Distributional Dominance with Trimmed Data

Cowell, Frank; Victoria-Feser, Maria-Pia

How to cite

COWELL, Frank, VICTORIA-FESER, Maria-Pia. Distributional Dominance with Trimmed Data. In: Journal of business & economic statistics, 2006, vol. 24, p. 291–300. doi: 10.1198/073500105000000207

This publication URL: https://archive-ouverte.unige.ch/unige:6461

Publication DOI: 10.1198/073500105000000207

Distributional Dominance With Trimmed Data

Frank A. COWELL
London School of Economics and Political Science, Houghton Street, London WC2A 2AE, U.K.
(*f.cowell@lse.ac.uk*)

Maria-Pia VICTORIA-FESER
HEC, University of Geneva, 40 Bd du Pont d'Arve, CH-1211, Geneva, Switzerland
(*maria-pia.victoriafeser@hec.unige.ch*)

Distributional dominance criteria are commonly applied to draw welfare inferences about comparisons, but conclusions drawn from empirical implementations of dominance criteria may be influenced by data contamination. We examine a nonparametric approach to refining Lorenz-type comparisons and apply the technique to two important examples from the Luxembourg Income Study database.

KEY WORDS: Distributional dominance; Lorenz curve; Robustness.

1. INTRODUCTION

This article addresses the issue of how practical comparisons of income distributions can be founded on a sound statistical and economic base when there is good reason to believe that the data in at least one of the distributions are "dirty." Dirt includes the possibility of obvious gross errors in the data (such as arise from coding or transcribing mistakes) and also other more innocuous observations that in some sense do not really belong to the income dataset. The problem is often handled pragmatically; some empirical studies have concentrated on a subset of the distribution delimited either by population subgroup (e.g., taking-prime-age males only) or by arbitrarily excluding some of the data in the tails (Gottschalk and Smeeding 2000). Although this research technique seems sensible, the question of whether it is appropriate remains open—"appropriateness" here being understood in terms of the statistical properties of the underlying economic criteria. This question matters because the economic criteria are used explicitly or implicitly to make normative judgments and perhaps policy recommendations.

Cowell and Victoria-Feser (2002) set the conditions under which welfare judgments are valid with model contamination (which includes the dirty data case) and concluded that second-order rankings, such as Lorenz curves, are highly sensitive to data contamination in the tail of the distributions. Starting from this result, we propose a formalization of the practical but ad hoc procedure of trimming with a family of ranking statistics.

In this article we develop the relationship between economic ranking principles and statistical tools to derive a practical method for making distributional comparisons in the presence of data contamination. This method uses a family of dominance comparisons based on the statistical concept of the trimmed mean. The basic methodology is set out in Section 2. Considerations of data contamination and their likely impact on the estimates of statistics associated with distributional dominance are discussed in Section 3. The application of these methods is demonstrated in Section 4 using second-order dominance and related Lorenz comparisons over time and between countries.

2. DISTRIBUTIONAL DOMINANCE WITH "DIRTY DATA"

2.1 Informal Methods

Empirical studies of income distribution use informal ranking criteria as a matter of routine. There are various good reasons for doing so. They usually involve easy computations, and they have a direct intuitive appeal; more important, they are usually connected to deeper points that are particularly relevant to applied welfare economists. Some prominent examples of the informal approach are as follows:

- Pragmatic indices involving *quantiles*. These include the semidecile ratio (Wiles 1974; Wiles and Markowski 1971) and the comparative function of Esberger and Malmquist (1972). An extreme example of the same type is the *range*, literally the maximum income minus the minimum income, but sometimes implemented in practice as a difference between extreme quantiles.
- The "*parade of incomes*" introduced by Pen (1971). This provides a persuasive picture of snapshot inequality and of the implications of an income distribution that is changing through time (see, e.g., Jenkins and Cowell 1994).
- The use of *distributive shares* (sometimes known as quantile shares).

The quantile method can be explicitly linked to formal welfare criteria. For example, in Rawls' work on a theory of justice there is a discussion of how to implement his famous "difference principle," which focuses on the least advantaged. To do this, Rawls himself suggested that it might be interpreted relative to the median of the distribution (see Rawls 1972, p. 98). So too can the distributive shares approach; changes in the relative income shares of, say, the richest and the poorest 10% slices of the distribution can be directly interpreted in terms of the principle of transfers (Dalton 1920).

2.2 A Formal Framework

Assume that the concepts of income and income receiver have been well defined. An individual's income is a number $x \in \mathfrak{X}$, where $\mathfrak{X} \subseteq \mathbb{R}$ and $\mathbb{R}$ is the real line. Let $\mathfrak{F}$ be the set of probability distributions (distribution functions) with support $\mathfrak{X}$. An *income distribution* is one particular member $F \in \mathfrak{F}$.

Journal of Business & Economic Statistics
???? 0, Vol. 0, No. 00
DOI 10.1198/00

In this approach a *statistic* of any distribution $F \in \mathfrak{F}$ is a functional $T(F)$; for example, the mean $\mu : \mathfrak{F} \mapsto \mathbb{R}$ given by

$$\mu(F) := \int x\, dF(x). \tag{1}$$

The properties of any functional T may play a role in both economic and statistical interpretations. Of particular interest here is the case where the range of T is a profile of values rather than a single number as in the example of (1); T is then a *family* of statistics. Individual family members may be of interest in their own right; the behavior of the whole family when applied to a pair of distributions F and G will provide important information about distributional comparisons that is richer than that provided by a single real-valued functional.

The basic distributional concept used here is a *ranking*, which amounts to a partial ordering on the space of distributions $\mathfrak{F}$. We use the symbol $\succeq_T$ to denote the ranking induced on $\mathfrak{F}$ by a statistic T, from which a number of the concepts of strict dominance, equivalence, and noncomparability are derived (see, e.g., Cowell and Victoria-Feser 2002). For example, T can be the quantile function

$$Q(F; q) = \inf\{x | F(x) \geq q\} = x_q \tag{2}$$

with $0 \leq q \leq 1$, which defines first-order dominance. The first-order dominance criterion $\succeq_Q$ is sometimes considered less than ideal, and so it is of interest to consider the second-order criterion (Cowell 2000). This is based on the *cumulative income functional*,

$$C(F; q) := \int_{\underline{x}}^{Q(F;q)} x\, dF(x), \tag{3}$$

with $0 \leq q \leq 1$ and $\underline{x} := \inf \mathfrak{X}$. By definition, $C(F; 0) = 0$ and $C(F; 1) = \mu(F)$. For a given $F \in \mathfrak{F}$, the graph of $C(F, q)$ against q describes the *generalized Lorenz curve* (GLC). The *relative Lorenz curve* (Lorenz 1905) is obtained by standardizing $C(F, q)$, the mean, that is,

$$L(F; q) := \frac{C(F; q)}{\mu(F)}, \tag{4}$$

and the *absolute Lorenz curve* (Moyes 1987) by

$$A(F; q) := C(F; q) - q\mu(F). \tag{5}$$

The graph of $L(F; q)$ against q is closely related to the first moment function, that is, the function $\Phi : \mathfrak{X} \mapsto [0, 1]$ defined for any $F \in \mathfrak{F}$ as $\Phi(x) = L(F; F(x)) = \frac{1}{\mu(x)} \int^x y\, dF(y)$ (Kendall and Stuart 1977). The (relative) Lorenz curve encapsulates the intuitive principle of the distributional-shares ranking referred to in Section 2.1. We examine the implementation of (3), (4), and (5) in Section 4.

2.3 The Approach

To assume that data will automatically give a reasonable picture of the "true" picture of a distributional comparison would obviously be reckless in the extreme. A prudent applied researcher will anticipate that because of miscoding and misreporting and other types of mistakes, some of the observations will be incorrect, and this may have a serious impact on distributional comparisons (Van Praag, Hagenaars, and Van Eck 1983). Obviously, if one had reason to suspect that this sort of error were extensive in the datasets under consideration, then the problem of distributional comparison might have to be abandoned because of unreliability. But it is possible that there might be a serious problem of comparison even if the amount of contamination were small, so that the data might be considered "reasonably clean."

Let us briefly review a standard model of this type of problem. This approach is based on the work of Hampel (1968, 1974), Hampel, Ronchetti, Rousseeuw, and Stahel (1986), and Huber (1981). Suppose that the "true" distributions that we wish to compare are denoted by F and G, but because of the problem of data-contamination we cannot assume that the data we have at hand have really been generated by F and G. What we actually observe instead of F is a distribution in some neighborhood of F. An elementary case is one in which a mixture distribution has been constructed by combining the "true" distribution F with a point mass at income z

$$F_\epsilon^{(z)} = [1 - \epsilon]F + \epsilon H^{(z)}, \tag{6}$$

where

$$H^{(z)}(x) = \begin{cases} 1 & \text{if } x \geq z \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

The degenerate distribution $H^{(z)}$ represents a simple form of *data contamination* at point z; ϵ indicates the importance of the contamination; the convex combination $F_\epsilon^{(z)}$ is the observed distribution, and F remains unobservable.

As we have noted, if ϵ is large, then we cannot expect to get sensible estimates of income distribution statistics. But what if the contamination were very small? To address this question for any given statistic T, we use the *influence function* (IF), given by

$$IF(z; T, F) := \lim_{\epsilon \to 0} \left[\frac{T(F_\epsilon^{(z)}) - T(F)}{\epsilon} \right]. \tag{8}$$

Then, under the given model of data contamination (6), the statistic T is *robust* if the IF in (8) is bounded for all $z \in \mathfrak{X}$.

A more reasonable formalization is when $H^{(z)}$ is generalized to any distribution H. In this case, all types of model deviations and measurement errors can be considered as being included in the mixture distribution. However, the formal study of the effect of such model deviations on the statistic T would then become impossible unless some structure were given to H. Moreover, Hampel et al. (1986) showed that

$$\sup_H \left\| T\big((1 - \varepsilon)F + \varepsilon H\big) - T(F) \right\| \simeq \varepsilon \sup_x \|IF(x; T, F)\|,$$

which means that the reduction of H to $H^{(z)}$ does not reduce the information on the bias of the statistic T under model deviation, such as measurement errors. In other words, the IF gives information on the behavior of T for any "contamination" distribution H.

Cowell and Victoria-Feser (1996, 2002) showed that most inequality measures are nonrobust, but most poverty indices with exogenous poverty lines are robust (see also Monti 1991). However, the nonrobustness problem is more pervasive than that which emerges in connection with inequality measures; the same type of approach can be used to show that although first-order dominance criteria are usually robust (for further discussion of the statistical implementation of first-order criteria, see Ben Horim 1990; Stein, Pfaffenberger, and French 1987), second- and higher-order dominance criteria (and associated ranking tools) are not (Cowell and Victoria-Feser 2002). Moreover, Cowell and Victoria-Feser (2002) showed that the worst bias on the second-order ranking statistics (as provided by the IF) occurs when the contamination is in the tails of the distribution. A case can then be made for adopting an approach based on trimmed distributions. This approach certainly will not prevent biases for any type of model deviation or measurement error, but will guarantee that the largest potential biases are kept under control.

3. ROBUST DISTRIBUTIONAL DOMINANCE

3.1 Trimmed Ranking Criteria

Because ranking criteria can be misleading in the presence of data contamination, it is desirable to have a procedure that enables one to control systematically for suspect values that may distort distributional comparisons using second-order ranking criteria. A natural approach would be to use an established tool in the statistical literature, the "trimmed mean," and extend the idea to Lorenz curve analysis. The trimmed mean of distribution F with trimming parameter α is

$$\bar{X}_\alpha(F) = \frac{1}{1-2\alpha}\int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} y\,dF(y)$$

$$= \frac{1}{1-2\alpha}\int_{\alpha}^{1-\alpha} F^{-1}(t)\,dt,$$

where $\alpha \in [0, \frac{1}{2})$ is the *balanced trimming proportion*. This estimator of location has intuitive appeal; one removes the αn smallest and the αn largest observations in a sample of size n and calculates the mean of the remaining observations. Note that $\lim_{\alpha\to .5} \bar{X}_\alpha(F) = Q(F, .5)$; in the limiting case, as α approaches 50%, the trimmed estimate of the mean approaches the median.

Likewise, consider *trimmed Lorenz curves* as estimators of Lorenz curves. One must interpret the quantile and income-cumulation functions (2) and (3). α-trimming the data means that $Q(F; q) \in (Q(F; \alpha), Q(F; 1-\alpha))$ and thus $q \in (\alpha, 1-\alpha)$. However, it makes sense to consider a more general trimming method that includes in particular the single-tailed trimming case. Indeed, this case is appropriate when one can form an a priori judgment about the nature of the contamination, for example, when contamination is assumed to affect only the lower tail of the distribution. Let $\underline{\alpha}$ and $1-\bar{\alpha}$ be the lower and upper trimming, and let $\alpha := \underline{\alpha} + (1-\bar{\alpha})$ be the total trimmed amount. Then the α-trimmed generalized Lorenz, Lorenz, and absolute Lorenz curves (see the similar concept of restricted dominance, discussed in Atkinson and Bourguignon 1989) for $q \in (\underline{\alpha}, 1-\bar{\alpha})$ are given by

$$c_{\alpha,q} := C_\alpha(F; q) = \frac{1}{1-\alpha}\int_{Q(F;\underline{\alpha})}^{Q(F;q)} u\,dF(u), \tag{9}$$

$$l_{\alpha,q} := L_\alpha(F; q) = \frac{C_\alpha(F; q)}{C_\alpha(F; 1-\bar{\alpha})}, \tag{10}$$

and

$$A_\alpha(F; q) = (1-\bar{\alpha}-\underline{\alpha})\cdot C_\alpha(F; q) - C_\alpha(F; 1-\bar{\alpha})\cdot(q-\underline{\alpha}) \tag{11}$$

[cf. (3), (4), and (5)]. From (9)–(11), we have that $C_\alpha(F; \underline{\alpha}) = 0$, $L_\alpha(F; \underline{\alpha}) = 0$, $A_\alpha(F; \underline{\alpha}) = 0$, and $L_\alpha(F; 1-\bar{\alpha}) = 1$, $A_\alpha(F; 1-\bar{\alpha}) = 0$.

The IF's of these trimmed Lorenz curves will be bounded for all q because extreme values in the data are automatically removed for all $\underline{\alpha}, 1-\bar{\alpha} > 0$. Trimmed Lorenz curves can be thought of as Lorenz curves on a restricted sample in which $100\underline{\alpha}$% of the bottom observations and $100(1-\bar{\alpha})$% of the top observations have been trimmed away. This practice is sometimes adopted in pragmatic discussions of inequality trends. (See also the discussion of related issues by Howes 1996.) Estimates can be obtained by replacing F with the empirical distribution $F^{(n)}(x) = \frac{1}{n}\sum_{i=1}^{n} H^{(x_i)}(x)$.

3.2 Confidence Intervals

When comparing distributions using ranking criteria, it is also important to be able to provide confidence intervals for the latter. Cowell and Victoria-Feser (2003) gave formulas for several statistics, including ranking criteria for full and trimmed samples. In particular, we have that the asymptotic covariance of $\sqrt{n}C_\alpha(F^{(n)}; q)$ and $\sqrt{n}C_\alpha(F^{(n)}; q')$ with $q \le q'$ is given by $\omega_{qq'}/(1-\alpha)^2$, where

$$\begin{aligned}\omega_{qq'} := {}& \big[qQ(F; q) - \underline{\alpha}Q(F; \underline{\alpha}) - [1-\alpha]c_{\alpha,q}\big] \\ & \times \big[[1-q']Q(F; q') - [1-\underline{\alpha}]Q(F; \underline{\alpha}) + [1-\alpha]c_{\alpha,q'}\big] \\ & - \big[Q(F; q)[1-\alpha]c_{\alpha,q} - [1-\alpha]s_{\alpha,q}\big] \\ & + Q(F; \underline{\alpha})\big[(q-\underline{\alpha})Q(F; q) - [1-\alpha]c_{\alpha,q}\big],\end{aligned} \tag{12}$$

with $s_{\alpha,q} := S(F; q) = \frac{1}{1-\alpha}\int_{Q(F;\underline{\alpha})}^{Q(F;q)} u^2\,dF(u)$. For the Lorenz curve ordinates, the asymptotic variance is

$$\upsilon_{qq'} = \frac{1}{(1-\alpha)^2\mu_\alpha^4} \times [\mu_\alpha^2\omega_{qq'} + c_{\alpha,q}c_{\alpha,q'}\omega_{\bar{\alpha}\bar{\alpha}} - \mu_\alpha c_{\alpha,q}\omega_{q'\bar{\alpha}} - \mu_\alpha c_{\alpha,q'}\omega_{q\bar{\alpha}}],$$

with $\mu_\alpha = C_\alpha(F; 1-\bar{\alpha})$. These covariances can be estimated by their empirical counterpart (see Sec. A.1).

3.3 Choosing the Trimming Proportions

The sampling properties of the key distributional statistics can provide a simple choice criterion. Let $\tilde{F}_\alpha$ be the trimmed distribution given by

$$\tilde{F}_\alpha(x) := \begin{cases} 0 & \text{if } x < Q(F, \underline{\alpha}) \\ \dfrac{F(x)-\alpha}{1-\alpha} & \text{if } Q(F, \underline{\alpha}) \le x < Q(F, \bar{\alpha}) \\ 1 & \text{if } x \ge Q(F, \bar{\alpha}), \end{cases}$$

and let T be a statistic of interest. We consider the ratio of the mean squared errors,

$$\kappa(\alpha) := \frac{(T(F) - \widehat{T}(F))^2 + \operatorname{var}\widehat{T}(F)}{(T(F) - \widehat{T}(\tilde{F}_\alpha))^2 + \operatorname{var}\widehat{T}(\tilde{F}_\alpha)}. \tag{13}$$

Indeed, there is no guarantee that $\widehat{T}(\tilde{F}_\alpha)$ is a consistent estimator of $T(F)$, so that not only the variances, but also the potential bias of the statistics should be taken into account. (We are gratefulul to two anonymous referees for pointing this out.) The implied trade-off of robustness against efficiency enables the researcher to make an informed choice about the extent of trimming that may be reasonable in making distributional comparisons.

Now (13) clearly implies that this choice is conditional on the specification of T. Which statistic would be appropriate? It seems reasonable to require that this be one of second-order distributional dominance, but this raises a further difficulty: There is an uncountable infinity of statistics $C(\cdot; q)$, and selecting one or a few of these appears to be arbitrary. However, there is a simple argument suggesting that one particular case is especially important. Not all values of q in the unit interval will be relevant in computing efficiency under trimming; the very process of trimming "nibbles away" some of the interval. If one is interested in trimming of arbitrary size, then it seems to be of particular interest to examine cases where $T(\tilde{F}_\alpha)$ is well defined for arbitrary α. In the case of a balanced trim, this implies focusing attention on $C(\cdot; .5)$ or its relative Lorenz counterpart $C(\cdot; .5)/\mu(\cdot)$.

$\kappa(\alpha)$ also depends on the underlying income distribution F. For the purposes of illustrating the technique and to obtain an idea of the efficiency losses involved, we used a number of examples of the Dagum type I distribution given by

$$f(x; \beta, \lambda, \delta) = \beta\delta\lambda^{-\beta}x^{\beta\delta-1}(1+\lambda^{-1}x^{\delta})^{-(\beta+1)}. \qquad (14)$$

We have (see Kleiber and Kotz 2003) $Q(F; q) = \lambda^{1/\delta}(q^{-1/\beta} - 1)^{-1/\delta}$ and $C(F; 1) = \lambda^{1/\delta}\Gamma(\beta + 1/\delta)\Gamma(1 - 1/\delta)/\Gamma(\beta)$, which were use to compute the theoretical values of $T(F)$. Two examples are illustrated in Figure 1. From these two simulated datasets, we computed the sampling variances for the trimmed and untrimmed cases, with lower, upper, and balanced trims.

The results are illustrated in Figure 2, where the vertical axis gives estimated values of $\kappa(\alpha)$ as defined in (13). One can see that the estimated efficiency loss (or gain) depends on the underlying model and the type of trim. For small trimming quantities, it is not very large; for larger trimming quantities, it can be either quite large or reasonable. The efficiency loss is smaller with lower trimming, which for more symmetric distributions [like the Dagum(8, 4, 8)] can even be greater than 1! It is difficult to draw a general conclusion, however, and the results presented here can provide at most a rough guideline.

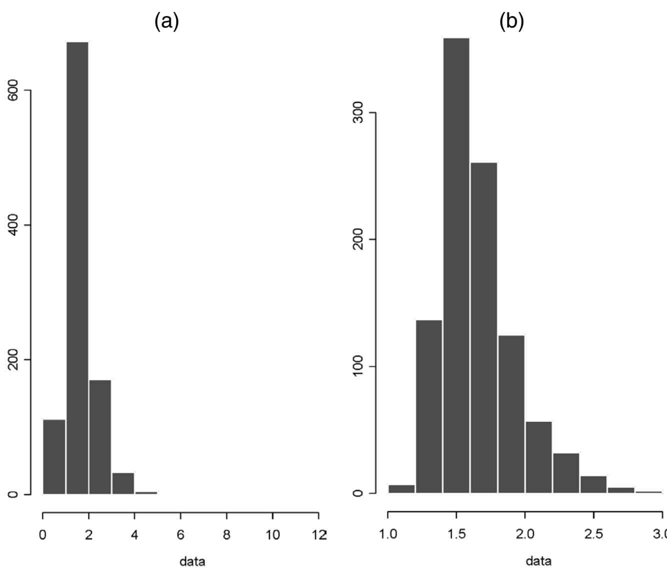

Figure 1. The Dagum Distribution: (a) (2, 2, 4); (b) (8, 4, 8).

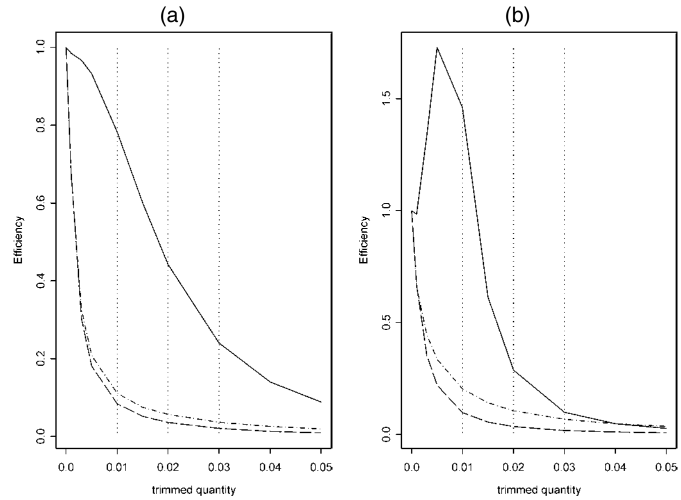

Figure 2. Efficiency Under Trimming for $C(\cdot, .5)/\mu(\cdot)$: (a) Dagum(2, 4, 2); (b) Dagum(4, 8, 4). (— lower trim; – – upper trim; -·-· balanced trim.)

4. EMPIRICAL APPLICATION

The trimming approach offers a practical tool for comparing income distribution when one wants an explicit control for taking into account the influence of outliers. We use the analysis of Section 3 to more carefully examine two aspects of conventional wisdom concerning comparisons of income distribution. In both cases the data are taken from the Luxembourg Income Study (LIS) database and refer to real income per equivalent adult distributed among individuals (see Sec. A.2).

4.1 Cross-Country Comparison: Sweden and Germany

The received wisdom suggests that 1980s Sweden was more equal than Germany. But is this actually borne out by the data, and if so, what are the implications for standard welfare comparisons? To investigate this, we use data for Sweden 1981 and (West) Germany 1983. As shown in Figure 3, we have $F_{\mathrm{GERMANY}} \succeq_C F_{\mathrm{SWEDEN}}$; the German income distribution second-order (generalized Lorenz) dominates that for Sweden. This conclusion is robust under trimming. But the picture is different if we attempt to apply the criterion of absolute Lorenz dominance. From the untrimmed data, we find $F_{\mathrm{SWEDEN}} \perp_A F_{\mathrm{GERMANY}}$ but under a very slight trimming of both tails, it is clear that Sweden absolute Lorenz dominates Germany $F_{\mathrm{SWEDEN}} \succeq_{A_{.005}} F_{\mathrm{GERMANY}}$ (compare Figs. 4 and 5). Currency units are 1981 U.S. dollars.

What of inequality? As Figure 6 shows, there is an ambiguity for the raw data $F_{\mathrm{SWEDEN}} \perp_L F_{\mathrm{GERMANY}}$, which is due to a single intersection of the relative Lorenz curves. Figure 7 depicts

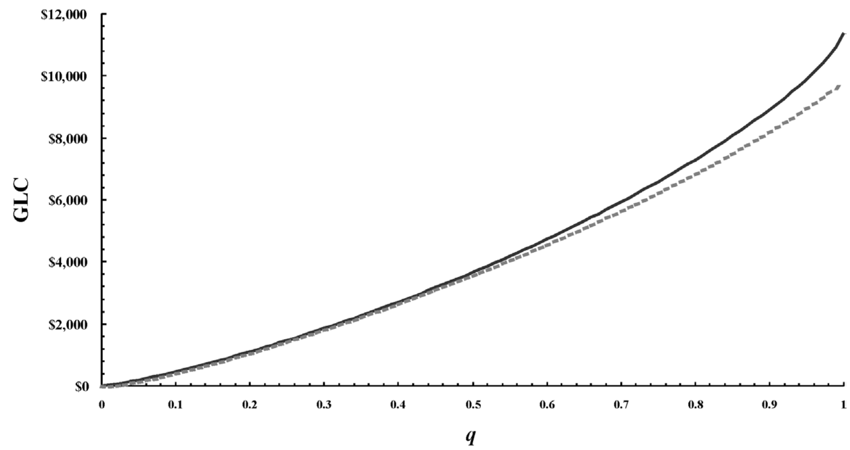

Figure 3. Generalized Lorenz Curves for Germany 1983 (——) and Sweden 1981 (- - -).

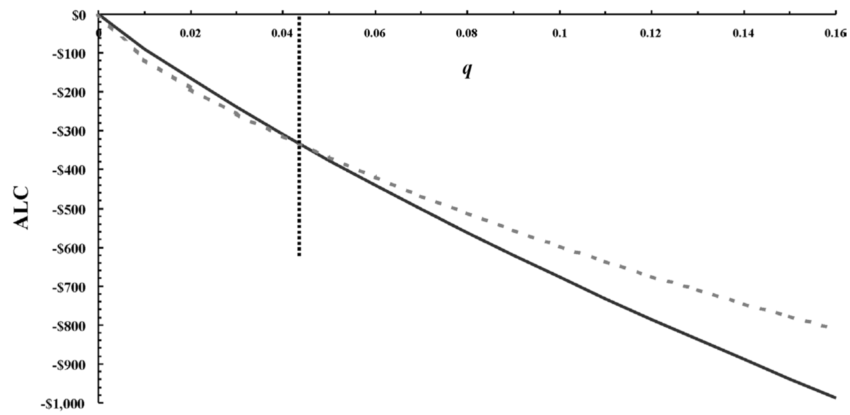

Figure 4. Absolute Lorenz Curves for Germany 1983 (——) and Sweden 1981 (- - -).

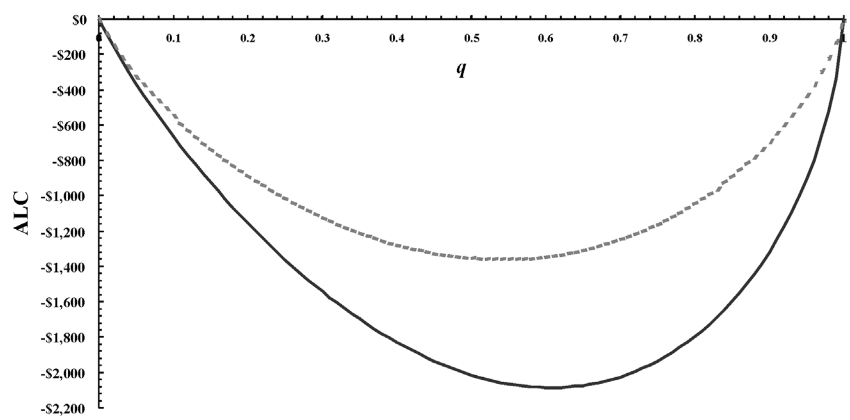

Figure 5. Absolute Lorenz Curves With .5% Balanced Trimming for Germany 1983 (——) and Sweden 1981 (- - -).

the *truncation profiles*, the position of the switchpoint (where the relative Lorenz curves intersect) for two types of trim, expressed as functions of $\alpha - q^{**}(\cdot)$ for the balanced two-tailed trim (solid curve) and $q^{*}(\cdot)$ for the one-sided lower-tailed trim (dotted curve). Denote the points where the truncation profiles intersect the horizontal axis by α^{**} and α^{*}. Then

$$q^{**}(0) = q^{*}(0) = .11,$$

$$q^{*}(\alpha) = 0, \qquad \alpha \geq \alpha^{*} = .030,$$

and

$$q^{**}(\alpha) = 0, \qquad \alpha \geq \alpha^{**} = .065.$$

We have $F_{\text{SWEDEN}} \succeq_{L_\alpha} F_{\text{GERMANY}}$ only if a trimming of 3% of the observations is carried out on the lower tail (see Fig. 8), or a balanced trimming of 6.5%. May we say that Sweden is less unequal than Germany? Consider two points here.

First, we apply the analysis of Section 3.2 to compute confidence intervals for the RLC of Germany 1983 and Sweden

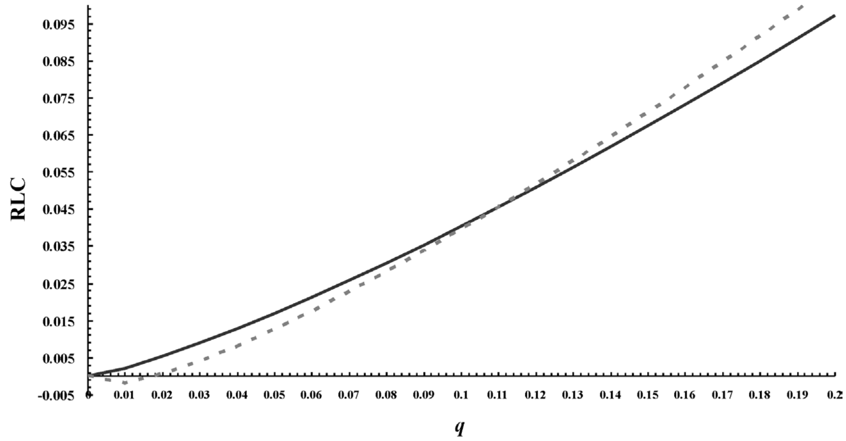

Figure 6. Relative Lorenz Curves for Germany 1983 (——) and Sweden 1981 (- - -).

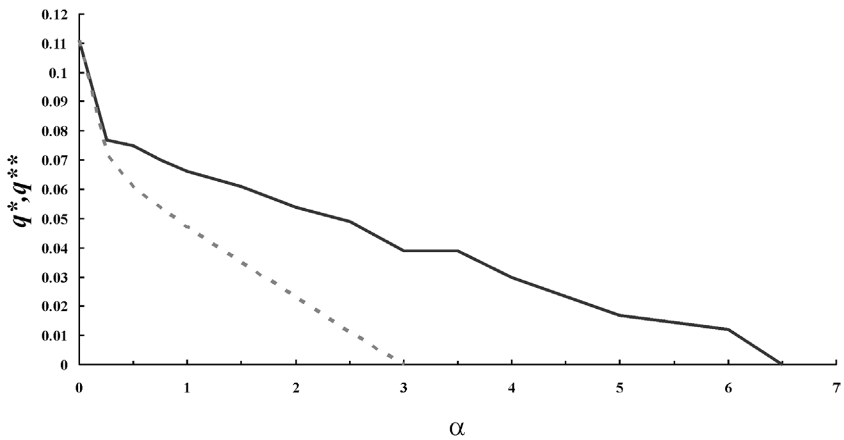

Figure 7. Is Sweden More Equal Than Germany? (——, two tail; - - -, one tail.)

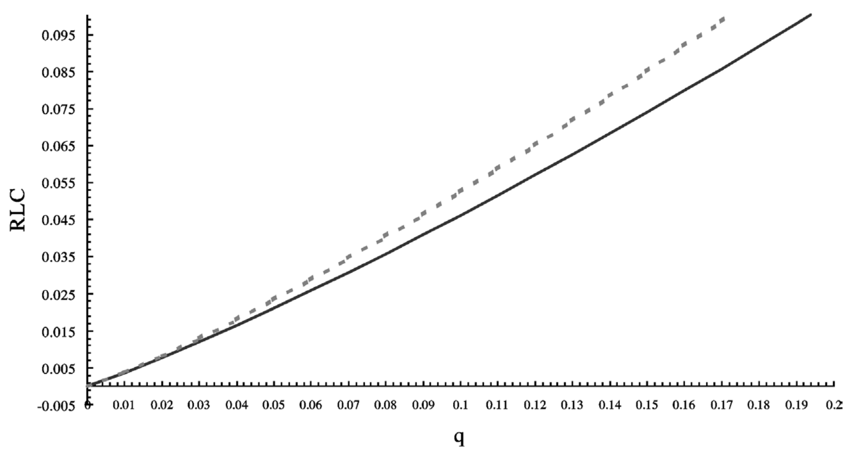

Figure 8. Relative Lorenz Curves for Germany 1983 (——) and Sweden 1981 (- - -) With a 3% Bottom Trim.

1981 on 3% bottom-tailed trimmed samples. The results are presented in Figure 9. The relative Lorenz dominance is indeed significant, except for the first q. This result is not surprising, because usually the sample sizes are large and thus the standard errors are small. If dominance is not significant, then this should appear at the smallest or the largest q values.

Second, note the behavior of the truncation profiles. Both q^{**} and q^{*} initially fall rapidly for α very close to 0 and thereafter decrease more gently. So the Lorenz comparison is certainly very sensitive to the presence or absence of the first few observations (in either the one- or two-tailed case), but the issue is clearly not just one of hypersensitivity to very small incomes. It seems unreasonable to suppose that the true picture is of strict Lorenz dominance, in that at least 1,000 observations would have to be discarded from the German data ($n \simeq 42{,}000$) for this conclusion to obtain.

JBES asa v.2004/12/08 Prn:29/10/2005; 16:12 F:jbes03m094.tex; (R) p. 7

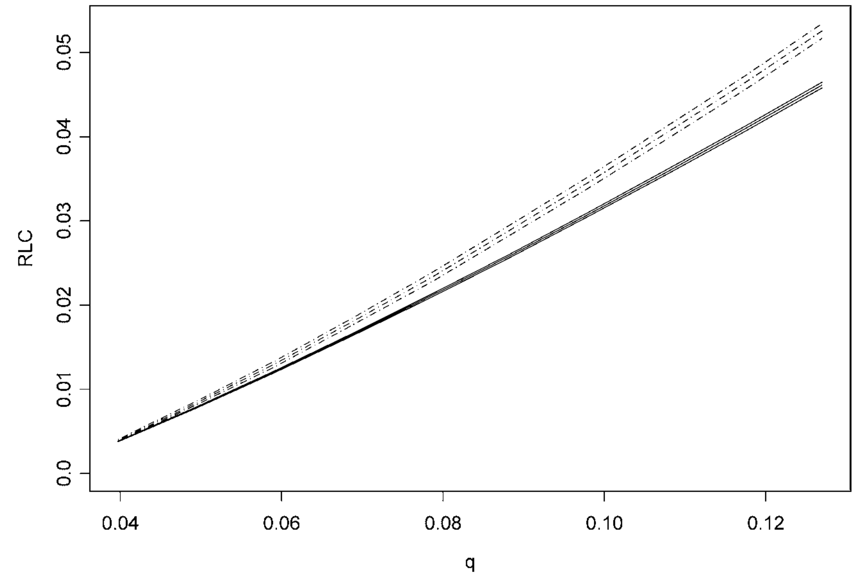

Figure 9. RLC of Germany 1983 (——) versus Sweden 1981 (- - -) With Confidence Intervals.

4.2 Inequality Over Time: The U.S. in the 1980s

Of course, the same technique may be applied to comparisons within one country but between two points in time. In the United States, the conventional wisdom is perhaps even sharper in its sketch of recent events; inequality rose over the 1980s (DeNavas-Walt and Cleveland 2002; Weinberg 1996). Again, the fact is—perhaps surprisingly—that the raw data do not reveal an unambiguous increase in inequality in the standard sense of relative Lorenz dominance. It may appear that this is due principally to the presence of negative incomes in the first centile group. As we will see, this is not quite the whole story. Note first that $F_{\text{US86}} \perp_C F_{\text{US79}}$; we do not have first- or second-order distributional dominance (see Fig. 10; the generalized Lorenz curves intersect at about $q = .02, .85$, etc.), but $F_{\text{US79}} \succeq_A F_{\text{US86}}$ (see Fig. 11). Again, monetary units are 1981 U.S. dollars. In addition, Figure 12 shows that $F_{\text{US86}} \perp_L F_{\text{US79}}$.

The trimming procedure is more complex. The issue of negative incomes is disposed of by a very modest (<.5%) trim, but there remains an issue of multiple intersections of the relative Lorenz curves at the bottom tail (with intersections between $q = .01$ and $q = .02$ and $q = .03$ and $q = .04$). Figure 13 plots $q^{**}(\alpha)$ and $q^{*}(\alpha)$ in this case. In view of the multiple intersections, these values are interpreted as the maximum switchpoint between the two Lorenz curves for each value of α. We find that $\alpha^{**} = .06$ and $\alpha^{*} = .03$. The outcome of the α-trimming procedure is interesting in that—in contrast to the Germany versus Sweden example—neither $q^{**}(\cdot)$ nor $q^{*}(\cdot)$ is monotonic. After dropping some 300–350 observations (3%) in the single-tailed trimming, or 600–700 observations (6%) in the two-tailed trimming, one may then conclude that $F_{\text{US79}} \succeq_{L_\alpha} F_{\text{US86}}$ (see Fig. 14). However, so much would have to be trimmed in either case that again it appears unreasonable to suppose that the true picture is one of strict Lorenz dominance.

There are some interesting points in common with the Germany-versus-Sweden example. First, for values of α in the range [0, .01], we find a relationship between the switchpoint and α, which is clearly different from the relationship that holds in the neighborhood of the points α^{**} and α^{*}. Second, the shape of the two-tailed trimming truncation profile closely follows that of the one-tailed trimming. On multiplying by $\frac{1}{2}$

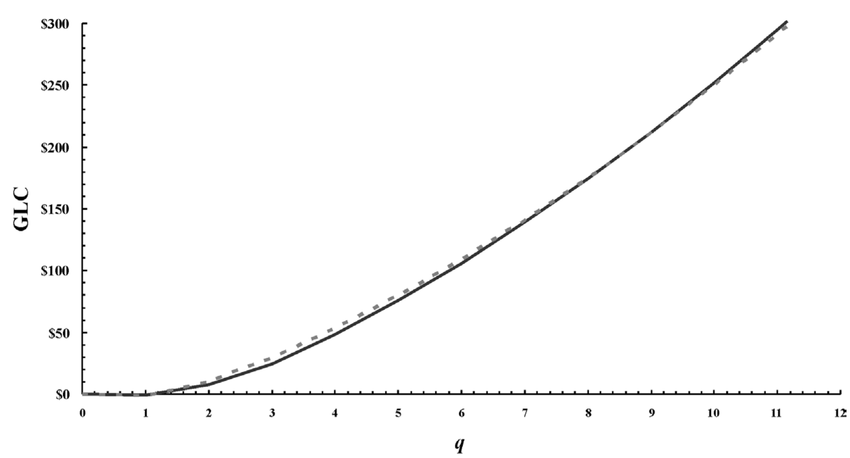

Figure 10. U.S. 1986 (——) Does Not Second-Order Dominate U.S. 1979 (- - -).

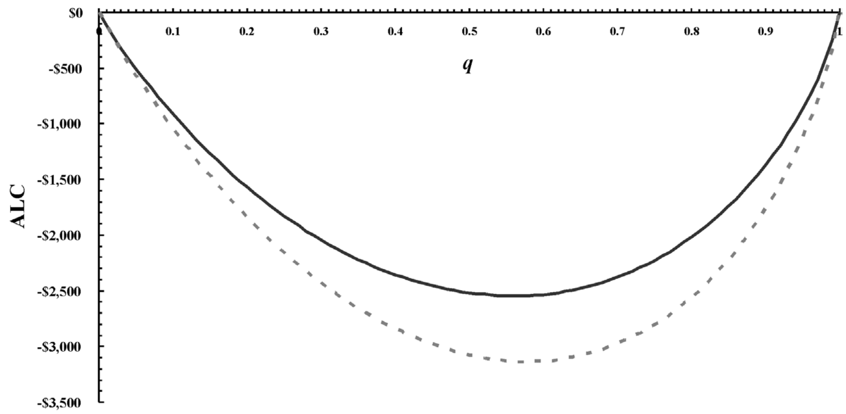

Figure 11. Absolute Lorenz Curves for U.S. 1979 (——) and 1986 (- - -).

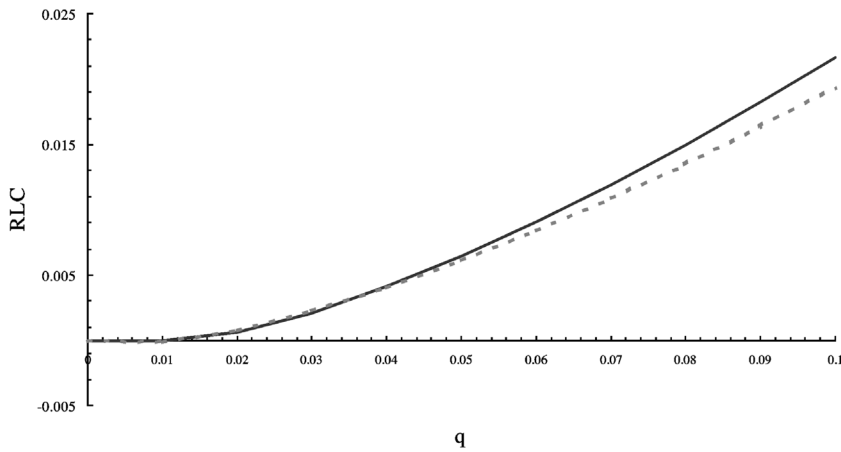

Figure 12. Relative Lorenz Curves for U.S. 1979 (——) and U.S. 1986 (- - -).

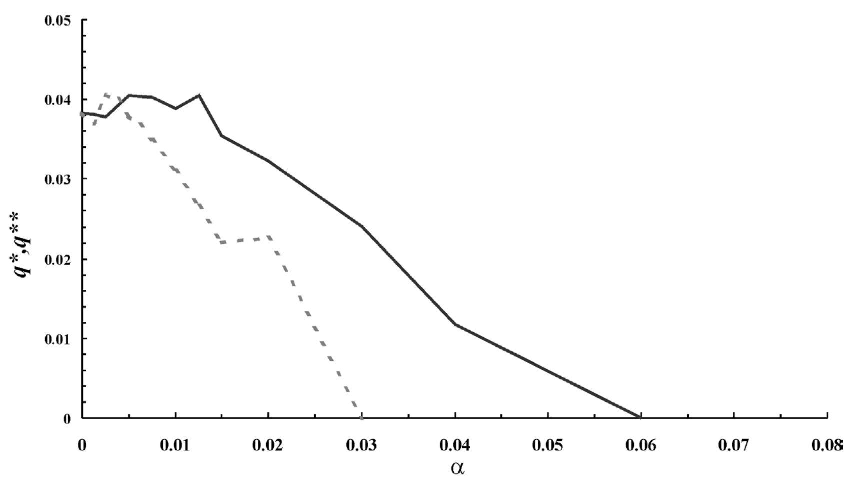

Figure 13. Did Inequality Rise in the U.S.? (——, two tail; - - -, one tail.)

the horizontal scale of the graph of $q^{**}(\cdot)$, we find that it lies extremely close to that of $q^{*}(\cdot)$; dropping $2\alpha\%$ of the sample in a two-tailed trimming has almost exactly the same impact on the Lorenz intersection as dropping $\alpha\%$ of the sample in a lower-tailed trimming. Third, all of the action appears to come from the lower tail. In the distributional comparisons reported in Sections 4.1 and 4.2 we also carried out an upper-tailed experiment; here the hypothesis is that the data contamination is concentrated in the high incomes and can be interpreted as potentially misreported data. However, in this case the ranking results turned out to be insensitive to the trim.

5. CONCLUSION

Given that second-order distributional-dominance criteria are known to be nonrobust, it is important to have practical meth-

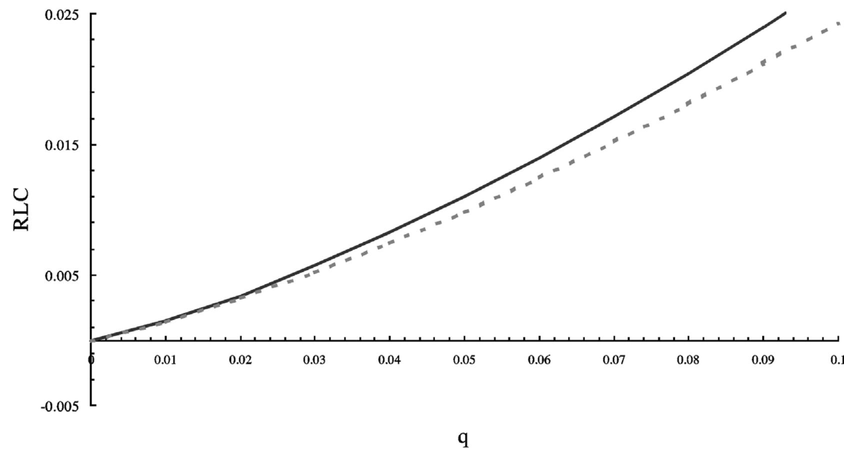

Figure 14. Relative Lorenz Curves for U.S. 1979 (——) and U.S. 1986 (- - -) With a 2% Bottom-Tailed Trimming.

ods of coping with the impact of potentially "dirty" data in either tail of an income distribution. One-tailed or two-tailed (i.e., balanced) trimming provides an obvious way to extend the simple distributional-dominance criteria. In effect, the researcher has the option of trading off efficiency of the distributional-dominance statistic with robustness. In this way one can place intuition about comparisons of empirical Lorenz curves on an appropriate analytical foundation. Another approach would involve trying to parameterize the upper or lower tail of the income distribution using robust estimation. We treated this approach in an earlier article (Cowell and Victoria-Feser 2001).

ACKNOWLEDGMENTS

This work was partially supported by ESRC grant R000 23 5725. The second author is supported in part by Swiss National Fund grants 610-057883.99 and PP001-106465. The authors thank the staff of the Luxembourg Income Study and Jean-Marc Falter for cooperation with the data analysis, as well as Julie Litchfield and Ceema Namazie for research assistance.

APPENDIX: **????**

A.1 Computational Method

When using a database such as the LIS database, from which the microdata cannot be recovered directly, LC or RLC with confidence intervals at each chosen q can be computed following this procedure:

1. Define the percentiles p (say $p = 0, .01, .02, \dots, 1$) and trimming proportions $\underline{\alpha}$ and $1 - \bar{\alpha}$.
2. Extract the personal incomes and weights from the database and sort the incomes and weights by incomes.
3. Define a new variable, **empperc**, composed of the cumulative weights, and divide all elements by the maximum, that is, the last element. Keep only the incomes and the weights for which **empperc** is between $\underline{\alpha}$ and $\bar{\alpha}$. This defines the trimmed incomes, **trinc**, and weights, **trwgt**.
4. Define **tottrwgt** as the sum of all **trwgt** and define **nbtrinc** as the number of elements in **trinc**.
5. Define a new variable, **trempperc**, composed of the cumulative **trwgt**, and divide all elements by the maximum, that is, the last element (and keep the value of the maximum of the cumulative **trwgt** in say **totweight**).
6. For each percentile $p > 0$, then do the following:

   a. Select the elements of **trinc** and **trwgt** for which **trempperc** is between p and the previous p; for example, for $p = .56$, **trempperc** is between .56 and 0.55. Call these **trincp** and **trwgtp**.

   b. Define $m1_p$ as the sum of **trincp·trwgtp** divided by **tottrwgt**, $m2_p$ as the sum of **trincp·trincp·trwgtp** divided by **tottrwgt**, and x_p as the maximum of **trincp**.
7. Define $q = \underline{\alpha} + (1 - \alpha)p$. Then for $q > \underline{\alpha}$, estimate $c_{\alpha,q}$ by the cumulative sum of the $m1_p$, $p \le \frac{q-\alpha}{(1-\alpha)}$ and $c_{\alpha,\underline{\alpha}} = 0$, and estimate $s_{\alpha,q}$ by the cumulative sum of the $m2_p$, $p \le \frac{q-\alpha}{(1-\alpha)}$ and $s_{\alpha,\underline{\alpha}} = 0$. Note that $\mu_\alpha = c_{\bar{\alpha}}$.

The 95% confidence intervals for the GLC and the RLC are $(c_q - 1.96\omega_{qq}; c_q + 1.96\omega_{qq})$ and $(c_q/\mu_\alpha - 1.96\upsilon_{qq}, c_q/\mu_\alpha + 1.96\upsilon_{qq})$, in which ω_{qq} (and thus υ_{qq}) are estimated using the estimates of $c_{\alpha,q}$ and $s_{\alpha,q}$. Note that $m1_p$ and/or $m2_p$ can take very large values depending on the measurement scale of the incomes. For numerical reasons, may be useful to divide all incomes by a properly chosen quantity.

A.2 Data Specification

LIS permits comparison of different countries' income distributions based on consistent international definitions of income and the income receiver. Accordingly, the same basic specifications were used for both the Germany and Sweden and the U.S. 1979 and U.S. 1986 comparisons in Section 4. The sample sizes were as follows:

Germany 1983:	42,752
Sweden 1981:	9,625
U.S. 1979:	15,928
U.S. 1986:	12,600.

The income distributions were formed using the following concept of equivalized incomes (Buhmann, Rainwater, Schmaus, and Smeeding 1988; Coulter, Cowell, and Jenkins 1992):

$$y = \frac{hhy}{hhsize^{\alpha}},$$

where *hhy* is net family (unit) income after tax, *hhsize* is the number of persons in the family unit, and $\alpha = .5$. Each observation is given a weight, $indwgt = hhsize * hweight$, to obtain distributions of income across individuals (Cowell 1984; Danziger and Taussig 1979). The variable *hweight* is the family unit sample weight.

For calculating distributions for different years and in dollars, the following data from the IMF Year Book 1994 were used:

	1981	*1983*
Price-level consumption		
Germany	106.3	115.6
Sweden	112.1	132.6
Dollar exchange rate		
Germany	2.260	2.553
Sweden	5.063	7.667

[Received May 2003. Revised August 2005.]

REFERENCES

Atkinson, A. B., and Bourguignon, F. (1989), "The Design of Direct Taxation and Family Benefits," *Journal of Public Economics*, 41, 3–29.

Ben Horim, M. (1990), "Stochastic Dominance and Truncated Sample Data," *Journal of Financial Research*, 13, 105–116.

Buhmann, B., Rainwater, L., Schmaus, G., and Smeeding, T. (1988), "Equivalence Scales, Well-Being, Inequality and Poverty: Sensitivity Estimates Across Ten Countries Using the Luxembourg Income Study (LIS) Database," *Review of Income and Wealth*, 34, 115–142.

Coulter, F. A. E., Cowell, F. A., and Jenkins, S. P. (1992), "Equivalence Scale Relativities and the Extent of Inequality and Poverty," *Economic Journal*, 102, 1067–1082.

Cowell, F. A. (1984), "The Structure of American Income Inequality," *Review of Income and Wealth*, 30, 351–375.

——— (2000), "Measurement of Inequality," in *Handbook of Income Distribution*, eds. A. B. Atkinson and F. Bourguignon, Amsterdam: North-Holland, pp. 87–166.

Cowell, F. A., and Victoria-Feser, M.-P. (1996), "Robustness Properties of Inequality Measures," *Econometrica*, 64, 77–101.

——— (2001), "Robust Lorenz Curves: A Semi-Parametric Approach," Distributional Analysis Discussion Paper 50, STICERD, London School of Economics.

——— (2002), "Welfare Rankings in the Presence of Contaminated Data," *Econometrica*, 70, 1221–1233.

——— (2003), "Distribution-Free Inference for Welfare Indices Under Complete and Incomplete Information," *Journal of Economic Inequality*, 1, 191–219.

Dalton, H. (1920), "Measurement of the Inequality of Incomes," *Economic Journal*, 30, 348–361.

Danziger, S., and Taussig, M. K. (1979), "The Income Unit and the Anatomy of Income Distribution," *Review of Income and Wealth*, 25, 365–375.

DeNavas-Walt, C., and Cleveland, R. W. (2002), *Money Income in the United States: 2001*, Current Population Reports: Consumer Income P60-191, Washington, DC: Census Bureau.

Esberger, S. E., and Malmquist, S. (1972), *En Statisk Studie av Inkomstutveklingen*, Stockholm: Statisk Centralbyr å n och Bostadssyrelsen.

Gottschalk, P., and Smeeding, T. M. (2000), "Empirical Evidence on Income Inequality in Industrialized Countries, in *Handbook of Income Distribution*, eds. A. B. Atkinson and F. Bourguignon, Amsterdam: North-Holland, Chap. 3.

Hampel, F. R. (1968), "Contribution to the Theory of Robust Estimation," unpublished doctoral thesis, University of California, Berkeley, Dept. of **???**.

——— (1974), "The Influence Curve and Its Role in Robust Estimation," *Journal of the American Statistical Association*, 69, 383–393.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., and Stahel, W. A. (1986), *Robust Statistics: The Approach Based on Influence Functions*, New York: Wiley.

Howes, S. R. (1996), "The Influence of Aggregation on the Ordering of Distributions," *Economica*, 63, 253–272.

Huber, P. J. (1981), *Robust Statistics*, New York: Wiley.

Jenkins, S. P., and Cowell, F. A. (1994), "Dwarfs and Giants in the 1980s: The UK Income Distribution and how It Changed," *Fiscal Studies*, 15, 99–118.

Kendall, M., and Stuart, A. (1977), *The Advanced Theory of Statistics*, London: Griffin.

Kleiber, C., and Kotz, S. (2003), *Statistical Size Distributions in Economics and Actuarial Sciences*, Hoboken, NJ: Wiley.

Lorenz, M. O. (1905), "Methods for Measuring Concentration of Wealth," *Journal of the American Statistical Association*, 9, 209–219.

Monti, A. C. (1991), "The Study of the Gini Concentration Ratio by Means of the Influence Function," *Statistica*, 51, 561–577.

Moyes, P. (1987), "A New Concept of Lorenz Domination," *Economics Letters*, 23, 203–207.

Pen, J. (1971), *Income Distribution*, London: Allen Lane.

Rawls, J. (1972), *A Theory of Justice*, Oxford, U.K.: Oxford University Press.

Stein, W. E., Pfaffenberger, R. C., and French, D. W. (1987), "Sampling Error in First-Order Stochastic Dominance," *Journal of Financial Research*, 10, 259–269.

Van Praag, B. M. S., Hagenaars, A. J. M., and Van Eck, W. (1983), "The Influence of Classification and Observation Errors on the Measurement of Income Inequality," *Econometrica*, 51, 1093–1108.

Weinberg, D. H. (1996), *A Brief Look at U.S. Postwar Income Inequality*, Current Population Reports: Household Economic Studies P60-191, Washington, DC: Census Bureau.

Wiles, P. J. D. (1974), *Income Distribution, East and West*, Amsterdam: North-Holland.

Wiles, P. J. D., and Markowski, S. (1971), "Income Distribution Under Communism and Capitalism," *Soviet Studies*, 22, 344–369, 485–511.