their

[11] See also Jenkins and Lambert (1998) for an extension of this to generalised poverty gap indices.

slopes equal z_2 are equal. Since these distances equal $D^2(z_2)$, the result follows.

A popular class of poverty measures that are linear in incomes can be easily obtained from $G(p;z)$. To see this, consider the class of indices $\Theta(z)$ that measure a weighted area beneath the CPG curve

$$\Theta(z) = \frac{1}{\tau(z) \cdot z} \int_0^{q(z)} \theta(p)\, G(p;z)\, dp \tag{12}$$

for various choices of the functions $\tau(\cdot)$, $q(\cdot)$, and $\theta(\cdot)$. $\Theta(z)$ is linear in incomes since $G(p;z)$ is itself a linear (cumulative) function of incomes.[12] Sen's (1976) index is given by setting $\theta(p) = 2$, $\tau(z) = D^1(z)$, and $q(z) = D^1(z)$. $\theta(p) = 2$, $\tau(z) = D^2(z)$ and $q(z) = 1$ yield the Takayama (1979) index. $\theta(p) = 2$, $\tau(z) = 1$ and $q(z) = 1$ give Thon's (1979), Shorrocks' (1995) and Chakravarty's (1997) poverty indices. Kakwani's (1980) index is obtained with $\theta(p) = \big(k(k+1)\big)\big(D^1(z) - p\big)^{k-1} / (D^1(z))^k$, with $k > 0$, $\tau(z) = 1$, and $q(z) = D^1(z)$. More generally, we can define any linear poverty index $\Theta(z)$ by defining $\theta(p)$ as some particular non-negative function of p. As for the FGT indices, we might also wish to infer the restricted ranges $[z^-, z^+]$ over which the additive or linear indices $A(z)$ and $\Theta(z)$ show more poverty in A than in B.

In the literature on the measurement of poverty, the poverty gap (4) is sometimes normalised by the poverty line.[13] For this, absolute poverty gaps $g(z,y)$ are replaced by relative poverty gaps[14] $g^r(z,y) = g(z,y)/z$, in the definitions of the poverty indices found in (5). We define classes P_r^s of relative poverty indices analogously to the classes P^s, with $g^r(z,y)$ in place of $g(z,y)$. The stochastic dominance conditions are obviously unchanged if poverty lines are common. It can be seen that there will be more poverty in A than in B for all indices in P_r^s if and only if

$$\frac{D_A^s(z_A x)}{z_A^{s-1}} - \frac{D_B^s(z_B x)}{z_B^{s-1}} \geq 0 \tag{13}$$

for all $x \in [0,1]$. The theoretically equivalent p–approach for class P_r^1 is given by checking whether

$$\frac{\big(z_A - Q_A(p)\big)_+}{z_A} - \frac{\big(z_B - Q_B(p)\big)_+}{z_B} \geq 0. \tag{14}$$

[12] This is analogous to the definition of linear inequality indices in Mehran (1976).

[13] It is not clear that this is desirable when poverty lines differ across groups or societies; see Atkinson (1991), p.7 and footnote 3.

[14] For a discussion of absolute versus relative poverty gaps and indices, see Blackorby and Donaldson (1978) and (1980).

For second-order dominance, the p–approach can be derived by redefining the CPG curve in terms of relative poverty gaps as follows:

$$G^r(p) = \int_0^p \left(\frac{\big(z - Q(q)\big)_+}{z} \right) \, dq \tag{15}$$

and checking whether $G^r_A(p) - G^r_B(p) \geq 0$ for all $0 \leq p \leq 1$.[15]

Finally, for indices of relative inequality, observe that $D^s(x)$ can be used to check both equality and welfare dominance when means are the same. When A and B have different means, μ_A and μ_B say, we can study equality dominance by comparing the mean-normalised distributions $F_A(x\mu_A)$ and $F_B(x\mu_B)$.[16] This implies checking whether

$$\frac{D^s_A(\mu_A x)}{\mu_A^{s-1}} - \frac{D^s_B(\mu_B x)}{\mu_B^{s-1}} \geq 0 \tag{16}$$

for all $x \geq 0$. For $s = 2$, this is equivalent to checking Lorenz dominance. Similarly to z_s, we can define critical common *proportions* x_s of the respective means up to which condition (16) is met at a given order s. When Lorenz curves cross, x_2 will give the slope of the line that is simultaneously tangent to both of the Lorenz curves.[17]

3. Estimation and Inference

Suppose that we have a random sample of N independent observations y_i, $i = 1, \ldots, N$, from a population. Then it follows from (2) that a natural estimator of $D^s(x)$ (for a nonstochastic x) is

$$\begin{aligned} \hat{D}^s(x) &= \frac{1}{(s-1)!} \int_0^x (x - y)^{s-1} \, d\hat{F}(y) \\ &= \frac{1}{N(s-1)!} \sum_{i=1}^N (x - y_i)^{s-1} I(y_i \leq x) = \frac{1}{N(s-1)!} \sum_{i=1}^N (x - y_i)_+^{s-1} \end{aligned} \tag{17}$$

where $\hat{F}$ denotes the empirical distribution function of the sample and $I(\cdot)$ is an indicator function equal to 1 when its argument is true and 0 otherwise. For $s = 1$, (17) simply estimates the population CDF by the empirical distribution function. For arbitrary s, it has the convenient property of being a sum of IID variables.

[15] See also Jenkins and Lambert (1998).

[16] This is also discussed in Foster and Shorrocks (1988c), Foster and Sen (1997) and Formby *et al* (1998).

[17] The argument for this is analogous to that used above in discussing Figure 2.

When comparing two distributions in terms of stochastic dominance, two kinds of situations typically arise. The first is when we consider two independent populations, with random samples from each. In that case,

$$\operatorname{var}\bigl(\hat{D}^s_A(x) - \hat{D}^s_B(x')\bigr) = \operatorname{var}\bigl(\hat{D}^s_A(x)\bigr) + \operatorname{var}\bigl(\hat{D}^s_B(x')\bigr). \tag{18}$$

The other typical case arises when we have N independent drawings of paired incomes, y_i^A and y_i^B, from the same population. For instance, y_i^A could be before-tax income, and y_i^B after-tax income for the same individual i, $i = 1, \ldots, N$. The following theorem allows us to perform statistical inference in both of these cases.

> **Theorem 1:** Let the joint population moments of order $2s-2$ of y^A and y^B be finite. Then $N^{1/2}(\hat{D}^s_K(x) - D^s_K(x))$ is asymptotically normal with mean zero, for $K = A, B$, and with asymptotic covariance structure given by $(K, L = A, B)$
>
> $$\begin{aligned}&\lim_{N\to\infty} N \operatorname{cov}\bigl(\hat{D}^s_K(x),\, \hat{D}^s_L(x')\bigr)\\ &= \frac{1}{\bigl((s-1)!\bigr)^2} E\left((x - y^K)_+^{s-1}(x' - y^L)_+^{s-1}\right) - D^s_K(x)\, D^s_L(x').\end{aligned} \tag{19}$$

Proof: For each distribution, the existence of the population moment of order $s-1$ lets us apply the law of large numbers to (17), thus showing that $\hat{D}^s(x)$ is a consistent estimator of $D^s(x)$. Given also the existence of the population moment of order $2s-2$, the central limit theorem shows that the estimator is root-N consistent and asymptotically normal with asymptotic covariance matrix given by (19). This formula clearly applies not only for y^A and y^B separately, but also for the covariance of $\hat{D}^s_A$ and $\hat{D}^s_B$.

If A and B are independent populations, the sample sizes N_A and N_B may be different. Then (19) applies to each with N replaced by the appropriate sample size. The covariance across the two populations is of course zero. ∎

Remarks: This theorem was proved for the case of independent samples as early as 1989 in an unpublished thesis, Chow (1989). The sampling distribution of the related estimator $\Delta^s(x)$ (see (8)) with a fixed x and independent samples is also found in Kakwani (1993), Bishop *et al* (1995) and Rongve (1997). For a different approach to inference on stochastic dominance, see Anderson (1996).

The asymptotic covariance (19) can readily be consistently estimated in a distribution-free manner by using sample equivalents. Thus $D^s(x)$ is estimated by $\hat{D}^s(x)$, and the expectation in (19) by

$$\frac{1}{N}\sum_{i=1}^{N}(x - y_i^K)_+^{s-1}(x' - y_i^L)_+^{s-1}. \tag{20}$$

If B does dominate A weakly at order s up to some possibly infinite threshold z, then, for all $x \leq z$, $D^s_A(x) - D^s_B(x) \geq 0$. There are various hypotheses that could serve either as the null or the alternative in a testing procedure. The most restrictive of these, which we denote H_0, is that $D^s_A(x) - D^s_B(x) = 0$ for all $x \leq z$. Next comes H_1, according to which $D^s_A(x) - D^s_B(x) \geq 0$ for $x \leq z$, and, finally, H_2, which imposes no restrictions at all on $D^s_A(x) - D^s_B(x)$. We observe that these hypotheses are nested: $H_0 \subset H_1 \subset H_2$.

McFadden (1989) proposes a test based on $\sup_{x \leq z}\big(\hat{D}^s_A(x) - \hat{D}^s_B(x)\big)$ for the null of H_0 against H_1 For $s = 1$, this turns out to be a variant of the Kolmogorov-Smirnov test, with known properties, for the identity of two distributions. Of higher values of s, McFadden considers only $s = 2$. Although it is easy to compute the statistic, its asymptotic properties under the null are not analytically tractable. However, a simulation-based method can provide critical values and P values.

In Kaur, Prakasa Rao, and Singh (1994) (henceforth KPS), a test is proposed based on the minimum or infimum of the t statistic for the hypothesis that $D^s_A(x) - D^s_B(x) = 0$, computed for each value of $x \leq z$. The minimum value is used as the test statistic for the null of *non*-dominance, $H_2 \backslash H_1$, against the alternative of dominance, H_1. Since the test can be interpreted as an intersection-union test, it is shown that the probability of rejection of the null when it is true is asymptotically bounded by the nominal level of a test based on the standard normal distribution.

Both the McFadden and the KPS statistics are calculated as the extreme value of the possibly very large set of values computed for $x = Y^A_i$ and $x = Y^B_j$ for all $i = 1, \ldots, N_A, j = 1, \ldots, N_B$. Other procedures make use of a predetermined grid of a much smaller number of points, x_j say, for $j = 1, \ldots, m$, at which $d_j \equiv D^s_A(x_j) - D^s_B(x_j)$, or some quantity related to it, like the t statistic considered above, is evaluated. The arbitrariness of the choice of the number of points m, and the precise values of the x_j, is an undesirable aspect of all procedures of this sort. At the very least, it is necessary that the x_j should constitute a grid covering the whole interval of interest.

Howes (1993) proposed an intersection-union test for the null of non-dominance, very much like the KPS test, except that the t statistics are calculated only for the predetermined grid of points. Its properties are similar to those of KPS.

The technique developed by Beach and Richmond (1985) allows us to test H_1 (dominance) against H_2 (no restriction). The alternative is *not* the hypothesis that A dominates B. That hypothesis can of course play the role of H_1 and be tested similarly against H_2. This technique was originally designed by Richmond (1982) to provide simultaneous confidence intervals for a set of variables asymptotically distributed as multivariate normal with known or consistently estimated asymptotic covariance matrix. It was extended by Bishop, Formby, and Thistle (1992), who suggested a union-intersection test of the hypothesis that one set of Lorenz curve decile ordinates dominates another. For a test of stochastic dominance, one can use the t statistics for the hypotheses that the individual d_j, $j = 1, \ldots, m$, are zero. The hypothesis H_1, which implies that they are all nonnegative, is rejected against the unconstrained alternative, H_2, if any of the t statistics is significant with the wrong sign (that is, in the direction of dominance of B by A), where significance is determined asymptotically by the critical values of the Studentised Maximum Modulus (SMM) distribution with m and an infinite number of degrees of freedom.

None of the tests discussed so far makes use of the asymptotic covariance structure provided by Theorem 1. As a result, they can be expected to be conservative, that is, lacking in power, relative to tests that do exploit that structure. Such tests, to date at least, all rely on a predetermined grid, on which stochastic dominance implies the set of m inequalities $d_j \geq 0$. Methods for testing hypotheses relating to such inequalities are developed in Robertson, Wright, and Dykstra (1988), in the context of order-restricted inference. For our purposes, the relevant methods can be found in Kodde and Palm (1986), and Wolak (1989). Wolak provides a variety of asymptotically equivalent tests of H_0 against H_1, and of H_1 against H_2, and provides the joint distribution under H_0 of the statistics corresponding to the two tests. He also shows that this allows us to bound the size of the test asymptotically when H_1 is the null, because, for any nominal level, the rejection probability under H_1 is maximised when H_0 is true.[18] The test of H_0 against H_2 is of less interest, and in any case it is a perfectly standard test of a set of equality restrictions.

A feature of the order restricted approach is that, if m is large, the mixture of chi-squared distributions followed by the test statistics under H_0 can, as Wolak remarks, be difficult to compute. However, he also proposes a Monte Carlo approach that works independently of the magnitude of m, and can be implemented with sufficient accuracy easily enough on present-day computers. The necessary ingredient for any of these procedures is the asymptotic covariance structure of the d_j.

[18] Note, however, that in the more general context of arbitrary nonlinear inequality restrictions on the parameters of a nonlinear model, it is not necessarily true that the rejection probability is maximised at the point at which all the restrictions hold with equality: see Wolak (1991) for full discussion of this point.

One should note that non-rejection of the null of dominance by either the Wolak or Bishop-Formby-Thistle approach can occur along with non-rejection of the null of non-dominance by the KPS or Howes approach. This occurs naturally if the $\hat{D}^s(\cdot)$ functions for the two populations are close enough over part of the relevant range. Such issues, and many others, are investigated in a valuable recent paper of Dardanoni and Forcina (1999), who also consider hypotheses according to which more than two distributions are ranked by a stochastic dominance criterion. They emphasise the intrinsically conservative nature of the KPS and Howes tests, and find that they are wholly lacking in power for comparisons with more than two distributions, although, as will be seen in our empirical illustration in Section 4, they remain useful with just two distributions when it is undesirable to infer dominance unless there is very strong evidence for it.

Dardanoni and Forcina investigate, in a set of Monte Carlo experiments, the power gain achieved by the tests of Wolak and of Kodde and Palm relative to tests that do not take account of the covariance structure of the d_j. They find that these are greatest when the d_j are negatively correlated. Although this can occur naturally in comparisons of more than two populations, the usual case with only two is that they are positively correlated. Even so, they find that methods like Wolak's are often worth the extra computational burden they impose – this conclusion is borne out by the results in our empirical illustration. They also advocate the use of tests that combine the information in a test of H_0 against H_1 with that in a test of H_1 against H_2, bearing in mind that the statistics are not independent. Their very interesting analysis is however beyond the scope of this paper.

In Theorem 1, it was assumed that the argument x of the functions $D^s(x)$ was nonstochastic. In applications, one often wishes to deal with $D^s(z-x)$, where z is the poverty line. In the next Theorem, we deal with the case in which z is estimated on the basis of sample information.

Theorem 2: Let the joint population moments of order $2s-2$ of y^A and y^B be finite. If $s=1$, suppose in addition that F_A and F_B are differentiable and let $D^0(x) = F'(x)$. Assume first that N independent drawings of pairs (y^A, y^B) have been made from the joint distribution of A and B. Also, let the poverty lines z_A and z_B be estimated by $\hat{z}_A$ and $\hat{z}_B$ respectively, where these estimates are expressible asymptotically as sums of IID variables drawn from the same sample, so that, for some function $\xi_A(\cdot)$,

$$\hat{z}_A = N^{-1}\sum_{i=1}^{N}\xi_A(y_i^A) + o(1) \quad \text{as } N \to \infty, \tag{21}$$

and similarly for B. Then $N^{1/2}(\hat{D}^s_K(\hat{z}_K - x) - D^s_K(z_K - x))$, $K = A, B$, is asymptotically normal with mean zero, and with covariance structure given by $(K, L = A, B)$

$$\lim_{N\to\infty} N \operatorname{cov}\big(\hat{D}^s_K(\hat{z}_K - x),\ \hat{D}^s_L(\hat{z}_L - x')\big) =$$
$$\operatorname{cov}\Big(D^{s-1}_K(z_K - x)\xi_K(y^K) + \big((s-1)!\big)^{-1}(z_K - x - y^K)^{s-1}_+,$$
$$D^{s-1}_L(z_L - x')\xi_L(y^L) + \big((s-1)!\big)^{-1}(z_L - x' - y^L)^{s-1}_+\Big). \tag{22}$$

If y^A and y^B are independently distributed, and if N_A and N_B IID drawings are respectively made of these variables, then, for $K = L$, N_K replaces N in (22), while for $K \neq L$, the covariance is zero.

Proof: See appendix. ∎

Remarks: The sampling distribution of the headcount when the poverty line is set to a proportion of a quantile is derived in Preston (1995), using results on the joint sampling distribution of quantiles. More generally, the sampling distribution of additive indices when the poverty line is expressed as a sum of IID variables is independently derived in Zheng (1997b) using the theory of U statistics.

Estimates of the poverty lines may be independent of the sample used to estimate the $D^s(z - x)$, as for example if they are estimated using different data. In that case, the right-hand side of (22) becomes

$$D^{s-1}_K(z_K - x)D^{s-1}_L(z_L - x')\ \operatorname{cov}\big(N^{1/2}(\hat{z}_K - z_K), N^{1/2}(\hat{z}_L - z_L)\big)$$
$$+ \operatorname{cov}\Big(\big((s-1)!\big)^{-1}(z_K - x - y^K)^{s-1}_+,\ \big((s-1)!\big)^{-1}(z_L - x' - y^L)^{s-1}_+\Big). \tag{23}$$

For indices based on relative poverty gaps, one needs the distribution of $\hat{D}^s(\hat{z}x)$ for positive x; see (13) and (16). The result of Theorem 2 can be used by first eliminating the additive x in that result, and then replacing $\hat{z}$ by $\hat{z}x$.

The covariance (22) can, as usual, be consistently estimated in a distribution-free manner, by the expression

$$N^{-1}\sum_{i=1}^{N}\Big(\Big(\hat{D}_K^{s-1}(\hat{z}_K - x)\xi_K(y_i^K) + \big((s-1)!\big)^{-1}(\hat{z}_K - x - y_i^K)_+^{s-1}\Big)$$

$$\times\Big(\hat{D}_L^{s-1}(\hat{z}_L - x')\xi_L(y_i^L) + \big((s-1)!\big)^{-1}(\hat{z}_L - x' - y_i^L)_+^{s-1}\Big)\Big)$$

$$-\Big(N^{-1}\sum_{i=1}^{N}\Big(\hat{D}_K^{s-1}(\hat{z}_K - x)\xi_K(y_i^K) + \big((s-1)!\big)^{-1}(\hat{z}_K - x - y_i^K)_+^{s-1}\Big)$$

$$\times N^{-1}\sum_{i=1}^{N}\Big(\hat{D}_L^{s-1}(\hat{z}_L - x')\xi_L(y_i^L) + \big((s-1)!\big)^{-1}(\hat{z}_L - x' - y_i^L)_+^{s-1}\Big)\Big).$$

The most popular choices of population dependent poverty lines are fractions of the population mean or median, or quantiles of the population distribution. Clearly any function of a sample moment can be expressed asymptotically as an average of IID variables, and the same is true of functions of quantiles, at least for distributions for which the density exists, according to the Bahadur (1966) representation of quantiles. For ease of reference, this result is cited as Lemma 2 in the Appendix. The result implies that $\hat{Q}(p)$ is root-N consistent, and that it can be expressed asymptotically as an average of IID variables. When the poverty line is a proportion k of the median, for instance, we have that:

$$\xi(y_i) = -k\left(\frac{I\big(y_i < Q(0.5)\big) - 0.5}{F'\big(Q(0.5)\big)}\right),$$

where $Q(0.5)$ denotes the median. When z is k times average income, we have

$$\xi(y_i) = ky_i.$$

This IID structure makes it easy to compute asymptotic covariance structures for sets of quantiles of jointly distributed variables.

For the purposes of testing for stochastic dominance, all the remarks following Theorem 1 regarding possible procedures continue to apply here. Only the asymptotic covariance structure is different, on account of the estimated poverty lines.

We turn now to the estimation of the threshold z_1 defined in (6). Assume that $\hat{F}_A(x)$ is greater than $\hat{F}_B(x)$ for some bottom range of x. If $\hat{F}_A(x)$ is smaller than $\hat{F}_B(x)$ for larger values of x, a natural estimator $\hat{z}_1$ for z_1 can be defined implicitly by

$$\hat{F}_A(\hat{z}_1) = \hat{F}_B(\hat{z}_1).$$

If $\hat{F}_A(x) > \hat{F}_B(x)$ for all $x \le z$, for some prespecified poverty line z, then we arbitrarily set $\hat{z}_1 = z$. If $\hat{z}_1$ is less than the poverty line z, we may define $\hat{z}_2$ by

$$\hat{D}_A^2(\hat{z}_2) = \hat{D}_B^2(\hat{z}_2)$$

if this equation has a solution less than z, and by z otherwise. And so on for $\hat{z}_s$ for $s > 2$: either we can solve the equation

$$\hat{D}_A^s(\hat{z}_s) = \hat{D}_B^s(\hat{z}_s), \tag{24}$$

or else we set $\hat{z}_s = z$. Note that the second possibility is a mere mathematical convenience used so that $\hat{z}_s$ is always well defined – we may set z as large as we wish. The following theorem gives the asymptotic distribution of $\hat{z}_s$ under the assumption that $z_s < z$ exists in the population.

Theorem 3: Let the joint population moments of order $2s-2$ of y^A and y^B be finite. If $s = 1$, suppose further that F_A and F_B are differentiable, and let $D^0(x) = F'(x)$. Suppose that there exists $z_s < z$ such that

$$D_A^s(z_s) = D_B^s(z_s),$$

and that $D_A^s(x) > D_B^s(x)$ for all $x < z_s$. Assume that z_s is a simple zero, so that the derivative $D_A^{s-1}(z_s) - D_B^{s-1}(z_s)$ is nonzero. In the case in which we consider N independent drawings of pairs (y^A, y^B) from one population in which y^A and y^B are jointly distributed, $N^{1/2}(\hat{z}_s - z_s)$ is asymptotically normally distributed with mean zero, and asymptotic variance given by:

$$\lim_{N\to\infty} \operatorname{var}\bigl(N^{1/2}(\hat{z}_s - z_s)\bigr) = \Bigl((s-1)!\bigl(D_A^{s-1}(z_s) - D_B^{s-1}(z_s)\bigr)\Bigr)^{-2} \times \Bigl(\operatorname{var}\bigl((z_s - y^A)_+^{s-1}\bigr) + \operatorname{var}\bigl((z_s - y^B)_+^{s-1}\bigr) - 2\operatorname{cov}\bigl((z_s - y^A)_+^{s-1}, (z_s - y^B)_+^{s-1}\bigr)\Bigr).$$

If y^A and y^B are independently distributed, and if N_A and N_B IID drawings are respectively made of these variables, where the ratio $r \equiv N_A/N_B$ remains constant as N_A and N_B tend to infinity, then $N_A^{1/2}(\hat{z}_s - z_s)$ is asymptotically normal with mean zero, and asymptotic variance given by

$$\lim_{N_A\to\infty} \operatorname{var}\bigl(N_A^{1/2}(\hat{z} - z)\bigr) = \frac{\operatorname{var}\bigl((z - y^A)_+^{s-1}\bigr) + r\operatorname{var}\bigl((z - y^B)_+^{s-1}\bigr)}{\Bigl((s-1)!\bigl(D_A^{s-1}(z) - D_B^{s-1}(z)\bigr)\Bigr)^2}.$$

Proof: See appendix. ∎

Remark: In this theorem, we assume that z_s exists in the population, and is a simple zero of $D_A^s(x) - D_B^s(x)$. Since the $\hat{D}_K^s$, $K = A, B$, are consistent estimators of the D_K^s, this implies that, in large enough samples, $\hat{z}_s$ exists and is unique. In general, in finite samples, it can happen that, although z_s exists, the estimated curves $\hat{D}_A^s(x)$ and $\hat{D}_B^s(x)$ do not intersect. In such cases, our definition gives $\hat{z}_s = z$, and no real harm is done. It may also happen that, even if no z_s exists, the estimated curves cross. In that case, the regularity condition of the theorem is not satisfied, and nothing simple can be said of the spurious estimate $\hat{z}_s$, except of course that, in large enough samples, $\hat{z}_s = z$ with high probability. Clearly, no *asymptotic* approach can handle these awkward cases, because asymptotically the true situation in the population is reflected in the sample. The situation is in fact analogous to what happens with parametric models for which the parameters may be identified asymptotically but not by a given finite data set, or *vice versa*.

The results of Theorems 1, 2 and 3 can naturally be extended to the additive poverty indices $A(z)$ of (9) by using $A(x)$ in place of $D^s(x)$, $\delta(y,x)$ for $\big((s-1)!\big)^{-1}(x-y)_+^{s-1}$, and $A'(x)$ for $D^{s-1}(x)$.

In order to perform statistical inference for p–approaches, we now consider the estimation of the ordinates of the cumulative poverty gap curve $G(p;z)$ defined in (11). The natural estimator, for a possibly estimated poverty line $\hat{z}$, is

$$\hat{G}(p;\hat{z}) = N^{-1} \sum_{i=1}^{N} (\hat{z} - y_i)_+ \, I\big(y_i \le \hat{Q}(p)\big)$$

where $\hat{Q}(p)$ is the empirical p–quantile. The asymptotic distribution of this estimator is given in the following theorem.

Theorem 4: Let the joint population second moments of y^A and y^B be finite, and let F_A and F_B be differentiable. Let $\hat{z}_A$ and $\hat{z}_B$ be expressible asymptotically as sums of IID variables, as in Theorem 2. If N independent drawings of pairs (y^A, y^B) are made from the joint distribution of A and B, then $N^{1/2}(\hat{G}_K(p;\hat{z}) - G_K(p;z))$, for $K = A, B$, is asymptotically normal with mean zero, and asymptotic covariance structure given by

$$\lim_{N\to\infty} N \operatorname{cov}\big(\hat{G}_K(p;\hat{z}_K)\,,\,\hat{G}_L(p';\hat{z}_L)\big) =$$

$$\operatorname{cov}\Bigg(\Big(I\big(y^K \le Q_K(p)\big)\,\big((z_K - y^K)_+ - (z_K - Q_K(p))_+\big) + \xi_K(y^K)\,\min\big(p, F_K(z_K)\big)\Big),\ \Big(I\big(y^L \le Q_L(p')\big)\,\big((z_L - y^L)_+ - (z_L - Q_L(p'))_+\big) + \xi_L(y^L)\,\min\big(p', F_L(z_L)\big)\Big)\Bigg). \tag{25}$$

If y^A and y^B are independently distributed, and if N_A and N_B IID drawings are respectively made of these variables, then, for $K = L$, N_K replaces N in (25). For $K \neq L$, the covariance is zero.

Proof: See appendix. ∎

Remarks: If $\hat{z}_A$ and $\hat{z}_B$ are independent of the drawings (y^A, y^B), then the right-hand side of (25) can be modified as in (23). The result of Theorem 4 for the special case of a deterministic poverty line and for independent samples can also be found in Xu and Osberg (1998).

The arguments used in Theorems 1–4 can be used to obtain the asymptotic distribution of all those indices considered in the previous section not already covered by the earlier theorems. First, when z is deterministically set to a level exceeding the highest income in the sample, Theorem 4 yields the sampling distribution of the generalised Lorenz curves, and of the ordinary Lorenz curves when we also take into account the asymptotic distribution of the sample mean $\hat{\mu}$. Second, for the first-order p–approach, based on quantiles (see (10)), the asymptotic covariance structure is easy to derive because the quantiles can be expressed asymptotically as averages of IID variables, by Bahadur's Lemma, as can the estimated poverty lines, by (21). Third, for the indices based on relative poverty gaps, inference on the expressions in (13), (14), (15) and (16) can be performed by using the asymptotic joint distributions of objects like $\hat{D}^s(x)$, $\hat{z} - \hat{Q}(p)$, $\hat{z}$ and $\hat{\mu}$. Fourth, the asymptotic distribution of estimates $\hat{\Theta}(\hat{z})$ of the general class of linear indices (12) can be readily obtained using the arguments of the proof of Theorem 4.[19] Fifth, the asymptotic distribution of estimators of critical poverty lines z (for $z = z^-, z^+$) for the linear indices $\Theta(z)$ can be obtained from Theorem 3 by replacing $(s-1)^{-2}\,\mathrm{var}\big((z_s - y^K)_+^{s-1}\big)$ by $\lim_{N\to\infty} N\,\mathrm{var}\big(\hat{\Theta}_K(z)\big)$ and $D_K^{s-1}(z_s)$ by $\Theta'_K(z)$, $K = A, B$. Sixth, the asymptotic distribution of estimators of critical *relative* poverty lines x_s in (16) can be derived from Theorem 3. Finally, the arguments found in the proof of Theorem 3 can also be used to provide the asymptotic distribution of the abscissae above which quantile, Lorenz, Generalised Lorenz, or CPG curves cross. For Lorenz curves, for instance, this would give the asymptotic distribution of the maximum proportion of the population for which it can be said that the share of total income is greater in B than in A.

[19] The statistical inference results for the special case of the Sen index with a deterministic poverty line can also be found in Bishop *et al* (1997).

4. Illustration

We illustrate our results using data drawn from the Luxembourg Income Study (LIS) data sets[20] of the USA, Canada, the Netherlands, and Norway, for the year 1991. The raw data were essentially treated in the same manner as in Gottschalk and Smeeding (1997). We take household income to be disposable income (*i.e.*, post-tax-and-transfer income) and we apply purchasing power parities drawn from the Penn World Tables[21] to convert national currencies into 1991 US dollars. As in Gottschalk and Smeeding (1997), we divide household income by an adult-equivalence scale defined as $h^{0.5}$, where h is household size, so as to allow comparisons of the welfare of individuals living in households of different sizes. Hence, all incomes are transformed into 1991 adult-equivalent US$. All household observations are also weighted by the LIS sample weights "hweight" times the number of persons in the household. Sample sizes are 16,052 for the US, 21,647 for Canada, 8,073 for Norway, and 4,378 for the Netherlands.

This illustration does not deal with important statistical issues. First, we assume that observations are drawn through simple random sampling. The LIS data, like most survey data, are actually drawn from a complex sampling structure with stratification, clustering and non-deterministic inclusion rates.[22] It would be possible, if messy, to adapt our methods to deal with complex sampling structures, provided of course that the design was known. Second, negative incomes are set to 0. This procedure, however, affects no more than 0.5% of the observations for all countries considered here. Finally, we ignore the measurement errors due to contaminated data; see Cowell and Victoria-Feser (1996) for a discussion of how to minimise the consequences of these.

Table 1 shows the estimates $\hat{D}^1(x)$ and $\hat{D}^2(x)$ for the selected countries and for poverty lines varying between US $2,000 and US $35,000 in adult-equivalent units, along with their asymptotic standard errors. For the purpose of comparisons, since the samples for the different countries are independent, asymptotic variance estimates for the differences $\hat{D}^s_A(x) - \hat{D}^s_B(x)$ are obtained by adding the variance estimates for countries A and B. Comparing the US with the other countries, we find that first-order dominance never holds everywhere in the samples.

[20] See http://lissy.ceps.lu for detailed information on the structure of these data.

[21] See Summers and Heston (1991) for the methodology underlying the computation of these parities, and http://www.nber.org/pwt56.html for access to the 1991 figures.

[22] See Cowell (1989) and Howes and Lanjouw (1998) for the consideration of such issues in applied distributional analysis.

It can be seen from the data for $s = 1$ that, with a conventional significance level of 5%, Canada has a significantly lower headcount ratio for all x less than or equal to \$25,000 (that is, a poverty line of \$50,000 for a family of 4); in other words, Canada has less poverty than the US for all poverty lines equal to or below \$25,000, and for all P^1 poverty indices. The American headcount is significantly lower than that of Norway only for those x no greater than \$15,000. As for the Netherlands, its headcount is initially significantly greater than that of the US (for x equal to \$2,000), it is lower than for the US for x between \$4,000 and \$8,000, and it is greater again subsequently. These results mean that, by use of Howes' intersection-union procedure, the null of non-dominance of the US by Canada can be rejected at the 5% level for all poverty lines up to \$25,000. The corresponding hypotheses for Norway and the Netherlands cannot be rejected. The null of dominance of the US by Canada, on the other hand, cannot be rejected by the Bishop-Formby-Thistle (BFT) union-intersection procedure until a poverty line of \$35,000. By use of the Wolak procedure, a similar but more precise result is obtained. The Wald test statistic of Kodde and Palm was calculated for the set of incomes given in Table 1 up to \$30,000 and up to \$35,000. The weights for the mixture of chi-squared distributions were obtained by running 10,000 simulations, and P values were calculated. The P value is 0.71 for the null of dominance of the US by Canada up to \$30,000, but only 0.0001 up to \$35,000.

For $s = 2$, the major difference from the results for $s = 1$ is that Canada now dominates the US for all values of x in the samples, and significantly so, so that Howes' procedure rejects the null of non-dominance at order 2. Since there is dominance in the sample, both the BFT and the Wolak method fail to reject the null of dominance. As for Norway, the initial range of values of x for which the US is dominated at second order is (as expected) larger for $s = 2$ than for $s = 1$. Compared to the US, the Netherlands have a significantly greater average poverty gap for $x = \$2,000$, a statistically indistinguishable average poverty gap for $x = \$4,000$, a lower one for x between \$6,000 and \$10,000, and a greater average poverty gap for x above \$15,000. For both Norway and the Netherlands compared with the US, non-dominance at second order is not rejected by the Howes procedure, and dominance is rejected by the BFT procedure. In this case, since the conclusions are clear, it is unnecessary to go to the trouble of using the Wolak procedure. The comparison of Canada and Norway, however, is less clear. By use of Wolak's procedure for $x = \$5,000$ and $x = \$10,000$, over which range Norway dominates Canada in the sample, the P value for the null that Canada dominates Norway is 0.078, so that the null cannot be rejected at the 5% level. If the income range is extended to \$15,000, the P value grows to 0.14. If it is extended all the way to \$35,000, the P value is 0.40. The closeness at the bottom of the range does of course rule out rejection of non-dominance by Howes' procedure.

Table 2 shows estimates of the thresholds z_s for dominance relations, and $[z_s^-, z_s^+]$ for restricted dominance relations, between the US and the other three countries for $s = 1, 2, 3, 4$. Not surprisingly, we find that Canada stochastically dominates the US for $s = 1$ up to a censoring threshold of $27,840, with a standard error on that threshold of $1,575. For higher values of s, Canada dominates the US everywhere up to an arbitrarily large threshold. For Norway, dominance up to $35,000 is attained for $s = 4$, but with a threshold income barely significantly greater than $35,000. Regarding the comparison of the Netherlands and the US, we can conclude that there is first-order dominance of the US for all poverty lines below $2,958 (with a standard error of $193), that there is restricted first-order poverty dominance by the Netherlands over the US for poverty lines between $2,958 and $8,470, and restricted first-order dominance by the US over the Netherlands for poverty lines above $8,470 (with a standard error of $203). Similar results hold for higher values of s.

Table 3 illustrates poverty rankings for the US, Canada, and the Netherlands when the poverty line is set to half median income in each country. Even without use of the Wolak procedure, it can be seen, for $s = 1$, that the null hypotheses of dominance of the US by Canada or by the Netherlands cannot be rejected, and that by Howes' procedure the null of non-dominance of the US by the other two countries can be rejected. Hence, there is more poverty in the US than in any of the other two countries for all P^1 indices, at poverty lines equal to half of median income and for all other pairs of lower poverty lines such that the absolute difference between the two poverty lines remains constant. For $s = 1$, the rankings of Canada and the Netherlands switch twice as x approaches 0. For $s = 2$, however, poverty becomes everywhere significantly greater (except perhaps for $x = \$4,000$ and $x = \$3,000$) in Canada than in the Netherlands.

Appendix

Lemma 1: If B dominates A for $s = 1$ up to some $w > 0$, with strict dominance over at least part of that range, then for any finite threshold z, B dominates A at order s up to z for s sufficiently large.

Proof: We have $F_A(x) - F_B(x) \geq 0$ for $0 \leq x \leq w$, with strict inequality over some subinterval of $[0, w]$. Thus

$$\int_0^w \big(F_A(y) - F_B(y)\big)\,dy \equiv a > 0.$$

We wish to show that, for arbitrary finite z, we can find s sufficiently large that $D^s_A(x) - D^s_B(x) > 0$ for $x \le z$, that is,

$$\int_0^x \left(1 - \frac{y}{x}\right)^{s-1} \big(dF_A(y) - dF_B(y)\big) > 0 \tag{26}$$

for $x < z$. For ease in the sequel, we have multiplied $D^s(x)$ by $(s-1)!/x^{s-1}$, which does not affect the inequality we wish to demonstrate.

Now the left-hand side of (26) can be integrated by parts to yield

$$\frac{s-1}{x} \int_0^x \big(F_A(y) - F_B(y)\big) \left(1 - \frac{y}{x}\right)^{s-2} dy.$$

We split this integral in two parts: the integral from 0 to w, and then from w to x. We may bound the absolute value of the second part: Since $|F_A(y) - F_B(y)| \le 1$ for any y and $1 - y/x \ge 0$ for all $y \le x$, we have

$$\begin{aligned} &\left| \frac{s-1}{x} \int_w^x \big(F_A(y) - F_B(y)\big) \left(1 - \frac{y}{x}\right)^{s-2} dy \right| \\ &\le \frac{s-1}{x} \int_w^x \left(1 - \frac{y}{x}\right)^{s-2} dy = \left(1 - \frac{w}{x}\right)^{s-1}. \end{aligned} \tag{27}$$

For the range from 0 up to w, we have, for $s \ge 2$,

$$\begin{aligned} &\frac{s-1}{x} \int_0^w \big(F_A(y) - F_B(y)\big) \left(1 - \frac{y}{x}\right)^{s-2} dy \\ &\ge \frac{s-1}{x} \left(1 - \frac{w}{x}\right)^{s-2} \int_0^w \big(F_A(y) - F_B(y)\big)\, dy \\ &= \frac{a(s-1)}{x} \left(1 - \frac{w}{x}\right)^{s-2} \end{aligned} \tag{28}$$

Putting (27) and (28) together, we find that, for $x \ge w$,

$$\begin{aligned} &\int_0^x \left(1 - \frac{y}{x}\right)^{s-1} \big(dF_A(y) - dF_B(y)\big) \\ &\ge \frac{a(s-1)}{x} \left(1 - \frac{w}{x}\right)^{s-2} - \left(1 - \frac{w}{x}\right)^{s-1} \\ &= \left(1 - \frac{w}{x}\right)^{s-2} \left(\frac{a(s-1)}{x} - 1 + \frac{w}{x}\right). \end{aligned} \tag{29}$$

If we choose s to be greater than $1 + (z-w)/a$, then, for all $w \le x \le z$, $\big(a(s-1) + w\big)/x - 1 > 0$. Thus for such s, the last expression in (29) is positive for all $w \le x \le z$. For $x < w$, the dominance at first order up to w implies dominance at any order $s > 1$ up to w. The result is therefore proved. ∎

Lemma 2: (Bahadur, 1966). Suppose that a population is characterised by a twice differentiable distribution function F. Then, if the p–quantile of F is denoted by $Q(p)$, and the sample p–quantile from a sample of N independent drawings y_i from F by $\hat{Q}(p)$, we have

$$\hat{Q}(p) - Q(p) = -\frac{1}{Nf\big(Q(p)\big)} \sum_{i=1}^{N} \Big(I\big(y_i < Q(p)\big) - \alpha\Big) + O\big(N^{-3/4}(\log N)^{3/4}\big),$$

where $f = F'$ is the density.

Proof of Theorem 2:

For distributions A and B, we have

$$(s-1)!\,\hat{D}^s(\hat{z}-x) = \int_0^{\hat{z}-x} (\hat{z}-x-y)^{s-1} d\hat{F}(y) \quad \text{and}$$
$$(s-1)!\,D^s(z-x) = \int_0^{z-x} (z-x-y)^{s-1} dF(y).$$

Thus

$$\begin{aligned}
(s-1)!\,&\big(\hat{D}^s(\hat{z}-x) - D^s(z-x)\big) = \\
&\int_{z-x}^{\hat{z}-x} \big((\hat{z}-x-y)^{s-1} - (z-x-y)^{s-1}\big)\,d\hat{F}(y) \\
&+ \int_{z-x}^{\hat{z}-x} (z-x-y)^{s-1}\,d(\hat{F}-F)(y) \\
&+ \int_{z-x}^{\hat{z}-x} (z-x-y)^{s-1}\,dF(y) \\
&+ \int_0^{z-x} \big((\hat{z}-x-y)^{s-1} - (z-x-y)^{s-1}\big)\,d(\hat{F}-F)(y) \\
&+ \int_0^{z-x} \big((\hat{z}-x-y)^{s-1} - (z-x-y)^{s-1}\big)\,dF(y) \\
&+ \int_0^{z-x} (z-x-y)^{s-1}\,d(\hat{F}-F)(y).
\end{aligned} \tag{30}$$

It follows from (21) that $\hat{z} - z = O(N^{-1/2})$, and by standard properties of the empirical distribution, $\hat{F} - F = O(N^{-1/2})$. Thus the first two terms and the fourth are of order N^{-1}, and the others are of order $N^{-1/2}$.

The third term can be expressed as:

$$\int_{z-x}^{\hat z - x} (z-x-y)^{s-1}\, dF(y) = \int_0^{z-\hat z} u^{s-1} dF(z-x-u) = O(N^{-s/2}),$$

from which we see that it contributes asymptotically only if $s = 1$. In that case, the term is

$$F(\hat z - x) - F(z-x) = D^0(z-x)(\hat z - z) + O(N^{-1}),$$

since we made the definition $D^0 = F'$.

The fifth term is obviously zero for $s = 1$. For $s > 1$, it can be expressed as

$$\begin{aligned}(\hat z - z)\int_0^{z-x} \sum_{k=0}^{s-2} (\hat z - x - y)^k (z-x-y)^{s-2-k}\, dF(y) = \\ (\hat z - z)(s-1)\int_0^{z-x} (z-x-y)^{s-2}\, dF(y) + O(N^{-1}) = \\ (\hat z - z)(s-1)!\, D^{s-1}(z-x) + O(N^{-1}).\end{aligned} \tag{31}$$

We see that expression (31) serves for the fifth term when $s > 1$ and for the third when $s = 1$.

Finally, the sixth term is

$$\begin{aligned}&\int_0^{z-x} (z-x-y)^{s-1}\, d(\hat F - F)(y) \\ =&\frac{1}{N}\sum_{i=1}^{N}\Big((z-x-y_i)_+^{s-1} - E\big((z-x-y)_+^{s-1}\big)\Big),\end{aligned}$$

and so it is the average of N IID variables of mean zero. Multiplying (30) by $N^{1/2}$, we see that

$$\begin{aligned}N^{1/2}\big(\hat D^s(\hat z - x) - D^s(z-x)\big) = D^{s-1}(z-x)N^{1/2}(\hat z - z) \\ +\frac{1}{(s-1)!}N^{-1/2}\sum_{i=1}^{N}\Big((z-x-y_i)_+^{s-1} - E\big((z-x-y)_+^{s-1}\big)\Big).\end{aligned} \tag{32}$$

The result of the theorem follows from (32) by simple calculation. ∎

Proof of Theorem 3:

Consider the general problem in which, for some population, a value z is defined implicitly by $h(z) = 0$, where the function h is defined in terms of the population distribution. For instance, if $Q(p)$ is the p–quantile of a distribution with CDF F, we have $F(Q(p)) = p$, and we can set $h(z) = F(z) - p$.

For z_s, the defining relationship, in terms of the populations A and B, is $D^s_A(z_s) = D^s_B(z_s)$, with $D^s_A(x) > D^s_B(x)$ for all $x < z_s$. Thus we set $h(x) = D^s_A(x) - D^s_B(x)$. According to (24), $\hat{z}_s$ is defined in terms of $\hat{h}(x) \equiv \hat{D}^s_A(x) - \hat{D}^s_B(x)$. Under the assumption that z_s exists in the population and is less than the poverty line z, $\hat{z}_s$ is clearly a consistent estimator of z_s, and, in particular, we need not consider the possibility that $\hat{z}_s = z$, since this will happen with vanishingly small probability as $N \to \infty$.

The proof is similar for all values of s, and so we drop s from our notations. Since $h(z) = 0$, we have by Taylor expansion that

$$h(\hat{z}) = h'(\tilde{z})(\hat{z} - z) \tag{33}$$

for some $\tilde{z}$ such that $|\tilde{z} - z| < |\hat{z} - z|$. We will show later that

$$\hat{h}(z) + h(\hat{z}) = o(N^{-1/2}). \tag{34}$$

It was assumed that $h'(z) \neq 0$, and, in fact, since $h(x) > 0$ for $x < z$, and $h(z) = 0$, it follows that $h'(z) < 0$. Since $\hat{z} \to z$ as $N \to \infty$, we have that $\tilde{z} \to z$ as $N \to \infty$ as well. Thus for large enough N, $h'(\tilde{z}) \neq 0$. It follows from (33) and (34) that

$$\hat{z} - z = -\frac{\hat{h}(z)}{h'(\tilde{z})} + o(N^{-1/2}). \tag{35}$$

Suppose first that the populations A and B are independent, and that we have N_A drawings from one and N_B drawings from the other. For the purposes of the asymptotic analysis, we assume that the ratio $r = N_A/N_B$ remains constant as N_A and N_B tend to infinity. We have that

$$E\left((z - y^A)_+^{s-1}\right) = (s-1)!\, D^s_A(z) = (s-1)!\, D^s_B(z) = E\left(z - y^B\right)_+^{s-1}\Big)$$

because $h(z) = 0$. It follows that

$$\begin{aligned} N_A^{1/2}\hat{h}(z) = \frac{1}{(s-1)!}\biggl(N_A^{-1/2}\sum_{i=1}^{N_A}\Bigl((z - y_i^A)_+^{s-1} - E\left((z - y^A)_+^{s-1}\right)\Bigr) \\ - r^{1/2}N_B^{-1/2}\sum_{j=1}^{N_B}\Bigl((z - y_j^B)_+^{s-1} - E\left(z - y^B\right)_+^{s-1}\Bigr)\Bigr)\biggr). \end{aligned} \tag{36}$$

The expression (36) consists of two independent sums of IID variables to which we may apply the central limit theorem since moments of order

$2s-2$ are assumed to exist. It follows immediately that $N_A^{1/2}\hat{h}(z) = O(1)$ in probability, and, from (35), that $\hat{z}-z = O(N^{-1/2})$. In addition, from (1),

$$h'(z) = D_A^{s-1}(z) - D_B^{s-1}(z). \tag{37}$$

If $s = 1$, (37) remains correct because we defined $D_A^0(z) = F_A'(z)$, the density associated with the CDF F_A. We now see from (35) and (36) that

$$\lim_{N_A\to\infty} \operatorname{var}\big(N_A^{1/2}(\hat{z}-z)\big) = \frac{\operatorname{var}\big((z-y^A)_+^{s-1}\big) + r\operatorname{var}\big((z-y^B)_+^{s-1}\big)}{\Big(\big(D_A^{s-1}(z) - D_B^{s-1}(z)\big)(s-1)!\Big)^2}, \tag{38}$$

Next, suppose that we have N paired observations y_i^A and y_i^B from one single population. (36) continues to hold with $N_A = N$ and $r = 1$. However, the two sums of IID variables are no longer independent in general, and so (38) must be replaced by

$$\lim_{N\to\infty} \operatorname{var}\big(N^{1/2}(\hat{z}-z)\big) = \\ \frac{\operatorname{var}\big((z-y^A)_+^{s-1}\big) + \operatorname{var}\big((z-y^B)_+^{s-1}\big) - 2\operatorname{cov}\big((z-y^A)_+^{s-1}, (z-y^B)_+^{s-1}\big)}{\Big(\big(D_A^{s-1}(z) - D_B^{s-1}(z)\big)(s-1)!\Big)^2}. \tag{39}$$

It remains to prove (34). Note that, because $\hat{h}(\hat{z}) = h(z) = 0$,

$$-\big(\hat{h}(z) + h(\hat{z})\big) = \hat{h}(\hat{z}) - \hat{h}(z) - \big(h(\hat{z}) - h(z)\big). \tag{40}$$

Consider the expression

$$\hat{h}(z+\delta) - \hat{h}(z) - \big(h(z+\delta) - h(z)\big). \tag{41}$$

for nonrandom δ. In the case of just one population and N paired drawings of y^A and y^B, we can write

$$\hat{h}(z+\delta) - \hat{h}(z) = \frac{1}{N(s-1)!}\Big(\sum_{i=1}^N (z+\delta-y^A)_+^{s-1} - (z+\delta-y^B)_+^{s-1} \\ - (z-y^A)_+^{s-1} + (z-y^B)_+^{s-1}\Big).$$

The expectation of this is $h(z+\delta) - h(z)$, and so (41) is the average of bounded IID variables with mean zero and finite variance of order δ^2. Consequently, by the central limit theorem, (41) times $N^{1/2}$ has mean zero and variance of order δ^2. Since $\hat{z}-z = O(N^{-1/2})$ in probability, it follows that (40) times $N^{1/2}$ tends to zero in mean square, and hence in probability.

An exactly similar argument applies when there are two populations. ∎

Proof of Theorem 4:

We have for both distributions A and B that

$$\hat{G}(p;\hat{z}) = \int_0^\infty (z-y)I(y<\hat{z})\, I(y<\hat{Q}(p))\, d\hat{F}(y) \\ + (\hat{z}-z)\int_0^\infty I(y<\hat{z})\, I(y<\hat{Q}(p))\, d\hat{F}(y). \tag{42}$$

The second term on the right-hand side of this is

$$\begin{aligned} &(\hat{z}-z)\hat{F}\big(\min(\hat{z},\hat{Q}(p))\big) \\ &= (\hat{z}-z)\min\big(\hat{F}(\hat{z}),p\big) \\ &= (\hat{z}-z)\min\big(F(z),p\big) + O(N^{-1}), \end{aligned}$$

and the first term is

$$\int_0^{\hat{Q}(p)} (z-y)_+\, d\hat{F}(y).$$

This kind of integral can be expressed asymptotically as a sum of IID variables using a technique developed in Davidson and Duclos (1997). The term becomes

$$p\big(z-Q(p)\big)_+ + N^{-1}\sum_{i=1}^N I\big(y_i<Q(p)\big)\big((z-y_i)_+ - (z-Q(p))_+\big) + O(N^{-1}),$$

which to leading order is a deterministic term plus an average of IID random variables. We can combine the two terms in (42) using (21) to get

$$\hat{G}(p;\hat{z}) = p(z-Q(p))_+ + N^{-1}\sum_{i=1}^N \Big(I\big(y_i<Q(p)\big)\big((z-y_i)_+ - (z-Q(p))_+\big) \\ + \big(\xi(y_i)-z\big)\min\big(F(z),p\big)\Big) + O(N^{-1}). \tag{43}$$

If z is known and not estimated, we can just set $\xi(y_i) = z$, and the last term in the sum will vanish.

It is easy to check that, whether $z < Q(p)$ or $z > Q(p)$, the expectation of the leading term of the above expression is just $G(p;z)$. The fact that $\hat{G}_A(p;\hat{z}_A)$ and $\hat{G}_B(p;\hat{z}_B)$ are sums of independently and identically distributed random variables with finite second moments leads to their asymptotic normality by the central limit theorem. The covariance structure is obtained by simple calculation. ∎

References

Anderson, G. (1996). "Nonparametric Tests of Stochastic Dominance In Income Distributions" *Econometrica*, **64**, 1183-1193.

Atkinson, A.B. (1970). "On the Measurement of Inequality" *Journal of Economic Theory*, **2**, 244–263.

Atkinson, A.B. (1987). "On the Measurement of Poverty" *Econometrica*, **55**, 749–764.

Atkinson, A.B. (1991). "Measuring Poverty and Differences in Family Composition" *Economica*, **59**, 1–16.

Atkinson, A.B. (1995). *Incomes and the Welfare State*, Cambridge University Press, Cambridge.

Bahadur, R.R. (1966). "A Note on Quantiles in Large Samples" *Annals of Mathematical Statistics*, **37**, 577–80.

Beach, C.M. and R. Davidson (1983). "Distribution-Free Statistical Inference with Lorenz Curves and Income Shares" *Review of Economic Studies*, **50**, 723–735.

Beach, C.M. and J. Richmond (1985). "Joint Confidence Intervals for Income Shares and Lorenz Curves" *International Economic Review*, **26**, 439–50.

Bishop, J.A., S. Chakraborti and P.D. Thistle (1989). "Asymptotically Distribution–Free Statistical Inference for Generalized Lorenz Curves" *The Review of Economics and Statistics*, **LXXI**, 725–27.

Bishop, J.A., K.V. Chow and B. Zheng (1995). "Statistical Inference and Decomposable Poverty Measures" *Bulletin of Economic Research*, **47**, 329–340.

Bishop, J.A, J.P. Formby, and P. Thistle (1992). "Convergence of the South and Non-South Income Distributions, 1969–1979" *American Economic Review*, **82**, 262–272.

Bishop, J.A., J.P. Formby and B. Zheng (1997). "Statistical Inference and the Sen Index of Poverty" *International Economic Review*, **38**, 381-387.

Blackorby, C. and D. Donaldson (1978). "Measures of Relative Equality and Their Meaning in Terms of Social Welfare" *Journal of Economic Theory*, **18**, 59–80.

Blackorby, C. and D. Donaldson (1980). "Ethical Indices for the Measurement of Poverty" *Econometrica*, **48**, 1053–1062.

Bourguignon, F. and G. Fields (1997). "Discontinuous Losses from Poverty, Generalized Measures, and Optimal Transfers to the Poor" *Journal of Public Economics*, **63**, 155–175.

Chakravarty, S.R. (1983). "A New Index of Poverty" *Mathematical Social Sciences*, **6**, 307–313.

Chakravarty, S.R. (1990). *Ethical Social Index Numbers*, New York, Springer-Verlag.

Chakravarty, S.R. (1997). "On Shorrocks' Reinvestigation of the Sen Poverty Index" *Econometrica*, **65**, 1241–1242.

Chow, K.V. (1989). *Statistical Inference for Stochastic Dominance: a Distribution Free Approach*, Ph.D. thesis, University of Alabama.

Clark, S., R. Hamming and D. Ulph (1981). "On Indices for the Measurement of Poverty" *The Economic Journal*, **91**, 515–526.

Cowell, F.A. (1989). "Sampling Variance and Decomposable Inequality Measures" *Journal of Econometrics*, **42**, 27–41.

Cowell, F.A. and M.-P. Victoria-Feser (1996). "Poverty Measurement with Contaminated Data: a Robust Approach" *European Economic Review*, **40**, 1761–1771.

Dardanoni, V. and A. Forcina (1999). "Inference for Lorenz Curve Orderings" *Econometrics Journal*, **2**, 48–74.

Davidson, R. and J.-Y. Duclos (1997). "Statistical Inference for the Measurement of the Incidence of Taxes and Transfers" *Econometrica*, **65**, 1453–1466.

Fishburn, P.C., and R.D. Willig (1984). "Transfer Principles in Income Redistribution" *Journal of Public Economics*, **25**, 323–328.

Formby, J.P., W.J. Smith and B. Zheng (1998). "Inequality Orderings, Stochastic Dominance and Statistical Inference" paper presented at the 1998 Chicago Meeting of the Econometric Society.

Foster, J.E., (1984). "On Economic Poverty: A Survey of Aggregate Measures" in R.L. Basmann and G.F. Rhodes, eds., *Advances in Econometrics*, **3**, Connecticut, JAI Press, 215–251.

Foster, J.E., J. Greer and E. Thorbecke (1984). "A Class of Decomposable Poverty Measures" *Econometrica*, **52**, 761–776.

Foster, J.E. and A. Sen (1997). "On Economic Inequality after a Quarter Century" in A. Sen, *On Economic Inequality*, (expanded edition), Oxford: Clarendon Press.

Foster, J.E. and A.F. Shorrocks (1988a). "Poverty Orderings" *Econometrica*, **56**, 173–177.

Foster, J.E. and A.F. Shorrocks (1988b). "Poverty Orderings and Welfare Dominance" *Social Choice Welfare*, **5**, 179–198.

Foster, J.E. and A.F. Shorrocks (1988c). "Inequality and Poverty Orderings" *European Economic Review*, **32**, 654–662.

Gottschalk, P. and T.M. Smeeding (1997). "Cross-National Comparisons of Earnings and Income Inequality" *Journal of Economic Literature*, **35**, 633–687.

Greer, J. and E. Thorbecke (1986). "A Methodology for Measuring Food Poverty Applied to Kenya" *Journal of Development Economics*, **24**, 59–74.

Gustafsson, B. and L. Nivorozhkina (1996). "Relative Poverty in Two Egalitarian Societies: A Comparison Between Taganrog, Russia During the Soviet Era and Sweden" *Review of Income and Wealth*, **42**, 321–334.

Hagenaars A. (1987). "A Class of Poverty Indices" *International Economic Review*, **28**, 583–607.

Horowitz, J.L. (1997). "Bootstrap Methods in Econometrics: Theory and Numerical Performance," in David M. Kreps and Kenneth F. Wallis, eds., *Advances in Economics and Econometrics: Theory and Applications*, Vol. 3, pp. 188–222, Cambridge, Cambridge University Press.

Howes, S. (1993). "Asymptotic Properties of Four Fundamental Curves of Distributional Analysis" Unpublished paper, STICERD, London School of Economics.

Howes, S., and J.O. Lanjouw (1998). "Does Sample Design Matter for Poverty Rate Comparisons?" *Review of Income and Wealth*, **44**, 99–109.

Jenkins, S.P. and P.J. Lambert (1997). "Three 'I's of Poverty Curves, With an Analysis of UK Poverty Trends" *Oxford Economic Papers*, **49**, 317–327.

Jenkins, S.P. and P.J. Lambert (1998). "Three 'I's of Poverty Curves and Poverty Dominance: TIPs for Poverty Analysis" *Research on Economic Inequality*, **8**, 39–56.

Kakwani, N. (1980). "On a Class of Poverty Measures" *Econometrica*, **48**, 437–446.

Kakwani, N. (1993). "Statistical Inference in the Measurement of Poverty" *Review of Economics and Statistics*, **75**, 632–639.

Kaur, A., B.L. Prakasa Rao, and H. Singh (1994). "Testing for Second-Order Stochastic Dominance of Two Distributions" *Econometric Theory*, **10**, 849–866.

Kodde, D.A. and F.C. Palm (1986). "Wald Criteria for Jointly Testing Equality and Inequality Restrictions" *Econometrica*, **54**, 1243–1248.

Kolm, S.C. (1976). "Unequal Inequalities, I" *Journal of Economic Theory*, **12**, 416–42.

McFadden, D. (1989). "Testing for Stochastic Dominance" in *Studies in the Economics of Uncertainty*, T.B. Fomby and T.K. Seo, editors, Springer-Verlag, New York.

Mehran, F. (1976). "Linear Measures of Income Inequality." *Econometrica*, **44**, 805–809.

Preston, I. (1995). "Sampling Distributions of Relative Poverty Statistics" *Applied Statistics*, **44**, 91–99.

Ravallion, M. (1994). *Poverty Comparisons*, Fundamentals of Pure and Applied Economics, Harwood Academic Publishers, Switzerland.

Ravallion, M. and B. Bidani (1994). "How Robust is a Poverty Profile?" *The World Bank Economic Review*, **8**, 75–102.

Richmond, J. (1982). "A General Method for Constructing Simultaneous Confidence Intervals" *Journal of the Americal Statistical Association*, **77**, 455–460.

Robertson, T., F.T. Wright, and R.L. Dykstra (1988). *Order Restricted Statistical Inference*, Wiley, New York.

Rongve, I. (1997). "Statistical Inference for Poverty Indices with Fixed Poverty Lines" *Applied Economics*, **29**, 387–392.

Sen, A. (1976). "Poverty: An Ordinal Approach to Measurement" *Econometrica*, **44**, 219–231.

Sen, A., (1981). *Poverty and Famine: An Essay on Entitlement and Deprivation*, Clarendon Press, Oxford; and Oxford University Press, New York.

Shorrocks, A.F. (1983). "Ranking Income Distributions" *Economica*, **50**, 3-17.

Shorrocks, A.F. (1987). "Transfer Sensitive Inequality Measures" *Review of Economic Studies*, **LIV**, pp. 485–497.

Shorrocks, A.F. (1995). "Revisiting the Sen Poverty Index" *Econometrica*, **63**, 1225–1230.

Shorrocks, A.F. (1998). “Deprivation Profiles and Deprivation Indices” ch. 11 in *The Distribution of Household Welfare and Household Production*, ed. S. Jenkins *et al*, Cambridge University Press.

Smeeding, T, L. Rainwater, and M. O'Higgins (1990). *Poverty, Inequality, and the Distribution of Income in an International Context: Initial Research from the Luxembourg Income Study (LIS)*, Wheatsheaf Books, London.

Spencer, B. and S. Fisher (1992). “On Comparing Distributions of Poverty Gaps” *Sankhya: The Indian Journal of Statistics Series B*, **54**, 114–126.

Summers, R. and A. Heston (1991). “The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988” *The Quarterly Journal of Economics*, **106**, 327–368.

Takayama, N. (1979). “Poverty, Income Inequality, and their Measures: Professor Sen's Axiomatic Approach Reconsidered” *Econometrica*, **47**, 747–759.

Thon, D. (1979). “On Measuring Poverty” *Review of Income and Wealth*, **25**, 429–439.

Van den Bosch, K., T. Callan, J. Estivill, P. Hausman, B. Jeandidier, R. Muffels, and J. Yfantopoulos (1993). “A Comparison of Poverty in Seven European Countries and Regions Using Subjective and Relative Measures” *Journal of Population Economics*, **6**, 235–259.

Watts, H.W. (1968). “An Economic Definition of Poverty” in D.P. Moynihan (ed.), *On Understanding Poverty*, New York: Basic Books.

Wolak, F.A. (1989). “Testing Inequality Constraints in Linear Econometric Models” *Journal of Econometrics*, **41**, 205–235.

Wolak, F.A. (1991). “The Local Nature of Hypothesis Tests involving Inequality Constraints in Nonlinear Models” *Econometrica*, **59**, 981–995.

Xu, K. and L. Osberg (1998). “A Distribution-Free Test for Deprivation Dominance” Department of Economics, Dalhousie University, Halifax.

Zheng, B. (1997a). “Aggregate Poverty Measures” *Journal of Economic Surveys*, **11**, 123–63.

Zheng, B. (1997b). “Statistical Inferences for Poverty Measures with Relative Poverty Lines” mimeo, Department of Economics, University of Colorado.

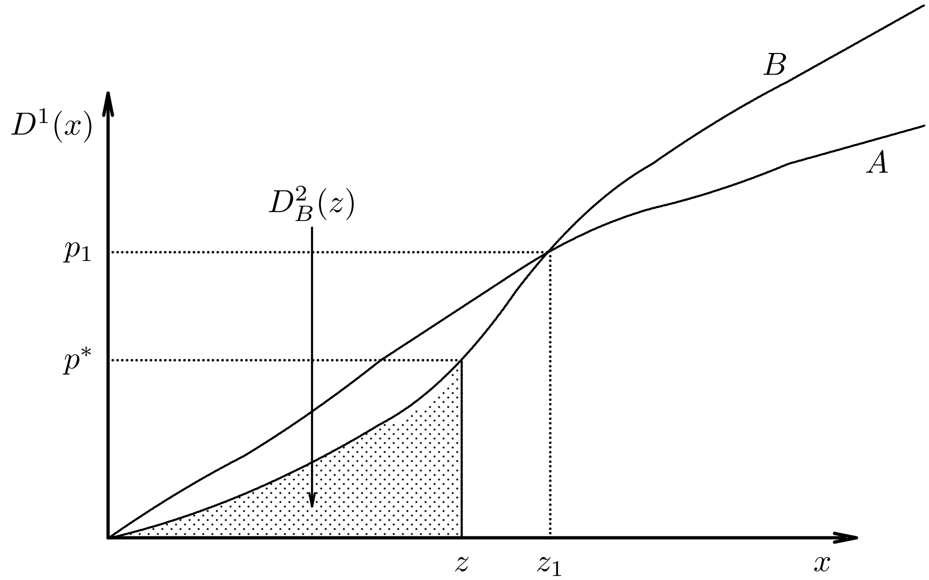

Figure 1. Poverty incidence curves for two distributions A and B

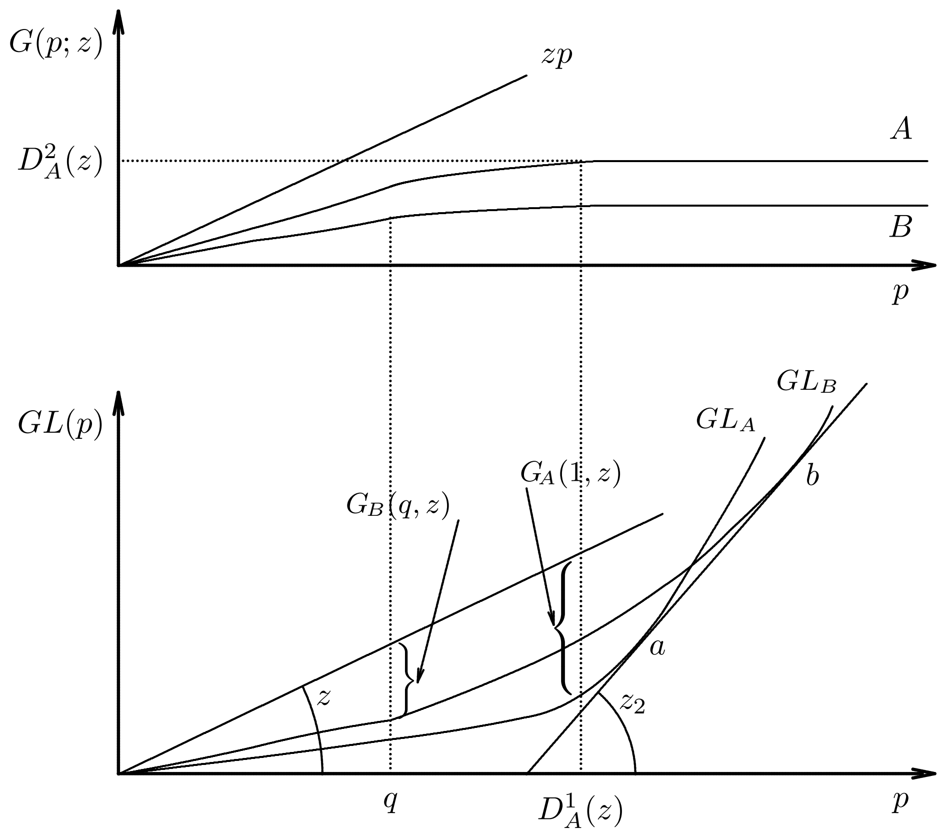

Figure 2. CPG and Generalised Lorenz Curves

Table 1.
Headcounts and average poverty gaps for various poverty lines

x	USA	Canada	Norway	Netherlands
2000	0.0184 (0.001) 23.0 (1.5)	0.0071 (0.0006) 8.7 (0.8)	0.0059 (0.0009) 8.5 (1.3)	0.0234 (0.002) 34.7 (3.7)
4000	0.0461 (0.002) 81.2 (3.7)	0.0162 (0.0009) 31.4 (2.0)	0.0122 (0.001) 27.4 (3.2)	0.0332 (0.003) 90 (8.2)
6000	0.104 (0.002) 227.7 (6.9)	0.042 (0.001) 86.7 (3.7)	0.026 (0.002) 64.7 (5.6)	0.061 (0.004) 177.4 (13.4)
8000	0.176 (0.003) 505.5 (11.1)	0.089 (0.002) 216.8 (6.2)	0.086 (0.003) 172.1 (9.1)	0.159 (0.006) 374.8 (19.8)
10000	0.250 (0.003) 933 (16)	0.149 (0.002) 453 (10)	0.173 (0.004) 429 (14)	0.310 (0.007) 843 (28)
15000	0.451 (0.004) 2694 (31)	0.366 (0.003) 1714 (21)	0.511 (0.006) 2101 (33)	0.660 (0.007) 3313 (54)
20000	0.625 (0.004) 5397 (45)	0.584 (0.003) 4112 (33)	0.796 (0.004) 5447 (50)	0.856 (0.005) 7153 (74)
25000	0.761 (0.003) 8884 (58)	0.751 (0.003) 7478 (44)	0.927 (0.003) 9811 (60)	0.944 (0.003) 11702 (85)
30000	0.854 (0.003) 12941 (60)	0.859 (0.002) 11531 (53)	0.970 (0.002) 14581 (65)	0.973 (0.002) 16506 (92)
35000	0.908 (0.002) 17362 (75)	0.923 (0.002) 16004 (58)	0.984 (0.001) 19468 (69)	0.985 (0.002) 21410 (97)

Notes: The first item in each box is $\hat{D}^1(x)$, beneath is its asymptotic standard error. Next is $\hat{D}^2(x)$, with its asymptotic standard error underneath. All amounts are in 1991 adult-equivalent US$. Data are for 1991, from the LIS data base.

Table 2.

Estimates of the thresholds z_s for dominance by three countries over the US

s	Canada	Norway	Netherlands $[z_s^-, z_s^+]$
$s = 1$	27840 (1575)	13190 (197)	$[2958, 8470]$ (193) (203)
$s = 2$		19708 (389)	$[4504, 11095]$ (486) (389)
$s = 3$		28051 (791)	$[6128, 13835]$ (741) (716)
$s = 4$		37533 (1299)	$[7839, 16530]$ (1071) (1145)

Asymptotic standard errors in parentheses. All amounts are in 1991 adult-equivalent US$. Data are for 1991, from the LIS data base.

Table 3.

Poverty ranking of the US, Canada, and the Netherlands with poverty line of half median income

x	$s=1$			$s=2$		
	Most Poverty	Medium Poverty	Least Poverty	Most Poverty	Medium Poverty	Least Poverty
7000	**USA** 0.012 (0.001)	**CAN** 0.0070 (0.0006)	**NL** 0 -	**USA** 10.2 (1.1)	**CAN** 8.2 (0.8)	**NL** 0 -
6000	**USA** 0.020 (0.001)	**NL** 0.012 (0.002)	**CAN** 0.0108 (0.0007)	**USA** 26.1 (2.1)	**CAN** 17.3 (1.4)	**NL** 3.1 (0.8)
5000	**USA** 0.030 (0.002)	**NL** 0.021 (0.002)	**CAN** 0.016 (0.001)	**USA** 50.0 (3.4)	**CAN** 30.4 (2.1)	**NL** 18.8 (2.5)
4000	**USA** 0.050 (0.002)	**CAN** 0.027 (0.001)	**NL** 0.024 (0.002)	**USA** 88.9 (5.3)	**CAN** 51.2 (3.0)	**NL** 41.0 (4.6)
3000	**USA** 0.077 (0.003)	**CAN** 0.041 (0.002)	**NL** 0.028 (0.003)	**USA** 152.1 (7.7)	**CAN** 84.2 (4.2)	**NL** 67.4 (6.8)
2000	**USA** 0.110 (0.004)	**CAN** 0.065 (0.002)	**NL** 0.035 (0.003)	**USA** 245 (11)	**CAN** 136.1 (5.8)	**NL** 99.1 (9.3)
1000	**USA** 0.144 (0.004)	**CAN** 0.088 (0.002)	**NL** 0.045 (0.004)	**USA** 372 (14)	**CAN** 211.6 (7.7)	**NL** 139 (12)
0	**USA** 0.181 (0.005)	**CAN** 0.116 (0.003)	**NL** 0.067 (0.005)	**USA** 534 (18)	**CAN** 312.8 (9.9)	**NL** 194 (15)

Rankings based on $D^s(z-x)$ for $s=1,2$. Asymptotic standard errors in parentheses. All amounts are in 1991 adult-equivalent US$. Data are for 1991, from the LIS data base.