Beach, Charles M.; Davidson, Russell; Slotsve, George A.

Working Paper

Distribution-Free Statistical Inference for Inequality Dominance with Crossing Lorenz Curves

Queen's Economics Department Working Paper, No. 912

Provided in Cooperation with:
Queen's University, Department of Economics (QED)

Suggested Citation: Beach, Charles M.; Davidson, Russell; Slotsve, George A. (1994) : Distribution-Free Statistical Inference for Inequality Dominance with Crossing Lorenz Curves, Queen's Economics Department Working Paper, No. 912, Queen's University, Department of Economics, Kingston (Ontario)

This Version is available at:
https://hdl.handle.net/10419/189233

Queen's Economics Department Working Paper No. 912

Distribution-Free Statistical Inference for Inequality Dominance with Crossing Lorenz Curves

Charles M. Beach Russell Davidson George A. Slotsve

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

11-1994

Discussion Paper #912

Distribution-Free Statistical Inference for Inequality Dominance with Crossing Lorenz Curves

by

Charles M. Beach
Queen's University

Russell Davidson
Queen's University & GREQE-EHESS

George A. Slotsve
Vanderbilt University

November 1994

Distribution-Free Statistical Inference for Inequality Dominance with Crossing Lorenz Curves

by

Charles M. Beach

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

Russell Davidson

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

GREQE–EHESS
Centre de la Vieille Charité
2 Rue de la Charité
13002 Marseille, France

and

George A. Slotsve

Department of Economics
Vanderbilt University
Nashville, TN, USA
37235

Abstract

Distribution-free techniques of statistical inference are developed for the cumulative coefficients of variation of an income distribution, thus allowing one to test for inequality dominance when Lorenz curves cross. The full covariance structure of the cumulative sample means and variances is worked out. As an illustration, the procedures are applied to the 1984 and 1990 earnings distributions of male paid workers in the United States, and it is found that the 1990 distribution was significantly less unequal than the 1984 distribution.

This research was supported, in part, by grants from the Social Sciences and Humanities Research Council of Canada. Subject to the usual disclaimers, the authors would like to thank Dan Bernhardt, Jean-Yves Duclos, Michael Hoy, and James Foster for helpful comments and advice.

November, 1994

1. Introduction

There has been considerable recent interest, in the United States and elsewhere, in the distribution of income, and changes induced in this distribution by the marked changes that have been observed in the labor market and family structure. (Blackburn and Bloom (1993), Bound and Johnson (1992), Karoly (1992), Levy and Murnane (1992), and Murphy and Welch (1992).) While techniques of analysis have varied widely across studies, there has been a growing use of various disaggregated dominance criteria, such as first- and second-order stochastic dominance, and Lorenz dominance, in order to rank income or earnings distributions. These criteria are based on very general principles, and by using them one avoids the temptation to draw conclusions that may be sensitive to the choice of a particular summary measure of inequality. (Bishop, Formby, and Smith (1991), Bishop, Formby, and Thistle (1992), Beach and Slotsve (1994), Lambert (1989), Howes (1993), and Richardson (1994).)

Parallel to the development of the economic welfare theory underlying the use of dominance criteria has been the development and application of the statistical theory needed to perform statistical inference in the context of the use of these criteria. As a result, researchers can make inferences on the basis of sets of sample data as to whether, according to some chosen criterion, the distribution from which one of the samples was drawn dominates another in a statistically significant manner. (Anderson (1994), Beach and Davidson (1983), Beach, Chow, Formby, and Slotsve (1994), Bishop, Formby, and Smith (1991), Bishop, Formby, and Thistle (1992), Howes (1994), and Xu (1994).)

One of the most frequently used dominance criteria for judging differences in inequality between income distributions is Lorenz dominance. This criterion is met if the Lorenz curve for one distribution dominates that for another (Atkinson (1970)). In this case, that is, if the Lorenz curve for a distribution f lies everywhere above that for another distribution g, then *any* aggregate inequality measure I, such as the Gini coefficient, that satisfies symmetry, mean independence, population homogeneity, and the Dalton-Pigou principle of transfers, will rank the income inequality of the distribution f lower than that of g: $I(f) < I(g)$ (Jenkins (1991)). Thus, if one distribution Lorenz dominates another, *all* aggregate inequality measures satisfying the above properties will agree in their ranking of the two distributions.

What, if anything, can be inferred if the Lorenz curves of two distributions cross, as they often do in empirical studies? Shorrocks and Foster (1987) show that, if the above conditions are strengthened by including a property that they call transfer sensitivity, then there exists a dominance

criterion which, if met, ensures that all inequality measures satisfying the strengthened conditions will agree in their rankings, *provided* that the two Lorenz curves have only a single crossing. Davies and Hoy (1994) generalize the Shorrocks-Foster result to provide for any number n of crossings of two Lorenz curves. They show that, for two distributions f and g, with, in general, different means, the following statements are equivalent:

a) $I(f) < I(g)$ for all inequality measures satisfying the conditions of Shorrocks and Foster (they refer to the Shorrocks-Foster "transfer sensitivity" as "aversion to downside inequality"); and

b) For all crossovers $i = 1, 2, \ldots, n+1$, $CV_i(f) \leq CV_i(g)$, where $CV_i(\cdot)$ denotes the cumulative coefficient of variation for incomes up to the i^{th} crossover point.

The $(n+1)^{\text{th}}$ crossover point is at $(1, 1)$, where all Lorenz curves "cross". Thus a necessary condition for transfer sensitivity dominance is that the more unequal distribution have higher unconditional variance.

In order to implement the test of Davies and Hoy on sample data, it is necessary to know the sampling distribution of any set of cumulative coefficients of variation for an income distribution. In this paper, we provide the statistical basis for testing for inequality dominance when Lorenz curves cross, by establishing the (asymptotic) sampling distribution of a vector of cumulative coefficients of variation for a distribution, under quite general conditions. The results are distribution-free in the sense that they do not require knowledge of the underlying population distribution from which the sample income data are drawn.

In the next section, the joint asymptotic variance-covariance structure is obtained for the cumulative means and variances of samples drawn from a given population, from which the distribution of the cumulative coefficients of variation can easily be calculated. In Section 3, these results are applied to samples of earnings of paid white U.S. males in 1984 and 1990. Although the earnings distribution shifted very little from one year to the other, it is still possible to show that, by the criterion of transfer sensitivity dominance, the 1990 distribution was significantly less unequal than the 1984 one. Section 4 concludes.

2. Asymptotic Distribution of the Conditional Coefficients of Variation

Let Y denote the random variable income for individuals or families, and let the population c.d.f of Y be $F(y)$, which is assumed to be at least once continuously differentiable. All incomes are assumed to be positive, and the

mean and variance of Y, μ and σ^2, are assumed to exist and be finite. We also assume that the third and fourth moments of Y exist and are finite.

Corresponding to any p, $0 \leq p \leq 1$, one defines the $p-$quantile ξ_p of the distribution F by the relation $F(\xi_p) = p$. In order that ξ_p be well defined for all p, we assume that F is strictly monotonic, as well as being differentiable. One may then define its inverse function, G say, with the properties

$$\begin{aligned} F(G(p)) &= p && \text{for } 0 \leq p \leq 1; \\ G(F(y)) &= y && \text{for positive income level } y; \text{ and} \\ \xi_p &= G(p) && \text{for any quantile.} \end{aligned}$$

Now select a set of K proportions, p_i, $i = 1, \ldots, K$, with $0 < p_1 < \ldots < p_K$. For deciles, for example, one would choose $K = 10$ and $p_1 = 0.1$, $p_2 = 0.2$, ... , $p_{10} = 1.0$. Then, corresponding to this set of proportions, we have a set of K population income quantiles, $\xi_{p_1} < \xi_{p_2} < \ldots \xi_{p_K}$, a set of K *cumulative means*, γ_i, defined by the equation

$$\gamma_i \equiv E\big(Y \mid Y \leq \xi_{p_i}\big) = \frac{1}{p_i} \int_0^{G(p_i)} y \, dF(y),$$

and a set of K *cumulative variances*, λ_i^2, defined by

$$\phi_i \equiv \lambda_i^2 + \gamma_i^2 \equiv E\big(Y^2 \mid Y \leq \xi_{p_i}\big) = \frac{1}{p_i} \int_0^{G(p_i)} y^2 \, dF(y),$$

It often turns out to be convenient to work with the set of *cumulative uncentered second moments*, ϕ_i.

Let a random sample of size N be taken from the population and let the observations be ordered by size from the smallest $(Y_{(1)})$ to the largest $(Y_{(N)})$. Then the *sample quantile*, $\hat{\xi}_p$, is defined as the r^{th} order statistic, $Y_{(r)}$, where $r = [Np]$ denotes the greatest integer not greater than Np. Since F is strictly monotonic, $\hat{\xi}_p$ has the property of strong or almost sure consistency: $\lim_{n \to \infty} \hat{\xi}_p = \xi_p$ with probability 1; see Rao (1965), pg 355. The sample estimates of the cumulative means are given by

$$\hat{\gamma}_i \equiv \frac{1}{r_i} \sum_{j=1}^{r_i} Y_{(j)}, \qquad i = 1, 2, \ldots, K, \tag{1}$$

where $r_i = [Np_i]$, and the sample estimates of the cumulative uncentered second moments are given by

$$\hat{\phi}_i \equiv \frac{1}{r_i} \sum_{j=1}^{r_i} Y_{(i)}^2, \qquad i = 1, 2, \ldots, K. \tag{2}$$

Similarly, corresponding to the third and fourth moments we define

$$p_i\chi_i \equiv \int_0^{G(p_i)} y^3\, dF(y), \qquad \text{and}$$
$$p_i\psi_i \equiv \int_0^{G(p_i)} y^4\, dF(y),$$

where the sample estimates of the cumulative uncentered third and fourth moments are given by

$$\hat{\chi}_i \equiv \frac{1}{r_i}\sum_{j=1}^{r_i} Y_{(j)}^3, \qquad i = 1, 2, \ldots, K, \qquad \text{and}$$
$$\hat{\psi}_i \equiv \frac{1}{r_i}\sum_{j=1}^{r_i} Y_{(j)}^4, \qquad i = 1, 2, \ldots, K.$$

The objective of this paper is to perform (asymptotic) statistical inference with the vector of sample conditional coefficients of variation,

$$\boldsymbol{\Phi}^\top \equiv \left[\begin{array}{cccc} \frac{\hat{\lambda}_1}{\hat{\gamma}_1} & \frac{\hat{\lambda}_2}{\hat{\gamma}_2} & \cdots & \frac{\hat{\lambda}_K}{\hat{\gamma}_K} \end{array}\right].$$

However, since $\hat{\lambda}_i^2 = \hat{\phi}_i - \hat{\gamma}_i^2$, this vector can be written as

$$\left[\begin{array}{cccc} \frac{\sqrt{\hat{\phi}_1 - \hat{\gamma}_1^2}}{\hat{\gamma}_1} & \frac{\sqrt{\hat{\phi}_2 - \hat{\gamma}_2^2}}{\hat{\gamma}_2} & \cdots & \frac{\sqrt{\hat{\phi}_K - \hat{\gamma}_K^2}}{\hat{\gamma}_K} \end{array}\right].$$

Thus it is necessary to establish the joint distribution of the $\hat{\gamma}_i$ and $\hat{\phi}_i$.

The first theorem is the following:

Theorem 1: Suppose that the $2K-$random vector

$$\hat{\theta}^\top \equiv \left[\hat{\theta}_1^\top \vdots \hat{\theta}_2^\top\right] = [p_1\hat{\gamma}_1 \quad \ldots \quad p_K\hat{\gamma}_K \quad \vdots \quad p_1\hat{\phi}_1 \quad \ldots \quad p_K\hat{\phi}_K]$$

where the $\hat{\gamma}_i$, $i = 1, \ldots, K$, are the conditional sample means defined in (1), the proportions p_i are such that $0 < p_1 < p_2 < \ldots < p_{K-1} < p_K = 1$, and the $\hat{\phi}_i$, $i = 1, \ldots, K$, are the conditional uncentered second moments defined in (2). Then, under the conditions that the first four moments of the population are finite, and that the c.d.f. F is strictly monotonic and continuously differentiable, $\hat{\theta}$ is asymptotically normal, in that $N^{1/2}(\hat{\theta} - \theta)$ has a limiting $2K-$variate normal distribution with mean zero and covariance matrix Ω, where

$$\Omega = \left[\begin{array}{cc} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{array}\right] \tag{3}$$

and where the (i,j) elements of the four submatrices are, for $i \leq j$:

$$(\Omega_{11})_{ij} = p_i[\phi_i - \gamma_i^2 + (1-p_j)(G(p_i)-\gamma_i)(G(p_j)-\gamma_j) + (G(p_i)-\gamma_i)(\gamma_j-\gamma_i)] \quad (4)$$

$$(\Omega_{12})_{ij} = p_i[\chi_i - \gamma_i\phi_i + (1-p_j)(G(p_i)-\gamma_i)(G^2(p_j)-\phi_j) + (G(p_i)-\gamma_i)(\phi_j-\phi_i)] \quad (5)$$

$$(\Omega_{21})_{ij} = p_i[\chi_i - \gamma_i\phi_i + (1-p_j)(G^2(p_i)-\phi_i)(G(p_j)-\gamma_j) + (G^2(p_i)-\phi_i)(\gamma_j-\gamma_i)] \quad (6)$$

$$(\Omega_{22})_{ij} = p_i[\psi_i - \phi_i^2 + (1-p_j)(G^2(p_i)-\phi_i)(G^2(p_j)-\phi_j) + (G^2(p_i)-\phi_i)(\phi_j-\phi_i)] \quad (7)$$

The elements of the off-diagonal blocks for which $i > j$ can be derived from the fact that the entire Ω matrix is symmetric. In fact, for $i > j$:

$$(\Omega_{12})_{ij} = (\Omega_{21})_{ji} \qquad \text{and} \qquad (\Omega_{21})_{ij} = (\Omega_{12})_{ji}.$$

Proof: See Appendix.

For the asymptotic distribution of the Lorenz curve ordinates and the conditional coefficients of variation, we have

$$\hat{\Phi}^\top = [\hat{\Phi}_1^\top \ \vdots \ \hat{\Phi}_2^\top]$$
$$= \left[\frac{p_1\hat{\gamma}_1}{\hat{\mu}} \ \cdots \ \frac{p_K\hat{\gamma}_K}{\hat{\mu}} \ \vdots \ \frac{\sqrt{\hat{\phi}_1 - \hat{\gamma}_1^2}}{\hat{\gamma}_1} \ \cdots \ \frac{\sqrt{\hat{\phi}_K - \hat{\gamma}_K^2}}{\hat{\gamma}_K} \right].$$

We now use a standard result in Rao (1965) (pg 231) on limiting distributions of differentiable functions of random variables. If $N^{1/2}(\hat{\theta}-\theta)$ (where $\theta = [\theta_1 \vdots \theta_2]$) has a multivariate normal limiting distribution with mean zero and covariance matrix Ω specified by (3), then the limiting distribution of $N^{1/2}(\hat{\Phi}-\Phi)$ is also multivariate normal with mean zero and covariance matrix $V_L = J\Omega J^\top$, where

$$J = \left[\frac{\partial \Phi_i}{\partial \theta_j}\right] = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix}.$$

The submatrices are given by

$$J_{11} = \frac{\partial \Phi_1}{\partial \theta_1} = \begin{bmatrix} 1/\mu & & \vdots & -\frac{p_1\gamma_1}{\mu^2} \\ & \ddots & \vdots & \vdots \\ & 1/\mu & \vdots & -\frac{p_{K-1}\gamma_{K-1}}{\mu^2} \\ 0 \quad \ldots & 0 & \vdots & 0 \end{bmatrix};$$

$$J_{12} = \frac{\partial \Phi_1}{\partial \theta_2} = [0];$$

$$J_{21} = \frac{\partial \Phi_2}{\partial \theta_1} = \begin{bmatrix} \frac{-\phi_1}{p_1\lambda_1\gamma_1^2} & & \\ & \ddots & \\ & & \frac{-\phi_K}{p_K\lambda_K\gamma_K^2} \end{bmatrix};$$

$$J_{22} = \frac{\partial \Phi_2}{\partial \theta_2} = \begin{bmatrix} \frac{1}{2p_1\lambda_1\gamma_1} & & \\ & \ddots & \\ & & \frac{1}{2p_K\lambda_K\gamma_K} \end{bmatrix};$$

where $p_K = 1$ and $\gamma_K = \mu$. It is worth noting that γ_i is a function of the conditional first moments alone, and that ϕ_i is a function of the conditional second moments alone. Thus

$$\frac{\partial \lambda_i^2(\phi_i)}{\partial \phi_i} = 1 \qquad \text{and} \qquad \frac{\partial \lambda_i^2(\phi_i)}{\partial \gamma_i} = 0.$$

As a result J_{12} is a zero matrix.

From the above results, we obtain that

$$V_L = \begin{bmatrix} V_{L_{11}} & V_{L_{12}} \\ V_{L_{21}} & V_{L_{22}} \end{bmatrix}, \tag{8}$$

with

$$\begin{aligned} V_{L_{11}} &= J_{11}\Omega_{11}J_{11}^{\top}; \\ V_{L_{12}} &= J_{11}\Omega_{11}J_{21}^{\top} + J_{11}\Omega_{12}J_{22}^{\top}; \\ V_{L_{21}} &= J_{21}\Omega_{11}J_{11}^{\top} + J_{22}\Omega_{21}J_{11}^{\top}; \\ V_{L_{22}} &= J_{21}\Omega_{11}J_{21}^{\top} + J_{22}\Omega_{21}J_{21}^{\top} + J_{21}\Omega_{12}J_{22}^{\top} + J_{22}\Omega_{22}J_{22}^{\top}. \end{aligned}$$

This provides the second result of the paper, which we state as Theorem 2.

Theorem 2: Under the conditions of Theorem 1, the vector $\hat{\Phi}$ is asymptotically normal, in the sense that $N^{1/2}(\hat{\Phi} - \Phi)$ has a limiting $2K-$variate normal distribution with mean zero and covariance matrix V_L specified in (8).

Asymptotic standard errors for the sample conditional coefficients of variation are given by

$$\left(\frac{V_{L_{22}})_{ii}}{N}\right)^{\frac{1}{2}} \quad \text{for } i = 1, \ldots, K.$$

Several things about these results are worthy of note. They are all distribution-free, in the sense that estimation of V_L does not require knowledge of the underlying distribution from which the data were drawn. It depends solely on the proportions p_i, the unconditional mean and variance μ and σ^2, the income quantiles ξ_p, the conditional means and variances γ_i and ϕ_i^2, and the uncentered third and fourth moments χ_i and ψ_i. These can all be estimated consistently from the sample without prior specification of the population distribution underlying the sample data. It is straightforward to write a computer program – see Beach and Slotsve (1994) – that integrates these calculations and provides the results needed for performing first-order, second-order, and Lorenz dominance tests.

3. Illustrative Example: U.S. Men's Earnings, 1984 and 1990

We now illustrate the calculations of the previous section with distribution data on the earnings of male paid workers in the United States for 1984 and 1990. The data come from the 1985 and 1991 CPS micro data files on individual male income recipients aged 16–62 with nonzero earnings. Self-employed workers were excluded. There are 34,896 observations in 1984 and 34,562 observations in 1990. Reported earnings figures that had been flagged or "hot-decked" by the Census Bureau were replaced by predictions from a conventional earnings regression (estimated from the non-allocated observations) as suggested by Lillard, Smith, and Welch (1986). The one or two percent of the earnings figures in the samples that had been top-coded were also replaced by the conditional mean earnings figures from a Pareto curve fitted over the top vigintile of earnings recipients. (Full details of these procedures and accompanying SAS programs can be obtained from the authors.) All earnings figures are expressed in real terms (using the CPI for urban consumers, with the base of 100 corresponding to the average for 1982–84).

Estimates of Lorenz curve ordinates and conditional coefficients of variation are reported in Table 1. The curves for the two years are sufficiently

similar that, if plotted in the usual manner, they would appear to coincide over most of their length. In Figure 1 accordingly, we plot the difference between the two Lorenz curves as a function of the usual Lorenz curve abscissa, p.

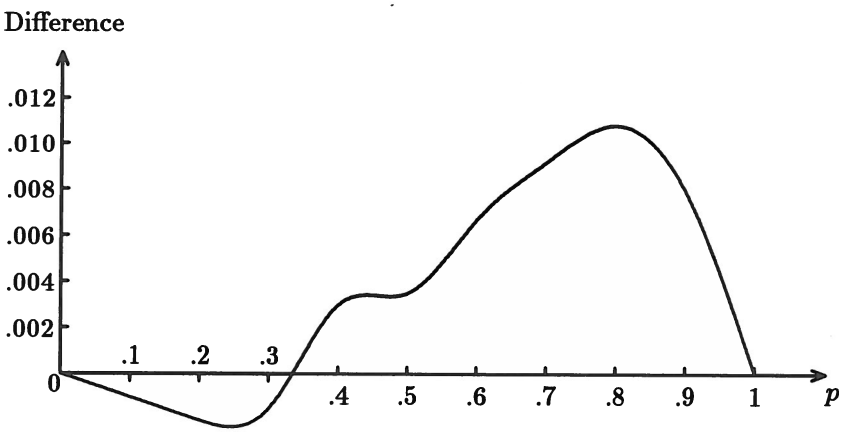

Figure 1

Difference between the Lorenz Curves, 1984 minus 1990

It can be seen that, between 1984 and 1990, Lorenz curve decile ordinates rose significantly in deciles 1 and 2: the $t-$statistics for the hypotheses that the decile ordinates do not differ are 3.31 and 2.89, with corresponding (asymptotic normal) $P-$values of 0.001 and 0.002 for two-tailed tests, or 0.0005 and 0.001 for one-tailed tests. On the other hand, the ordinates fell in deciles 5–9, with two-tailed $P-$values ranging from 0.02 to 0 (for a $t-$statistic of 6.06). The crossover seen in the Figure thus seems to be statistically significant.

Bishop, Formby, and Thistle (1992) suggested a union-intersection test of the hypothesis that one set of Lorenz curve decile ordinates dominates another. This test, based on the work of Beach and Richmond (1985), examines the set of $t-$statistics for the hypotheses that the individual decile ordinates do not differ. Assuming that one may reject the joint hypothesis that the full set of decile ordinates are the same, Bishop, Formby, and Thistle propose that one accept the hypothesis that one set of ordinates dominates the other, against the alternative of non-comparability, if at least one of the $t-$statistics has the appropriate sign and is significant, and none of the $t-$statistics (if any) that has the wrong sign is significant. Significance is determined asymptotically by the critical values of the Studentized Maximum Modulus (SMM) distribution with 9 (for deciles) and

Table 1

Male Paid Workers in the United States, 1984–1990

	Lorenz Curve Ordinates			Coefficient of Variation		
Decile	1984	1990	1984–90	1984	1990	1984–90
1	0.0062	0.0072	-0.0010	0.9553	0.9290	0.0263
	0.0002	0.0002	0.0003	0.0161	0.0164	0.0229
2	0.0348	0.0368	-0.0020	0.7315	0.6994	0.0321
	0.0005	0.0005	0.0007	0.0075	0.0076	0.0106
3	0.0826	0.0842	-0.0015	0.6518	0.6222	0.0295
	0.0007	0.0007	0.0010	0.0054	0.0052	0.0075
4	0.1477	0.1474	0.0003	0.6129	0.5857	0.0272
	0,0009	0.0009	0.0013	0.0043	0.0042	0.0060
5	0.2295	0.2260	0.0035	0.5916	0.5665	0.0252
	0.0011	0.0012	0.0015	0.0037	0.0036	0.0051
6	0.3282	0.3215	0.0067	0.5796	0.5603	0.0193
	0.0012	0.0012	0.0017	0.0032	0.0031	0.0045
7	0.4452	0.4360	0.0092	0.5759	0.5633	0.0125
	0.0013	0.0012	0.0018	0.0029	0.0028	0.0045
8	0.5831	0.5723	0.0108	0.5793	0.5734	0.0060
	0.0013	0.0013	0.0018	0.0027	0.0026	0.0037
9	0.7486	0.7407	0.0079	0.5946	0.6002	-0.0056
	0.0012	0.0011	0.0016	0.0025	0.0025	0.0036
10	1.0000	1.0000	0.0000	0.7259	0.7405	-0.0146
	0.0000	0.0000	0.0000	0.0040	0.0037	0.0055

Note: Figures reported below the point estimates are (asymptotic) standard errors.

---

an infinite number of degrees of freedom. The critical value for a 5% test is 2.80, and for a 1% test 3.29. For the two sets of ordinates that we consider here, for 1984 and 1990, one may reject the hypothesis of dominance in favor of non-comparability at both the 1% and the 5% level.

Between 1984 and 1990 the conditional coefficients of variation significantly decreased in deciles 2–7, with t–statistics for the individual

differences varying between 3.02 and 4.93, the corresponding two-tailed $P-$values being 0.0025 and 0.000001. For decile 10, however, the coefficient of variation increased, with a $t-$statistic of 2.67 for the difference. Using the union-intersection test once more, we observe that each of the $t-$statistics for deciles 1–8 is of the same sign, and, for deciles 3–6, exceeds the 1% critical value of 3.29. For deciles 9 and 10, the $t-$statistics are not of the same sign as for deciles 1–8, but neither exceeds the critical value of 3.29 (or the 5% value of 2.80 for that matter). Thus, on the basis of transfer sensitivity, one can conclude that the earnings distribution for 1984 inequality dominates the earnings distribution for 1990: there was more inequality in 1984 than in 1990.

4. Conclusion

This paper has extended the techniques of statistical inference to the cumulative coefficients of variation of an income distribution. It thus provides the statistical basis for testing inequality dominance when Lorenz curves cross. We give the full (asymptotic) variance-covariance structure of the cumulative sample means and variances (jointly), and hence also of the Lorenz curve ordinates and the conditional coefficients of variation. The results are distribution-free and easy to compute.

The procedures are applied to the 1984 and 1990 earnings distributions of male paid workers in the United States. While inequality in the two distributions cannot be compared on the sole basis of Lorenz curve ordinates, since the curves for the two years cross, we found that, on the basis of transfer sensitivity, the 1990 earnings distribution was significantly less unequal than the 1984 distribution.

Appendix

In order to prove Theorem 1, we will demonstrate a more general result. Let a random variable Y be characterized by a strictly increasing continuously differentiable cumulative distribution function F, with inverse function G. Then let γ_p denote the expectation of some function h of Y, conditional on Y being in the low $p-$quantile of its distribution:

$$p\gamma_p = \int_0^{G(p)} h(y)\, dF(y). \tag{9}$$

If there is a set of independent drawings Y_i, $i = 1, \ldots, N$ from the distribution F, then we may estimate γ_p by $\hat{\gamma}_p$, defined as follows:

$$p\hat{\gamma}_p = \int_0^{\hat{G}(p)} h(y)\, d\hat{F}(y), \tag{10}$$

where the empirical distribution function $\hat{F}$ is defined as

$$\hat{F}(y) = N^{-1} \sum_{i=0}^{N} I_{[0,y]}(Y_i).$$

Here the indicator function satisfies

$$I_{[0,y]}(Y) = \begin{cases} 1 & \text{if } Y \in [0, y]; \\ 0 & \text{otherwise.} \end{cases}$$

Clearly $\hat{F}(y)$ is the fraction of the drawings Y_i which are smaller than y. $\hat{G}$ is the function that gives estimated quantiles:

$$\hat{G}(p) = Y_{([Np])}, \tag{11}$$

where $Y_{(i)}$ denotes the i^{th} order statistic, and $[Np]$ denotes the largest integer not greater than Np. $\hat{G}$ and $\hat{F}$ are inverse functions in the sense that

$$\hat{F}(\hat{G}(p)) = \frac{[Np]}{N}, \text{ and } \hat{G}(\hat{F}(y)) = \max_i \{Y_i \mid Y_i \leq y\}. \tag{12}$$

With these definitions the estimator (10) becomes

$$p\hat{\gamma}_p = N^{-1} \sum_{i=1}^{[Np]} h(Y_{(i)}), \tag{13}$$

a very easy expression to calculate from ordered sample data.

The disadvantage of both (10) and (13) for determining the asymptotic properties of $p\hat{\gamma}_p$ is that not only are the summands or the integrand random, but also the largest value of y or $Y_{(i)}$ included in the integral or sum respectively. Further, order statistics are correlated, and so laws of large numbers and central limit theorems based on series of i.i.d. variables are not applicable.

Fortunately, a simple trick allows us to overcome both these problems. Consider (10) with a nonrandom upper limit:

$$\int_0^{G(p)} h(y)\, d\hat{F}(y). \tag{14}$$

In order to make (12) take on a simple form, suppose that $Np = [Np]$, so that $p = r/N$ for some positive integer r. In contrast to (13), (14) can be written in terms of i.i.d. variables, as follows:

$$N^{-1} \sum_{i=0}^{N} h(Y_i)\, I_{[0,G(p)]}(Y_i), \tag{15}$$

where the indicator function ensures that only terms for which $Y_i \leq G(p)$ are counted in the sum. Whereas in (13) there are always exactly r terms, the number of nonzero terms in (15) is random. By the law of large numbers, (15) tends almost surely to

$$p\gamma_p \equiv E\big(h(Y)\, I_{[0,G(p)]}(Y)\big) = \int_0^{G(p)} h(y)\, dF(y) \tag{16}$$

Asymptotic normality of (15) can be proved similarly by use of the central limit theorem.

Next consider the difference between (10) and (14), which is

$$\int_{G(p)}^{\hat{G}(p)} h(y)\, d\hat{F}(y). \tag{17}$$

The estimated quantiles $\hat{G}(p)$ are root$-n$ consistent – see Wilks (1962), p. 273. Thus the estimation error $\hat{G}(p) - G(p)$ is $O(N^{-1/2})$. So too therefore is expression (17), since it is the integral of a finite function over an interval of length $O(N^{-1/2})$. Further, for $y \in [G(p), \hat{G}(p)]$,

$$h(y) = h(G(p)) + O(N^{-1/2}),$$

under weak smoothness conditions on h. Thus (17), which is itself $O(N^{-1/2})$, can be approximated, with an error of order only $O(N^{-1})$, by the following expression:

$$\begin{aligned}\int_{G(p)}^{\hat{G}(p)} h(G(p))\, d\hat{F}(y) &= h(G(p))\big(\hat{F}(\hat{G}(p)) - \hat{F}(G(p))\big) \\ &= -h(G(p))\big(\hat{F}(G(p)) - p\big) \quad \text{(by (12))} \\ &= ph(G(p)) - h(G(p)) \int_0^{G(p)} d\hat{F}(y).\end{aligned}$$

Adding this last expression to (14) yields the following expression for (10), valid with error of at most $O(N^{-1})$:

$$\begin{aligned}p\hat{\gamma}_p &= ph(G(p)) + \int_0^{G(p)} \big(h(y) - h(G(p))\big)\, d\hat{F}(y) \\ &= ph(G(p)) + N^{-1} \sum_{i=1}^{N} \big(h(Y_i) - h(G(p))\big)\, I_{[0,G(p)]}(Y_i). \qquad (18)\end{aligned}$$

Clearly, to order N^{-1},

$$\begin{aligned}E(p\hat{\gamma}_p) &= ph(G(p)) + E\big((h(y) - h(G(p)))\, I_{[0,G(p)]}(y)\big) \\ &= ph(G(p)) + p\gamma_p - ph(G(p)) = p\gamma_p,\end{aligned}$$

where the second equality follows from (16) and the fact that

$$E\big(I_{[0,G(p)]}(y)\big) = \Pr(y \le G(p)) = p.$$

Thus $p\hat{\gamma}_p$ is an asymptotically unbiased estimator of $p\gamma_p$.

The next task is to compute the covariance structure of estimators like $p\hat{\gamma}_p$. By analogy with (9) and (10), let

$$p\delta_p \equiv \int_0^{G(p)} k(y)\, dF(y) \quad \text{and} \quad p\hat{\delta}_p \equiv \int_0^{\hat{G}(p)} k(y)\, d\hat{F}(y), \qquad (19)$$

for some function k with the same properties as h. Then, asymptotically, the covariance of $p\hat{\gamma}_p$ and $p'\hat{\delta}_{p'}$ is, from (18),

$$\begin{aligned}N^{-1}\Big(&E\Big(\big((h(y) - h(G(p)))\, I_{[0,G(p)]}(y)\big)\big((k(y) - k(G(p')))\, I_{[0,G(p')]}(y)\big)\Big) \\ &-E\big((h(y) - h(G(p)))\, I_{[0,G(p)]}(y)\big)E\big((k(y) - k(G(p')))\, I_{[0,G(p')]}(y)\big)\Big). \qquad (20)\end{aligned}$$

Now

$$E\big((h(y) - h(G(p)))\ I_{[0,G(p)]}(y)\big) = p(\gamma_p - h(G(p))) \tag{21}$$

by the definition of γ_p, and, similarly,

$$E\big((k(y) - k(G(p')))\ I_{[0,G(p')]}(y)\big) = p'(\delta_{p'} - k(G(p'))). \tag{22}$$

Suppose without loss of generality that $p \le p'$. Then

$$\begin{aligned} E\Big(\big((h(y) - h(G(p)))\ I_{[0,G(p)]}(y)\big)\big((k(y) - k(G(p')))\ I_{[0,G(p')]}(y)\big)\Big) \\ = E\Big(h(y)\ k(y)\ I_{[0,G(p)]}(y)\Big) - p\gamma_p k(G(p')) \\ -p\delta_p h(G(p)) + ph(G(p))k(G(p')) \end{aligned} \tag{23}$$

For ease of notation, let

$$\begin{aligned} p\phi_p \equiv pE\big(h(y)\ k(y) \mid y \le G(p)\big) &= E\big(h(y)\ k(y)\ I_{[0,G(p)]}(y)\big) \\ &= \int_0^{G(p)} h(y)\ k(y)\ dF(y). \end{aligned}$$

Then substituting (21), (22), and (23) into (20) gives N times the covariance of $p\hat{\gamma}_p$ and $p'\hat{\delta}_{p'}$ as

$$\begin{aligned} p\Big\{ \phi_p - \gamma_p k(G(p')) - \delta_p h(G(p)) + h(G(p))\ k(G(p')) \\ - p'\big(\gamma_p - h(G(p))\big)\big(\delta_{p'} - k(G(p'))\big)\Big\}. \end{aligned}$$

This can be rearranged to yield a somewhat more convenient expression:

$$\begin{aligned} p\Big\{ \phi_p - \gamma_p\delta_p + (1-p')\big(h(G(p)) - \gamma_p\big)\big(k(G(p')) - \delta_{p'}\big) \\ + \big(h(G(p)) - \gamma_p\big)\big(\delta_{p'} - \delta_p\big)\Big\}. \end{aligned} \tag{24}$$

The equations (4), (5), (6), and (7) of Theorem 1 are readily seen to be special cases of (24), with $h(y) = y$ and $k(y) = y^2$.

Everything in (24) can be estimated consistently in a distribution-free manner: γ_p, δ_p, and $\delta_{p'}$ by $\hat{\gamma}_p$, $\hat{\delta}_p$, and $\hat{\delta}_{p'}$ respectively ((10) and (19)); $G(p)$ and $G(p')$ by $\hat{G}(p)$ and $\hat{G}(p')$, that is, the sample p and p' quantiles ((11)); and ϕ_p by $\hat{\phi}_p$, with

$$p\hat{\phi}_p \equiv \int_0^{\hat{G}(p)} h(y)\ k(y)\ d\hat{F}(y) = N^{-1}\sum_{i=1}^{[Np]} h(Y_{(i)})k(Y_{(i)}).$$

References

Anderson, G. (1994). "Nonparametric Tests of Stochastic Dominance in Income Distributions." Unpublished paper, University of Toronto.

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

Beach, C., V. Chow, J. Formby, and G. Slotsve (1994). "Statistical Inference for Decile Means." *Economics Letters*, **45**, 161–167.

Beach, C. and R. Davidson (1983). "Distribution-Free Statistical Inference with Lorenz Curves and Income Shares." *Review of Economic Studies*, **50(3)**, 723–735.

Beach, C. M., and J. Richmond (1985). "Joint Confidence Intervals for Income Shares and Lorenz Curves." *International Economic Review*, **26**, 439–450.

Beach, C. M. and G. A. Slotsve (1994). *Lorenz Program*, Department of Economics, Queen's University, Kingston, Ontario.

Bishop, J., J. Formby, and J. Smith (1991). "Lorenz Dominance and Welfare: Changes in the U.S. Distribution of Income, 1967–1986." *Review of Economics and Statistics*, **73**, 134–141.

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

Blackburn, M. and D. Bloom (1993). "The Distribution of Family Income: Measuring and Explaining Changes in the 1980s for Canada and the United States," in D. Card and R. Freeman (eds), *Small Differences that Matter: Labor Markets and Income Maintenance in Canada and the United States*, the University of Chicago Press, Chicago, pp. 233–265.

Bound, J. and G. Johnson (1992). "Changes in the Structure of Wages in the 1980s: an Evaluation of Alternative Explanations." *American Economic Review*, **82**, 371–392.

Davies, J. and M. Hoy (1994). "Making Inequality Comparisons when Lorenz Curves Intersect." *American Economic Review*, forthcoming.

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

Howes, S. (1994). "Testing for Dominance: Inferring Population Rankings from Sample Data." Unpublished paper, Policy Research Department, World Bank.

Jenkins, S. (1991). "The Measurement of Income Inequality," in L. Osberg (ed.), *Economic Inequality and Poverty: International Perspectives*, M. E. Sharpe Inc., Armonk, N.Y., pp. 3–38.

Karoly, L. (1992). "Changes in the Distribution of Individual Earnings in the United States: 1967–1986." *The Review of Economics and Statistics*, **74**, 107–115.

Lambert, P. (1989). *The Distribution and Redistribution of Income: a Mathematical Analysis*, Basil Blackwell, Cambridge, Mass.

Levy, F. and R. Murnane (1992). "U.S. Earnings Levels and Earnings Inequality: a Review of Recent Trends and Proposed Explanations." *Journal of Economic Literature*, **30**, 1333–1381.

Lillard, L., J. P. Smith, and F. Welch (1986). "What do we Really Know about Wages? The Importance of Nonreporting and Census Imputation." *Journal of Political Economy*, **94**, 489–506.

Murphy, K. and F. Welch (1992). "The Structure of Wages," *Quarterly Journal of Economics*, **107**, 285–326.

Rao, C. R. (1965). *Linear Statistical Inference and its Applications*, John Wiley, New York.

Richardson, D. (1994). "Changes in the Distribution of Wages in Canada, 1981–1992." University of British Columbia Discussion Paper no. 94–22.

Shorrocks, A. and J. Foster (1987). "Transfer Sensitive Inequality Measures." *Review of Economic Studies*, **54**, 485–497.

Wilks, S. S. (1962). *Mathematical Statistics*, John Wiley, New York.

Xu, K. (1994). "Dominance Testing in Economics and Finance." Unpublished Ph.D. dissertation, Department of Economics, Concordia University, Montreal, Canada.