Queen's Economics Department Working Paper No. 1530

Empirically Implementing a Social Welfare Inference Framework

Charles Beach

Russell Davidson
McGill University

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

2-2025

Empirically Implementing a Social Welfare Inference Framework*

By

Charles M. Beach

Department of Economics

Queen's University

Kingston, Ontario, Canada K7L 3N6

email: beach.chaz3@gmail.com

and

Russell Davidson

Department of Economics and CIREQ
McGill University
Montreal, Quebec, Canada

Aix-Marseille Universite
CNRS, EHESS, AMSE
13205 Marseilles cedex 01, France

Email: russell.davidson@mcgill.ca

Key words: social welfare tests, income distribution comparisons, implementing social welfare criteria_JEL codes: C10, D31, D63, 131

*The authors thank Stefan Fassler for remarkably able computer RA work and Roda Mendoza for excellent technical assistance for this paper, and James Davies for helpful comments on an earlier working paper on this topic by the first author. The authors bear full responsibility for any shortcomings of this paper.

February 2025

ABSTRACT

This paper builds on recent econometric developments establishing distribution-free statistical inference methods for quantile means and income shares for a sample distribution of microdata to propose an approach to empirically Implement several dominance criteria for comparing economic well-being and general income inequality between distributions. It provides straightforward variance-covariance formulas in a set of practical empirical procedures for formally testing economic well-being and inequality comparisons such as rank dominance, Lorenz dominance and generalized Lorenz dominance between distributions.

The tests and procedures are illustrated with Canadian census data between 2000 and 2020 on women's and men's incomes. It is found that both women's and men's economic well-being statistically significantly improved over this period, while income inequality significantly increased over 2000-15 and then fell over 2015-20.

1. Introduction

Since about 1980 to the early twenty-first century, income inequality in many developed economies rose dramatically to historic levels (Guvenen et al., 2022; Hoffman et al., 2020). Structural factors behind these changes have been extensively reviewed in, for example, Acemoglu et al. (2016); Autor, Dorn and Hanson (2013); Beach (2016); Goos, Manning and Salomons (2014); and Saez and Veall (2007), and have highlighted the major roles of automation and technological advances, globalization and trade, as well as demographics, deregulation and policy changes such as minimum wages. More recently, world events such as the COVID pandemic, the war in Ukraine and growing trade restrictions with China have elicited major economic adjustments as supply chains are reworked and much basic and high-tech manufacturing is being reshored, and as income support programs burgeon (at least temporarily) and then labour markets tightened — resulting in notable improvements in lower incomes and a new "unstuck middle" (The Economist, 2023). Such complex developments give rise to quite different effects over different groups of workers in the labour market and over different regions of the income distribution. Clearly, a disaggregative approach to examining distributional changes is much more informative than conventional summary measures of inequality (such as a Gini coefficient or coefficient of variation).

Such major developments have given rise to issues of "equitable growth"(Drummond, 2021), and "common prosperity" (The Economist, 2021a, 2021b) and ways of linking GDP and social welfare measurement (Fleurbaey, 2009; Jones and Klenow, 2016; and Jorgenson, 2018). Statistical agencies and National Accounts analysts are part of a burgeoning interest in the broad endeavour of distributional National Accounts (see, eg., Alvaredo et al., 2018, 2020;

Zucman et al., 2018; and Rapley, 2024). The theoretical economics literature has been examining the distributional incidence of growth and its social welfare implications (Palmisano and Peragine, 2015) and its related implications for equality of opportunity in a growing economy (Peragine, 2004; Bosmans and Ozturk, 2021). Empirically implementing such concerns has been reviewed, for example, in Cowell (2021), and Duclos and Araar (2006). In both Canada and the United States, federal governments have also been focusing policies to better target low-incomes and middle-class families. And in Canada, dramatic increases in immigration rates and inflows of low-wage temporary foreign workers have led to debate on whether the growth objective should be for GDP or GDP per capita as economic growth rates broadly decline (Rendell, 2024). All these developments call for an empirically implementable set of criteria or tools to blend these various objectives and to better evaluate if people are indeed better (or worse) off, income inequality has been reduced, or economic opportunities have been improved.

The availability of such tools — as provided in this paper — is based on three advances in the area. First is the advent of large publicly-available microdata sets — such as census files in Canada, the United States and elsewhere — that allow for very detailed analysis of different groups within the overall income distribution. The second is the development in the theoretical social welfare literature of various criteria for judging whether a change in income distributions leaves the population as a whole better (or worse) off. Thirdly, in the econometrics literature, Beach and Davidson (2024a) establish the basis for drawing statistical inferences on disaggregative inequality measures by deriving the (asymptotic normality and asymptotic) variance-covariance structure of the sample means and income shares of quantile groups across a distribution (eg., for each decile group within a distribution).

This paper applies the formulas of Beach and Davidson (2024a) and offers an empirical approach to evaluate distributional changes in an easily implementable (distribution-free) framework of statistical inference based on well-known disaggregative distributional statistics. The paper shows how various theoretical criteria can be implemented in practice in straightforward fashion using explicit formulas and without having to know the underlying income distribution function. It imbeds this implementation within a standard statistical inference framework in terms of a practical empirical criterion (PEC) for empirically testing hypotheses of distributional dominance and establishing whether such distributional comparisons are indeed statistically significant. And it helps unify the applicability of various social welfare and general inequality comparisons within a common framework of formulas and procedures for hypothesis testing.

The paper is organized as follows. The next section sets out the quantile function approach for establishing statistical inference formulas for quantile-group means and income shares. These are the basis for the rest of the paper. Section 3 explains the normative perspective underlying the various distributional dominance criteria used in the paper. Then Sections 4-7 apply these developments to a range of dominance criteria for both social welfare and general inequality comparisons typically involving Lorenz curves and their extensions. The distributional statistics and test criteria are then empirically illustrated with Canadian census public use microdata files over 2000 - 2020 in Sections 8 and 9. The final section reviews the main results of the paper and draws some implications.

2. Quantile Function Approach for Quantile Means and Income Shares

Empirical measures of economic well-being and income inequality are built up from disaggregative statistics on quantile income shares and quantile mean income levels. Quantile statistics are those that are expressed in terms of given percentage groups of the ranked or ordered observations in a microdata sample. In the case of income distribution statistics, the data observations in a sample are ordered by income from the lowest income observation to the highest income observation. The ordered observations are then divided into non-overlapping income groups, say, in terms of ten deciles or twenty vigintiles (generically referred to as quantile income groups or simply quantiles). So the first decile group consists of those observations with the 10 percent lowest incomes, the second decile group includes the next 10 percent lowest income recipients, and so on up to the top or tenth decile group which includes those 10 percent of income recipients with the highest income levels in the sample. The standard Lorenz curve (of cumulated income shares), for example, is based around such percentile groups. Quantile means and income shares can be expressed in terms of integrals of underlying distribution functions, what we will call quantile functions. The quantile function approach used in this paper involves suitably approximating the quantile functions that the distributional statistics are based on in order to identify and establish the asymptotic properties of the corresponding sample statistics. The approach is applied to quantile means and quantile income shares.

The key feature of these quantile statistics is that the relative sizes of the quantile groups are *given* percentages of the sample or distribution. This turns out to simplify quite dramatically the sampling properties of quantile-based statistics. (Contrast this with, say, median-based income groups where the middle-income group consists of those with incomes between, say, 50 percent and 200 percent of the median income level — a common designation of the so-called Middle Class. In this case, the size of the group is not given, but is a consequence of the sampling process. As a result, the sampling properties of median-based statistics are considerably more complex and are distribution-dependent, so are less convenient to deal with (Beach and Davidson, 2024b).

2.1 Quantile Function Approach for Quantile Means

To formalize the quantile function approach taken in this paper, we summarize the analytical statistical development in Beach and Davidson (2024a).

Suppose the distribution of income Y is divided into K ordered income groups. Let the dividing proportions of recipients be $p_1 < p_2 < \ldots < p_{K-1}$ (with p_0=0 and p_K=1). The mean income of recipients with incomes between the p_i and p_{i-1} quantiles, for i=1, 2,..., K, is given by the quantile function

$$\mu_i = \int_{\xi_{i-1}}^{\xi_i} y\, dF(y) \Big/ \int_{\xi_{i-1}}^{\xi_i} d\,F(y) = \left(\frac{1}{p_i - p_{i-1}}\right) \int_{\xi_{i-1}}^{\xi_i} y\, dF(y)$$

where ξ_i is the p_i quantile cut-off income level of the distribution F, and ξ_0 is taken to be the smallest (possibly negative) income in the support of the distribution. Let

$$n_i = \int_{\xi_{i-1}}^{\xi_i} y\, dF(y) = (p_i - p_{i-1})\, \mu_i$$

for i = 1,..., K, and consider the constructed variable

$$W_i = Y \cdot I\,(\xi_{i-1} < Y \le \xi_i) - \xi_i \cdot I\,(Y \le \xi_i) + \xi_{i-1} \cdot I\,(Y \le \xi_{i-1})$$

where $I\,(\cdot)$ is an indicator function which takes the value 1 if the condition in the parentheses holds, and 0 otherwise. $\widehat{n_i}$ is the sample estimate of n_i and $\hat{\mu}_i$ is the sample estimate of μ_i.

Then it can be shown that the set of K $\widehat{n_i}$'s are asymptotically joint normally distributed, and the asymptotic variance of

$$\hat{n}_i = (p_i - p_{i-1})\,\hat{\mu}_i$$

is the same as the regular variance of the constructed random variable W_i which in turn is derived to be

$$Asy.var\,(\hat{n}_i) = Var\,(W_i)$$

$$= \sigma_i{}^2 + \xi_i{}^2 p_i\,(1 - p_i) + \xi_{i-1}{}^2 p_{i-1}(1 - p_{i-1}) - 2\,\xi_{i-1}\,\xi_i\,p_{i-1}\,(1 - p_i)$$

$$- 2\,n_i\,[\xi_i\,(1 - p_i) + \xi_{i-1}\,p_{i-1}] \qquad (1)$$

where $\sigma_i{}^2 = Var\,(Y \mid \xi_{i-1} < Y \leq \xi_i)$. The first term in this expression represents the variability of Y within the $\xi_i - \xi_{i-1}$ range, the next three terms capture the randomness of the sample estimates $\hat{\xi}_i$ and $\hat{\xi}_{i-1}$, and the last combined term captures the covariance of the first two effects. Thus

$$Asy.var\,(\hat{\mu}_i) = \left(\frac{1}{p_i - p_{i-1}}\right)^2 \cdot Asy.var\,(\hat{n}_i). \qquad (1a)$$

Similarly, for the covariances, for $j < i$,

$$Asy.\widehat{\ cov}\left(\hat{n}_i, \hat{n}_j\right) = Cov\left(W_i, W_j\right)$$

$$= (\xi_i - \xi_{i-1})\left(p_j\,\xi_j - p_{j-1}\,\xi_{j-1} - n_j\right)$$

$$- \left(n_j - p_j\,\xi_j + p_{j-1}\,\xi_{j-1}\right)(n_i - p_i\,\xi_i + p_{i-1}\,\xi_{i-1})$$

$$= \left(p_j\,\xi_j - p_{j-1}\,\xi_{j-1} - n_j\right)[n_i + \xi_i\,(1 - p_i) - \xi_{i-1}\,(1 - p_{i-1})]. \qquad (2)$$

The variance and covariances are consistently estimated by replacing all unknown terms by their sample counterparts. Also,

$$Asy.cov\left(\hat{\mu}_i, \hat{\mu}_j\right) = \frac{Asy.cov\left(\hat{n}_i,\hat{n}_j\right)}{(p_i - p_{i-1})(p_j - p_{j-1})} \tag{3}$$

and $S.E(\hat{\mu}_i) = \left[\frac{Asy.\widehat{v}ar\,(\hat{\mu}_i)}{N}\right]^{1/2}$

$$= \left[\frac{Asy.\widehat{v}ar\,(\hat{n}_i)}{N\,(p_i - p_{i-1})^2}\right]^{1/2}.$$

Denote the asymptotic variance-covariance matrix of the $\hat{\mu}_i$'s in (1a) and (3) by V_s and the estimated regular variance-covariance matrix of the $\hat{\mu}_i$'s as $\hat{V}$, whose elements are gotten by dividing the estimated elements of (1a) and (3) by sample size N.

Note that all of these expressions are distribution-free in that they do not involve F( ) or its density f( ) in their calculations. Thus one avoids the need for computer-based estimation techniques for density ordinate evaluation such as kernel estimation methods; and since we have specific variance-covariance estimation formulas, computer-based bootstrapping methods are no longer needed as well.

2.2 Quantile Function Approach for Income Shares

In similar fashion, one can express the income share that accrues to recipients with income between ξ_{i-1} and ξ_i in terms of the quantile function

$$IS_i = \left(\frac{1}{\mu}\right)\int_{\xi_{i-1}}^{\xi_i} ydF(y) = (p_i - p_{i-1})\frac{\mu_i}{\mu} = \frac{n_i}{\mu}$$

for i=1, ..., K. If $\widehat{IS}_i$ is the natural sample estimator of IS_i, then again it can be shown that the set of $\widehat{IS}_i$'s are asymptotically joint normal, and the asymptotic variance of $\widehat{IS}_i$ can be shown to be the same as the regular variance of the constructed variable

$$V_i = \left(\frac{1}{\mu}\right)\left(W_i - \frac{n_i\, Y}{\mu}\right).$$

As useful definitions, let

$m_i = \sum_{k=1}^{i} n_k$ and $n_{2i} = \int_{\xi_{i-1}}^{\xi_i} y^2\, dF(y)$.

Then it can be shown that

$$Asy.var\left(\widehat{IS}_i\right) = Var\left(V_i\right)$$

$$= \left(\frac{1}{\mu^4}\right)\left[\mu^2 \cdot Var\left(W_i\right) + n_i^2 \cdot Var\left(Y\right) - 2\mu\, n_i \cdot Cov\left(W_i, Y\right)\right] \qquad (4)$$

where $Var\left(W_i\right)$ is given in (1), $Var\left(Y\right) = \sigma^2$ for the distribution, and

$Cov\left(W_i, Y\right) = n_{2i} - \mu\, n_i - \xi_i\left(m_i - p_i\mu\right) + \xi_{i-1}\left(m_{i-1} - p_{i-1}\mu\right)$.

Then the standard error of $\widehat{IS}_i$ is given by

$$S.E.\left(\widehat{IS}_i\right) = \left[\frac{Asy.\widehat{var}\left(\widehat{IS}_i\right)}{N}\right]^{1/2}$$

where again the estimated asymptotic variance has all its unknown elements replaced by their sample estimates.

As for the asymptotic covariances,

$$Asy.cov\left(\widehat{IS}_i, \widehat{IS}_j\right) = Cov\left(V_i, V_j\right)$$

$$= \left(\frac{1}{\mu^2}\right) \cdot Cov\left(W_i, W_j\right) + \left(\frac{n_i\, n_j}{\mu^4}\right) \cdot Var(Y) \qquad (5)$$

$$- \left(\frac{n_i}{\mu^3}\right) \cdot Cov\left(W_j, Y\right) - \left(\frac{n_j}{\mu^3}\right) \cdot Cov\left(W_i, Y\right)$$

and again $\widehat{Cov}\left(\widehat{IS}_i, \widehat{IS}_j\right) = \frac{Asy.\widehat{cov}\left(\widehat{IS}_i, \widehat{IS}_j\right)}{N}$.

Note as well that all the above formulas are again distribution-free. Again, denote the asymptotic variance-covariance matrix of the $\widehat{IS}_i$'s in eqs (4) and (5) by W_s and the estimated regular variance-covariance matrix of the $\widehat{IS}_i$'s by $\widehat{W}$, whose elements are gotten by dividing the estimated elements in (4) and (5) by N.

Note also the flexibility of the quantile function approach applied to percentile income groups. The latter groups do not have to be of uniform size. Depending on the empirical analysis being undertaken, one may wish to have narrower groups (eg., deciles or vigintiles) towards the two ends of the distribution and wider groups (eg., quintiles) over the middle range of the distribution.

2.3 Illustration: Mean Differences and Growth Rates

One can also apply the above results to look at differences in individual quantile means between different population groups — such as quantile mean earnings *differences* between female and male workers in the labour market — and at *changes* in separate quantile means between time periods. So long as the estimates being compared are from independent samples, the variance of the difference in sample estimates is simply the sum of the respective variances, and the standard error of the differences is given by

$$S.E.(\hat{\mu}_i(b) - \hat{\mu}_i(a)) = \left[\frac{Asy.\widehat{var}\,(\hat{\mu}_i(b))}{N(b)} + \frac{Asy.\widehat{var}\,(\hat{\mu}_i(a))}{N(a)}\right]^{1/2}$$

where designators a and b refer to the two separate sample estimates. A quantile analysis thus allows for potentially quite detailed disaggregative examination of differences between distributions. Beach and Davidson (2024a) sets out a range of toolbox measures available to further the perspective and flexibility of such analyses.

Following a concern for issues of equitable growth, one could also express quantile mean differences in relative or percentage terms — or what The Economist (2021c, p.24) refers to as Piketty lines of different growth rates of quantile means across the various regions of the income distribution. In this case, if

$$\hat{q}_i = (\hat{\mu}_i(b) - \hat{\mu}_i(a))/\hat{\mu}_i\,(a) = \left(\frac{\hat{\mu}_i(b)}{\hat{\mu}_i\,(a)}\right) - 1,$$

between time periods a and b, then approximately

$$\widehat{Var}(\hat{q}_i) = \left(\frac{-\hat{\mu}_i(b)}{(\hat{\mu}_i(a))^2}\right)^2 \cdot \widehat{Var}\big(\hat{\mu}_i(a)\big) + \left(\frac{1}{\hat{\mu}_i(a)}\right)^2 \cdot \widehat{Var}(\hat{\mu}_i(b))$$

$$= \left(\frac{\hat{\mu}_i(b)}{(\hat{\mu}_i(a))^2}\right)^2 \cdot \left[\frac{Asy\widehat{\ }var\big(\hat{\mu}_i\,(a)\big)}{N\,(a)}\right]$$

$$+ \left(\frac{1}{\hat{\mu}_i(a)}\right)^2 \cdot \left[\frac{Asy\widehat{.\,v}ar\,\big(\hat{\mu}_i(b)\big)}{N\,(b)}\right],$$

and again

$$S.E.(\hat{q}_i) = \left[\hat{V}ar\,(\hat{q}_i)\right]^{1/2}.$$

The quantile means and income shares serve as the basis for operationally implementing the evaluation of changes in social welfare and general income inequality in the next several sections of this paper.

3. A Normative Perspective for Evaluating Changes in Social Welfare and Inequality

The traditional way of measuring income inequality in an income distribution is in terms of some summary or aggregate measure of inequality such as the Gini coefficient (G), coefficient of variation (C), relative mean (absolute) deviation (M), or the standard deviation of the logs of income (L). But such measures are subject to two basic criticisms. First is the aggregation problem: various summary measures aggregate income differences in different ways, so that different measures can give different results when comparing two distributions. One way (partially) to address this is to identify several desirable properties we may want such summary measures to satisfy. These could include for example:

i) Symmetry (or Anonymity) - An inequality measure depends only on incomes in a distribution and not on who has which incomes;

ii) Mean Independence - An inequality measure is invariant to proportional changes (eg., doubling) of all incomes (ie., it is a relative measure of inequality);

iii) Population Homogeneity - An inequality measure is invariant to replication of the population (eg., doubling the number of persons in the distribution while keeping the shape of the distribution the same);

iv) Principle of Transfers - Any transfer of $x from a richer person to a poorer person so that $y(i) + x < y(j) - x$ if initially $y(i) < y(j)$ should reduce inequality;

v) Transfer Sensitivity - A transfer of $x such as envisioned in (iv) should reduce inequality more if it occurs among a lower-income pair of individuals than if it occurs among a higher-income pair of individuals. This is obviously a stronger form of the Principle of Transfers.

It turns out that (i), (ii) and (iii) are satisfied by all the above four inequality measures, but (iv) is satisfied only by C and G, and (v) is not satisfied by any of them.

Alternatively, another way to address the aggregation problem is to rely on a disaggregative measure of inequality or a set of multiple statistics linked up graphically such as a Lorenz curve. A problem here, though, is that two empirical Lorenz curves being compared often cross, so a clear comparison is not straightforward.

The second basic criticism of conventional summary measures of inequality is the implicit value judgement problem. That is, any summary inequality measure involves implicit value judgements or weightings of different persons' incomes (or economic well-being), and thus contains embedded in it an implicit social welfare function (SWF). For example, various inequality measures differently emphasize income differences at the bottom, middle or upper end of the distribution. Consequently, it can be argued, it would be better to choose desirable SWF properties explicitly and then derive the implied inequality measure from the desired SWF. To do so is to take a *normative* approach to measuring inequality (originated by Atkinson, 1970) rather than the traditional descriptive approach. This normative approach is the one followed in the current paper.

To implement such a normative approach, one first needs to define a social welfare function and its basic properties. For a much more expansive discussion of the normative approach, see for example, Boadway and Bruce (1984), Blackorby et al. (1999), Lambert (2001), or Cowell (2011). Specifically, social welfare function W(.) is any function

$$W = f(U_1, \ldots, U_N)$$

that has as arguments U_i individual (or household) utility functions and that incorporates social values used to aggregate economic well-being across the population. The U_i can thus be viewed as a 'social income valuation function". To do this, we require that:

- the U_i's must be at least cardinal scale measurable in order to be aggregated across persons;
- the U_i's must have at least some degree of comparability across persons in the population (ie., if utilities are cardinally measured for each individual, the units of measurement must be the same across individuals); and
- for technical convenience, each U_i depends only on incomes and indeed only on individual i's income (ie., U_i = U_i (Y_i), so there is no envy or altruism).

One can then identify several possible desirable properties for such a social welfare function:

i) Pareto Principle - State X is socially preferred to state Y if at least one person strictly prefers X to Y and no one prefers state Y to X (ie., ∂U_i / ∂Y_i > 0 and social indifference curves in Y_i, Y_j space have negative slopes);

ii) Symmetry or Anonymity - Everyone's incomes are evaluated by using the same U(.) function (ie. U_i (.) = U(.) for all i =1, ..., N);

iii) Population Invariance - If the population is replicated K times, then social welfare increases K-fold (ie., W(Y_1, ..., Y_{KN}) =K. W(Y_1, ..., Y_N);

iv) Strict Concavity of the SWF or the Principle of Transfers - A strictly concave SWF is such that $\partial^2 U_i$ / ∂Y_i^2 < 0 for all i (this implies that social indifference curves are strictly convex to the origin).[1] This is sometimes referred to as an "egalitarian SWF";

[1] Actually, strict concavity is sufficient for the Principle of Transfers to hold. A weaker necessary and sufficient property for the Principle of Transfers is referred to as Schur concavity. See general discussion in Lambert (2001).

v) Transfer Sensitivity or the Principle of Diminishing Transfers – A transfer-sensitive SWF is such that $\partial^3 U_i \,/\, \partial Y_i^{\,3} > 0$. Again, this is a stronger version of the Principle of Transfers.

Atkinson (1970) uses this normative approach to show that, under properties (i) - (iv), an empirical proxy of social welfare or economic well-being (SW_p) can be expressed as

$$SW_p = \bar{Y}.\ (1 - I_A)$$

$$= \bar{Y}.\ \mathrm{E}$$

where $\bar{Y}$ is the mean income level of a distribution and I_A is a measure of inequality based on the above four properties and where it turns out that $0 \leq I_A \leq 1$ where higher values indicate greater levels of inequality in the distribution. That is, SW_p can be decomposed into two (multiplicative) components — an efficiency dimension ($\bar{Y}$) or average per capita income and an equity dimension (E) where E = 1- I_A.

If one further assumes a specific functional form for U(.) — in the convenient form of an iso-elastic social welfare function — Atkinson (1970) then derives a specific formula for the calculation of I_A. An iso-elastic SWF is general and flexible enough to incorporate a wide range of social attitudes to income inequality from a Benthamite utilitarian SWF to Rawls' maxi-min SWF.

But I_A is still a summary or aggregate measure of income inequality. What the social choice literature since Atkinson's (1970) paper has tried to do is to extend or apply Atkinson's normative perspective to develop a set of disaggregative criteria for comparing different income distributions based on the above properties, so that both criticisms of traditional inequality measures are addressed. The rest of this paper examines several such disaggregative criteria from the theoretical social choice literature and proposes ways to operationalize or empirically implement these criteria in terms of vectors of quantile means and income shares and related disaggregative distributional statistics. The paper also develops inference procedures to allow for formal statistical testing for these criteria. This development is applied to four such criteria in the following sections.

4. Applications to Rank Dominance and a Practical Empirical Criterion

One early example of a disaggregative normative ranking criterion for distributions comes from Saposnik (1981) and involves what may be called the quantile curve (Duclos and, Araar, 2006, p. 45) which is essentially the inverse of the cumulative distribution function of an income distribution. Saposnik's rank dominance theorem says that, for any social welfare function satisfying the properties of symmetry, population invariance and the Pareto principle (ie., social welfare conditions (i) - (iii)), distribution A is socially preferred to distribution B if the quantile curve for A is everywhere higher than that for B. Note that there is no egalitarianism built into this criterion. It essentially says that, if everyone has higher incomes in A than in B, then they must be better off. This is useful in comparing distributions many years apart, say for example, the Canadian income distributions for 1970 versus 2020. But in many practical cases faced by empirical researchers, this situation doesn't apply.

Nonetheless, it is useful to begin our application of dominance criteria with this relatively simple criterion. To empirically implement it, one represents the two distributions or quantile curves being compared by their respective vectors of sample quantile means, $\hat{\mu}_i$, for the i = 1, ..., K quantiles. The actual decision rule for determining the outcome of the comparison of vectors requires some practical empirical criterion (or PEC) based on the principles of statistical inference.

4.1 A Practical Empirical Criterion for Quantile Means

Following Beach, Davidson and Slotsve (1994) and Davidson and Duclos (2000), one can set out a two-step test procedure for the PEC. It is assumed that the data samples for the two distributions being compared are independent and hence do not overlap. Examples are, say, two different years of data (with no overlapping samples) being compared or two different (non-overlapping) population groups such as age, racial, or sex groups.

Step 1 - Test the joint null hypothesis of equality of the two (population) quantile mean vectors versus the alternative hypothesis of non-equality. This can be done by a standard (but asymptotic) chi-square test with K degrees of freedom, where K is the number of quantiles. For a meaningful disaggregative analysis, it makes sense to let K = 10 or 20, say, rather than a small number such as 4 or 5. If the null hypothesis is not rejected, then the two sets of quantile means can be said to be not statistically significantly different, and further comparison is not pursued. This is taken as an empirical proxy for comparison of the two underlying quantile curves.

Step 2 - If, however, the null hypothesis in Step 1 is rejected — which is the typical case when using large microdata sets for the sample distributions — then proceed to calculate separate t-statistics for differences in respective means for each of the individual quantile means. These K individual t-statistics, however, are correlated, and hence comparing each test statistic to the critical value on an (asymptotic) normal distribution would not be appropriate. One has to recognize that this Step 2 involves correlated multiple comparisons. Following the work of Beach and Richmond (1985) and Bishop, Formby and Thistle (1989, 1992) on multiple comparison testing, one should compare the K separate t-statistics (for differences in quantile means) to critical values on the Studentized Maximum Modulus (or SMM) distribution. If at least one of the quantile mean difference t-statistics has the appropriate sign and is statistically significant (based on the SMM distribution) and none of the t-statistics of the remaining quantile mean differences has the wrong sign and is significant, then conclude that the quantile curve with the higher sample quantile means rank dominates (or is socially preferred to) that with the lower quantile means. If not, then one can say only that the two quantile curves are statistically significantly different and not reach a preferred or dominance conclusion. Note that this is an asymptotic test and critical values from the SMM distribution correspond to K and infinite degrees of freedom. Typically useful critical values from the SMM distribution are:

	alpha = .01	alpha = .05	alpha = .10
K = 5	3.289	2.800	2.560
K = 10	3.691	3.254	3.043
K = 20	4.043	3.643	3.453

Source: Stoline and Ury (1979), Tables 1-3.

4.2 Full Variance-Covariance Matrix for Quantile Means

The first step in the above practical empirical criterion (PEC) involves a joint test of the difference between two vectors or sets of estimated quantile means. If the two quantile curves being compared are designated A and B, then the vectors of quantile means can be represented as

$$\hat{\mu}^a = (\hat{\mu}_1^a, \dots, \hat{\mu}_K^a)^1, \quad \mu^a = (\mu_1^a, \dots, \mu_K^a)^1$$

and

$$\hat{\mu}^b = \left(\hat{\mu}_1^b, \dots, \hat{\mu}_K^b\right)^1, \quad \mu^b = \left(\mu_1^b, \dots, \mu_K^b\right).$$

A standard result from statistics, then, shows that, if the random vector $\hat{\mu}^a$ is normally distributed with mean μ^a and variance-covariance matrix V^a, $\hat{\mu}^b$ is normally distributed with mean μ^b and variance-covariance matrix V^b, and $\hat{\mu}^a$and $\hat{\mu}^b$ are statistically independent, then $\hat{\mu}^b - \hat{\mu}^a$ is also normally distributed with mean $\mu^b - \mu^a$ and variance-covariance matrix $V^a + V^b$. Under the null hypothesis that the two vectors μ^a and μ^b are the same (ie., $\mu^b - \mu^a = 0$) then the quadratic form

$$\left(\hat{\mu}^b - \hat{\mu}^a\right)^1 [V^a + V^b]^{-1} \left(\hat{\mu}^b - \hat{\mu}^a\right)$$

is distributed as a chi-squared random variable with K degrees of freedom. If V^a and V^b are estimated consistently, then the test statistic for Step 1 of the PEC,

$$\left(\hat{\mu}^b - \hat{\mu}^a\right)^1 \left[\hat{V}^a + \hat{V}^b\right]^{-1} \left(\hat{\mu}^b - \hat{\mu}^a\right) \tag{6}$$

is asymptotically distributed as a chi-squared variate with K degrees of freedom. The elements of $\hat{V}^a$ and $\hat{V}^b$ are gotten from eq (1a) and (3), where each

$$\hat{v}_{i\,j} = \frac{Asy.\widehat{cov}\left(\hat{\mu}_i, \hat{\mu}_j\right)}{N}$$

and N is the respective sample size.

To perform the individual tests in Step 2 of the PEC, compute the standard "t-statistic" ratio for the difference between two independent random variables ($\hat{\mu}_i^a$ and $\hat{\mu}_i^b$) as

$$t_i = \frac{\hat{\mu}_i^b - \hat{\mu}_i^a}{\left[\hat{v}^a(i,i) + \hat{v}^b\,(i,i)\right]^{1/2}}$$

and compare this to the appropriate critical value on the SMM distribution.

Note that, in the above test procedure, the quantile income curve is approximated or represented by a vector of quantile means ($\hat{\mu}_i$'s) and not by a vector of quantile cut-off levels ($\hat{\xi}_i$'s). This is because the latter have a variance-covariance structure that is distribution-dependent (ie., involving f(·)'s), while the variance-covariance structure of the former has been shown to be (asymptotically) distribution-free. So the use of quantile means is convenient and sensible.

It could be argued that the test reliance on a set of quantile points is arbitrary and doesn't provide adequate coverage of the income distribution as a whole. But such a choice of quantiles is quite conventional and even standard in the income distribution literature — as witness by published official distribution statistics in terms of deciles and quintiles. This is in contrast, say, to the ranking of investment opportunities on financial portfolios where all alternatives need to be examined. The availability of large microdata files nowadays also allows considerable disaggregative detail (such as vigintiles, say, or even percentiles in large census or administrative files) as well as a flexible differentiated focus (such as vigintiles in the tails of the distribution and quintiles or deciles over the mid-range of the distribution).

5. Application to Lorenz Dominance

The same approach can be applied to an inequality-based dominance criterion (for more extensive discussions of Lorenz curve comparisons, see Maasoumi, 1998; Lambert, 2001; and Aaberge, 2000, 2001). Atkinson, in his famous 1970 paper, forwarded what has come to be known as the Lorenz dominance theorem. For any (summary) inequality measure (such as a Gini coefficient) satisfying symmetry, mean independence, population homogeneity, and the principle of transfers (ie., essentially inequality criteria (i) – (iv) above), if the Lorenz curve for distribution A lies everywhere above the Lorenz curve for distribution B, then all inequality measures satisfying these properties will indicate that (summary) inequality in A is less than in B. Note that this theorem does not say anything about social welfare; it refers only to inequality. It also does not say anything if the two Lorenz curves cross. Interestingly, comparing quantile curves and Lorenz curves can also serve as the basis for the measurement of first- and second-order earnings discrimination, such as between men and women (Le Breton et al., 2012).

To empirically implement this dominance criterion, one can again represent a Lorenz curve or cumulative income shares by a vector of its estimated ordinates. Testing between Lorenz curves then amounts to tests of differences between the estimated ordinate vectors. Again, if the two distributions whose inequality is being compared are designated A and B, then the vectors of Lorenz curve ordinates can be represented by

$\widehat{I^a}$ = ($\widehat{I^a}$ (1), ... $\widehat{I^a}$ (K-1))′, l^a= (l^a(1), ..., l^a(K-1))′

and $\widehat{I^b}$ = ($\widehat{I^b}$ (1), ..., $\widehat{I^b}$ (K-1))′, l^b = (l^b(1), ..., l^b(K-1))′

and their respective variance-covariance matrices by Φ^a and Φ^b. The ordinates l(1), ..., l(K-1) correspond to the given (cumulative) proportions p(1), ..., p(K-1). Since the two end points on a Lorenz curve are fixed at p(0) = 0 and p(K) = 1, only K -1 ordinates are random variables.

The actual decision rule or PEC for comparing the vectors of Lorenz curve ordinates again involves two steps. And again, it is assumed that the two sets of ordinate estimates are statistically independent and based on two quite separate samples.

Step 1 — Test the joint null hypothesis of equality of the two ordinate vectors (ie., $l^b - l^a$ = 0) versus the alternative hypothesis of non-equality. In this case, the test statistic is

$$(\widehat{I^b} - \widehat{I^a})' \, [\widehat{\Phi}^a + \widehat{\Phi}^b]^{-1} \, (\widehat{I^b} - \widehat{I^a}) \tag{7}$$

which is distributed asymptotically as a chi-square random variable with K - 1 degrees of freedom. If the null hypothesis is not rejected, then the two estimated Lorenz curves can be said to be not statistically significantly different, and further comparison is not pursued.

Step 2 — If, however, the null hypothesis in Step 1 is rejected, then undertake separate t-statistic calculations for differences on each of the individual estimated Lorenz curve ordinates. If at least one of the t-statistics has the appropriate sign and is statistically significant compared to critical values on the SMM distribution with K – 1 and infinite degrees of freedom and none of the t-statistics (if any) that has the wrong sign is statistically significant (again based on the SMM distribution), then one can conclude that one set of estimated ordinates statistically dominates the other. If statistical dominance is found, this implies dominance for all summary inequality measures satisfying inequality properties (i) – (iv). Again, typical useful SMM critical values are:

	alpha = .01	alpha = .05	alpha = .10
K – 1 = 4	3.430	2.631	2.378
K – 1 = 9	3.634	3.190	2.976
K – 1 = 19	4.018	3.615	3.425

Source: Stoline and Ury (1979), Tables 1-3.

This test procedure raises the issue of how to determine the statistical properties of the estimated Lorenz curve ordinates in order to make statistical inference decisions.

5.1 Inference for Lorenz Curve Ordinates

Recall that Lorenz curve ordinates are simply cumulative income shares (which have already been considered in Section 2 above). Let the K-vector of individual income share statistics be

$\hat{s}$ = ($\hat{s}$ (1), …, $\hat{s}$ (K))′

with corresponding population shares s = (s(1), …, s(K))′. Then it can be seen that

$\hat{I}$ = U. $\hat{s}$ (8)

where U is a (K-1)xK summation matrix with ones on its i-i elements and below, and zeros elsewhere. U is given and non-random. Since (8) is a linear transformation, if $\hat{s}$ is (asymptotically) joint normally distributed with mean vector s and asymptotic variance-covariance matrix Ws, then $\hat{I}$ is also (asymptotically) joint normally distributed with mean I = U.s and asymptotic variance-covariance matrix

Φs = U.Ws.U′

and the corresponding estimated regular variance-covariance matrix is given by

$\widehat{\Phi}$ = U. $\widehat{W}$.U′ (9)

where W is given by (4) and (5) with each element divided by N.

So to perform Step 1 of the PEC for comparing two vectors of estimated Lorenz curve ordinates, $\widehat{I^a}$ and $\widehat{I^b}$, first calculate estimates of all the asymptotic variances and covariances for the two estimation samples from equation (5) by replacing population parameters by their consistent sample estimates, rescale the (asymptotic) variance and covariance estimates to the estimated regular variances and covariances ($\widehat{W^a}$ and $\widehat{W^b}$), calculate the Lorenz curve ordinates by $\hat{I}$ = U. $\hat{s}$ from equation (8) and the regular estimated variances and covariance for the ordinates from

$\widehat{\Phi}$ = U. $\widehat{W}$.U′ ,

and then finally calculate the joint chi-squared test statistic in equation (7).

To perform the individual tests in Step 2 of the PEC, again use the standard t-statistic ratio for the difference between two independent random variates ($\widehat{I^a}$ (i) and $\widehat{I^b}$ (i)) as

t(i) = ($\widehat{I^b}$ (i) - $\widehat{I^a}$ (i)) / [$\widehat{\Phi}^a(i,i)$ + $\widehat{\Phi}^b$ $(\mathrm{i,i})]^{1/2}$

and compare this to the relevant critical value on the SMM tables.

As an aside, it can be noted that the Gini coefficient of overall income inequality can be approximated by (twice) the area between the estimated Lorenz curve ordinate segments and the 45 degree equality diagonal. If one represents each of these segments as a quadrilateral, one can calculate

$\hat{G} = \sum_{i=1}^{K}(1/K)$. [((i/K) – $\hat{I}$ (i)) + ((i-1/K) – $\hat{l}$ $(i-1)$)].

Since this is a linear function of the ordinates $\hat{I}$ (i), one can calculate the estimated variance of $\hat{G}$ as a simple quadratic form in the estimated variance-covariance matrix $\widehat{\Phi}$ of the Lorenz curve ordinates. Once again, the standard error of $\hat{G}$ is the square root of the estimated variance of $\hat{G}$, and once again the standard error is distribution-free.

6. Application to Generalized Lorenz Dominance

A blending of the first two dominance criteria is provided in a third application of empirically implementing curve-based dominance criteria. Shorrocks (1983) uses a transformed Lorenz curve as a basis for social welfare inferences, not just inequality conclusions. The generalized Lorenz dominance theorem of Shorrocks (1983) says that, for any additively separable social welfare function essentially satisfying social welfare conditions (i) - (iv) including the principle of transfers, distribution A is socially preferred to distribution B if the generalized Lorenz curve for A lies everywhere above the generalized Lorenz curve for B. The generalized Lorenz curve ordinates for an income distribution are obtained by scaling up the Lorenz curve ordinates of the distribution by the distribution's overall mean income level:

$$g(i) = \mu \, . \, l(i) \qquad \text{and } \hat{g}(i) = \hat{\mu} \, . \, \hat{l}(i). \tag{10}$$

Essentially, the argument is that, if the mean income of the distribution A is sufficiently higher than that in distribution B, this can compensate for some greater degree of inequality in A than in B, so that social welfare will still be greater in distribution A than in B. It turns out that this rule is very convenient for ranking social welfare among quite disparate countries, or for ranking income distributions in a given country (or group) over long periods of time (eg., the Canadian income distribution across the decades of 1950, 1960, 1970, and 1980, say). The use of generalized Lorenz curves for dominance comparisons has also been applied to earnings discrimination analysis over an entire distribution (Jenkins, 1994: del Rio et al., 2011; and Salas et al., 2018).

To implement this dominance criterion, one can again represent a generalized Lorenz curve by a vector of its estimated ordinates:

$$g = (g(1), \ldots, g(K\text{-}1))' \qquad \text{and } \hat{g} = (\hat{g}(1), \ldots, \hat{g}\,(K\text{-}1))'.$$

Testing between generalized Lorenz curves then amounts to tests of differences between the estimated ordinate vectors $\hat{g}^a$ and $\hat{g}^b$. The respective generalized Lorenz curve estimated ordinate variance-covariance matrices may be labelled $\widehat{\Psi}^a$ and $\widehat{\Psi}^b$.

The corresponding decision rule or PEC for comparing vectors $\hat{g}^a$ and $\hat{g}^b$ once again involves two steps where, as above, the estimation samples are statistically independent.

Step 1 — Test the joint null hypothesis of equality of two estimated generalized Lorenz curve ordinate vectors (ie., $g^b - g^a = 0$) versus the alternative hypothesis of non-equality. In this case, the test statistic is

$$(\hat{g}^b - \hat{g}^a)' \, [\hat{\Psi}^a + \hat{\Psi}^b]^{-1} \, (\hat{g}^b - \hat{g}^a) \tag{11}$$

which, under the null hypothesis, is asymptotically distributed as a chi-square random variable with K-1 degrees of freedom. If the null hypothesis is not rejected, then the two generalized Lorenz curves can be said to be not statistically significantly different, and further comparison is not warranted.

Step 2 — If, however, the null hypothesis in Step 1 is rejected, then proceed to compute separate t-statistic ratios for differences on each of the respective individual generalized Lorenz curve ordinates. If at least one of the t-statistics has the appropriate sign and is statistically significant compared to critical values on the SMM distribution with K-1 and infinite degrees of freedom and none of the t-statistics (if any) that have the wrong sign is statistically significant (again on the SMM critical values), then one can conclude that the distribution with the higher sample generalized Lorenz curve ordinates dominates (or is socially preferred to) the distribution with the corresponding curve with lower ordinates. If not, then one can say only that the social welfare of the two distributions are statistically significantly different, but not reach a preferred or dominance social welfare conclusion.

6.1 Inference for Generalized Lorenz Curve Ordinates

Since Lorenz curve ordinates are calculated from income shares, it makes sense to consider the relationship of generalized Lorenz curve ordinates to these underlying income shares as well. To go back to first principles, consider μ.IS(i) as the dollar contribution of the ith income group to the overall mean income of the distribution. So we can represent it by the "contribution"

$$c(i) = \mu \,.\, IS(i)$$

$$= D(i) \,.\, \mu(i) \qquad (12)$$

where $\mu(i)$ is the quantile mean of the ith income group and D(i) = p(i) - p(i-1). So c(i) is simply a scalar transform of the quantile mean $\mu(i)$ and $\hat{c}(i) = D(i) \,.\, \hat{\mu}(i)$. Or more generally, c = D . μ and $\hat{c}$ = D . $\hat{\mu}$ where D is a (KxK) diagonal matrix with elements D(i) on its principal diagonal. Thus the set of $\hat{c}(i)$'s are also asymptotically joint normally distributed and the elements of the asymptotic variance-covariance matrix of the vector of sample contributions $\hat{c} = (\hat{c}\,(1), \ldots, \hat{c}\,(K))'$ are simply scalar transforms of the corresponding elements of the asymptotic variance-covariance matrix of the vector of quantile means $\hat{\mu} = (\hat{\mu}\,(1), \ldots, \hat{\mu}\,(K))'$ given by equations (1) and (3) above. That is, if $\hat{V}$ is the estimated variance-covariance matrix of vector $\hat{\mu}$, then the estimated variance-covariance matrix of $\hat{c}$ is

$$\hat{T} = D.\, \hat{V}.\, D. \qquad (13)$$

Note, interestingly, that the economic contribution term c(i) here is exactly the statistical term n(i) in section 2.1 above. Thus the asymptotic variances and covariances for the $\hat{n}(i)$'s in equations (1) and (2) are exactly those for the $\hat{c}(i)$'s as well.

The ordinates of the generalized Lorenz curve can be readily obtained from the $\hat{c}(i)$'s by straightforward cumulation:

$\hat{g}$ (i) = $\sum_{j=1}^{i} \hat{c}\,(j)$ and g(i) = $\sum_{j=1}^{i} c(j)$

or more generally,

$\hat{g}$ = U . $\hat{c}$ and g = U . c

where again U is a (K-1)xK non-random matrix with ones on its principal diagonal and below, and zeros above the diagonal. Again, since the $\hat{g}$(i)'s are linear functions of the $\hat{c}$(i)'s, the $\hat{g}$(i)'s are also asymptotically joint normally distributed with means g(i)'s and a regular variance-covariance matrix that can be consistently estimated by

$$\hat{\Psi} = \text{U}.\, \hat{\Gamma}.\, \text{U}' \tag{14}$$

where the elements in $\hat{\Gamma}$ are given by equation (13). Once again, all terms in Γ and Ψ are distribution-free, and thus can be readily estimated consistently and directly.

To perform Step 1 of the PEC for comparing two vectors of generalized Lorenz curve ordinates $\hat{g}^a$ and $\hat{g}^b$, calculate $\hat{g}^a$ and $\hat{g}^b$ and all the estimated variances and covariances in $\hat{\Gamma}^a$ and $\hat{\Gamma}^b$ by (i) calculating the results in equations (1) and (2), and (ii) rescaling these equation results by their respective sample sizes, then compute $\hat{\Psi}^a$ and $\hat{\Psi}^b$ from equation (14), and finally calculate the joint chi-squared test statistic in equation (11).

To perform the individual tests in Step 2 of the PEC, again use the standard t-statistic ratio for the difference between two independent variates ($\hat{g}^a$(i) and $\hat{g}^b$(i)) as

$$\text{t(i)} = (\hat{g}^b(\text{i}) - \hat{g}^a(\text{i})) \,/\, [\hat{\Gamma}^a(i,i) + \hat{\Gamma}^b(i,i)]^{1/2}$$

and compare this to the relevant critical value on the SMM distribution.

7. Inequality and Crossing Lorenz Curves

7.1 Single Lorenz Curve Crossing

What can one infer if Lorenz curves cross? This is not uncommon when comparing two empirical curves, and the above Lorenz dominance criterion is of no help in such situations. However, Shorrocks and Foster (1987) extend the latter criterion to cases of single crossings of Lorenz curves. If the Lorenz curve for distribution A crosses the Lorenz curve for distribution B once from above, then all inequality measures satisfying the inequality properties (i) - (iv) plus property (v) — transfer sensitivity — will indicate that summary inequality in A is less than that in B if the coefficient of variation for distribution A is lower than that for distribution B. The coefficient of variation is the ratio of the standard deviation of the distribution to the mean; ie., $\hat{C} = \hat{\sigma} / \hat{\mu}$ in the estimation sample. Thus, by adding the one further property of transfer sensitivity, one can get a stronger practical result that helps rank aggregate income inequality across distributions even when their Lorenz curves cross (once). Again, this provides a potential ranking of overall income inequality between distributions, and not of social welfare more generally.

Implementing this stronger dominance rule is also feasible in light of the approach in this paper. All it requires is some revision of the Lorenz dominance PEC of Section 5.

A transfer sensitive Lorenz dominance PEC criterion can now be stated as follows:

Step 1 — Same as before. Test the joint null hypothesis of equality of the two Lorenz curve ordinate vectors (ie., $l^b - l^a = 0$) versus the alternative hypothesis of non-equality. As before, the test statistic is

$(\hat{I}^b - \hat{I}^a)' \cdot [\hat{\Phi}^a + \hat{\Phi}^b]^{-1} \cdot (\hat{I}^b - \hat{I}^a)$.

If the null hypothesis is not rejected, the two Lorenz curves can be said to be not statistically significantly different, and further comparison is not pursued.

Step 2 — If the null hypothesis in Step 1 is rejected and there is a single crossing of Lorenz curve ordinates, then undertake separate t-statistic calculations for differences in each of the respective individual estimated Lorenz curve ordinates up to the cross-over point. If at least one of the t-statistics below the cross-over point is statistically significant compared to critical values on the SMM distribution with K-1 and infinite degrees of freedom, one can conclude that the distribution with the initially higher ordinates initially Lorenz dominates the lower Lorenz curve distribution. If no such t-statistic is statistically significant, further comparison is not pursued.

Step 3 — If one distribution indeed initially Lorenz dominates the other, then compare the estimated coefficients of variation of the two distributions. If the coefficient of variation for the distribution with the initially higher Lorenz curve ordinates (say $\widehat{C^a}$) is statistically significantly smaller than the coefficient of variation for the other distribution ($\widehat{C^b}$) on the basis of a one-tailed asymptotically normal "t-test" (ie., Ho: $C^a - C^b = 0$ vs H1: $C^a - C^b < 0$), then one can conclude that distribution A transfer-sensitive Lorenz dominates distribution B. That is, summary inequality in distribution A is statistically significantly smaller than in distribution B for all inequality measures satisfying inequality properties (i) - (v). If $\widehat{C^a}$ is not statistically significantly smaller than $\widehat{C^b}$, do not draw any inequality dominance inference.

In order to perform the above test, one can make use of Ahn and Fessler (2003) who show that $\hat{C}$ from an i.i.d sample is asymptotically normally distributed with a standard error given by

$$\text{S.E.}(\hat{C}) = 100\ \hat{C}\ .\ [1 + 2\hat{C}^2 / 2\text{N}]^{1/2}$$

where $\hat{C}$ is expressed as a proportion. Thus for independent samples, the estimated variance of $\widehat{C^a}$ - $\widehat{C^b}$ is

$$\widehat{Var}\ (\widehat{C^a} - \widehat{C^b}) = [\text{S.E.}\ (\widehat{C^a})]^2 + [\text{S.E.}\ (\widehat{C^b})]^2,$$

and the (asymptotic) normal t-ratio test statistic is

$$t = (\widehat{C^a} - \widehat{C^b})\ /\ [\widehat{Var}\ (\widehat{C^a} - \widehat{C^b})]^{1/2}.$$

7.2 Multiple Lorenz Curve Crossings

But what of the situation where two Lorenz curves cross more than once? The transfer-sensitivity-based approach of Shorrocks and Foster (1987) has been extended by Davies and Hoy (1994) to address just this situation. They posit a coefficient of variation condition for each cross-over point (including the top right-hand (1,1) point on the Lorenz curve).

More specifically, Davies and Hoy (1994, 1995) argue that a sufficient condition for all summary measures of inequality satisfying inequality properties (i) - (v) — ie., including transfer sensitivity — to show a reduction in inequality is that, at all cross-over points k = 1, 2, ... of the Lorenz curves for two distributions, the cumulative coefficients of variation at points k are smaller for the distribution with the initially higher Lorenz curve[2]. The cumulative coefficient of variation (CC) is the ratio of the standard deviation over the mean where both are calculated over all observations in the distribution with incomes less than or equal to the cross-over income level at point k[3].

To empirically implement a test for this situation, one follows a similar set of PEC steps as for the single-crossing situation, but with some refinement.

Step 1 — Same as before. Do a chi-square test of the joint null hypothesis of equality of the two Lorenz curve ordinate vectors versus the alternative hypothesis of non-equality. If the null hypothesis is not rejected, the two Lorenz curves can be said to be not statistically significantly different, and further comparison is not pursued.

Step 2 — Similar to the single-crossing case. If the null hypothesis in Step 1 is rejected and there are one or more crossings of the Lorenz curve ordinates, then undertake separate t-statistic calculations for differences in each of the respective individual estimated Lorenz ordinates up to the lowest or first cross-over point. If at least one of the t-statistics below this first cross-over point is statistically significant compared to critical values on the SMM distribution with K-1 and infinite degrees of freedom, one can conclude that the distribution with the initially higher ordinates initially Lorenz dominates the lower Lorenz curve distribution. If no such t-statistic is statistically significant, further comparison is not pursued.

[2] For a fuller explication of the necessary and sufficient conditions for this multi-crossing point criterion, see Chiu (2007) and Davies, Hoy and Lin (2022).

[3] Note that this is a stronger sufficient condition than in Shorrocks and Foster (1987) in that a single-crossing test now involves two coefficient of variation tests rather than one.

Step 3 — Here is where the PEC for crossing Lorenz curves changes. Now compare the cumulative coefficients of variation ($\widehat{CC}$ (i)) for the two distributions (A and B) for all quantile points i = 1, …, K (note that $\widehat{CC}$(K) = $\hat{C}$) and undertake separate "t-statistic" calculations for differences of each of the individual estimated quantile cumulative coefficients of variation ($\widehat{CC}^a$ (i) – $\widehat{CC}^b$ (i)). If at each cross-over point, the immediately following quantile $\widehat{CC}$(i) is statistically significantly smaller for the initially higher Lorenz (as well as for the full-sample $\hat{C}$) on the basis of the SMM distribution with K and infinite degrees of freedom, then one can conclude that inequality in the distribution with the initially higher Lorenz curve (distribution A, say), statistically significantly dominates inequality in distribution B for all summary inequality measures satisfying inequality properties (i) – (v) — including transfer sensitivity. If one or more of these individual t-statistic tests is not one-tailed statistically significant, do not draw any inequality dominance inferences.

To implement Step 3 of this PEC involves doing statistical inference on the $\widehat{CC}_i$'s. To do so, one can make use of results in Beach, Davidson and Slotsve (1994) who establish the asymptotic joint normal distribution of the full set of $\widehat{CC}_i$'s and derive their full asymptotic variance-covariance structure when calculated from a random sample. More specifically, they show that, for i=1, …,K, the asymptotic variance of $\widehat{CC}_i$ is given by

$$Asy.var\left(\widehat{CC}_i\right) = \left(\frac{\phi_i}{p_i\,\lambda_i\,\gamma_i{}^2}\right)^2 \cdot\ p_i\left[\phi_i - \gamma_i^2 + (1-p_i)(\xi_i - \gamma_i)^2\right]$$

$$-2\left(\frac{\phi_i}{p_i\,\lambda_i\,\gamma_i{}^2}\right)\left(\frac{1}{2p_i\,\lambda_i\,\gamma_i}\right)\cdot\ p_i\left[\chi_i - \gamma_i\phi_i + (1-p_i)(\xi_i - \gamma_i)(\xi_i^2 - \phi_i)\right]$$

$$+\left(\frac{1}{2p_i\,\lambda_i\,\gamma_i}\right)^2\cdot\ p_i\left[\psi_i - \phi_i^2 + (1-p_i)\left(\xi_i^2 - \phi_i\right)^2\right]$$

where ξ_i is the i'th quantile cut-off income level,

$\gamma_i = E(Y \mid Y \leq \xi_i)$ is the cumulative mean,

$\lambda_i^2 = E[(Y-\gamma_i)^2 \mid Y \leq \xi_i]$ is the cumulative variance,

$\phi_i = \lambda_i^2 + {\gamma_i}^2 = E(Y^2 | Y \leq \xi_i)$ is the cumulative second moment,

$\chi_i = E(Y^3 | Y \leq \xi_i)$ is the cumulative third moment, and

$\psi_i = E(Y^4 | Y \leq \xi_i)$ is the cumulative fourth moment.

All of these terms can be estimated consistently in straightforward fashion (that is again distribution-free) by their sample analogues to yield

$As\hat{y}.var\,(\widehat{CC}_i)$. Thus

$$S.E.(\widehat{CC}_i) = \left[\frac{Asy.\hat{v}ar\,(\widehat{CC}_i)}{N}\right]^{1/2}.$$

Again, for independent samples, the estimated variance of $\widehat{CC}_i^a - \widehat{CC}_i^b$ is

$$\hat{V}ar\left(\widehat{CC}_i^a - \widehat{CC}_i^b\right) = \left[\frac{As\widehat{y.v}ar\left(\widehat{CC}_i^a\right)}{N^a}\right] + \left[\frac{As\widehat{y.v}ar\left(\widehat{CC}_i^b\right)}{N^b}\right]$$

and the (asymptotic) normal t-ratio statistic is

$$t = \frac{\widehat{CC}_i^a - \widehat{CC}_i^b}{\left[\widehat{Var}\left(\widehat{CC}_i^a - \widehat{CC}_i^b\right)\right]^{1/2}}.$$

It should be noted that Stage 3 tests take place over the pre-specified quantile points and not at estimated cross-over points. This is done for convenience because of the known and distribution-free (asymptotic) variance-covariance structure of the quantile $\widehat{CC}$ (i) statistics. In the authors' experience, one cross-over point is quite common, two such points have been observed on rare occasions, but more than two have never been encountered when using decile or vigintile intervals. We would thus argue that, for the purpose of inequality rankings for income distributions as a whole, decile breakdowns are quite sufficient to pick up relevant Lorenz curve cross-overs and few, if any, $\widehat{CC}$ (i) difference tests are required.

Note also that the above PEC empirical implementation criteria are only sufficient and specified to be fairly strong in order to yield pretty convincing dominance conclusions when they are indeed satisfied. Certainly, alternative weaker PEC criteria could also be specified for Steps 2 and 3 that are more likely to lead to more — but perhaps less convincing — dominance conclusions.

Note further that quantile mean dominance, Lorenz curve dominance and transfer sensitivity dominance essentially correspond to first-order, second-order and third-order stochastic dominance of one distribution by another. This analytical framework has generated a huge literature, especially in the finance field where testing for stochastic dominance efficiency is applied over a class of portfolio returns and investor preference toward risk (Post, 2003; Scaillet and Topaloglou, 2010; and Linton et al., 2014). Application of the stochastic dominance framework to income distributions and income inequality comparisons are found in Anderson (1996) based on Pearson goodness-of-fit tests and in Maasoumi and Heshmati (2000), Barrett and Donald (2003), and Linton et al. (2005) based on forms of Kolmogorov-Smirnov goodness-of-fit tests and on resampling/simulation procedures and bootstrapping methods. The stochastic dominance approach also underlies Davidson and Duclos' (2000) application of statistical inference to poverty rankings along with welfare and inequality comparisons in a unified analytical framework. While they rely on bootstrapping for inference results, they note that use of (asymptotic) variance information is known to improve inference results (ie., make tests more powerful). Applications of their test procedures are found in Duclos and Araar (2006).

8. Empirical Application: Basic Distributional Statistics

We now illustrate the application of the various tests of this paper with Canadian census data over 2000-2020. We consider four censuses for the years 2001, 2006, 2016, and 2021 —— in the 2011 Census, Statistics Canada changed their methodology, so its data are not completely comparable to the other years' data; consequently, the results for the 2011 Census are not included. We make use of the Statistics Canada Public Use Microdata Files (PUMFs) on Individuals and consider the total income of individuals, separately for women and men. Since income refers to that reported for the previous full calendar year, the income years are 2000, 2005, 2015, and 2020. Since we are testing for changes in economic well-being, families or household may be considered the more natural record unit. But there is not sufficient consensus in the literature on how a family unit is defined (the PUMF files include three different definitions) or on how best to adjust for family size and composition (say, on an adult-equivalent scale). In the name of simplicity in the application of the various tests of this paper and to focus on the tests themselves, we opt for total income of individual income recipients. Since the labour market activity patterns and experiences have considerable differences between men and women, we treat these two groups separately. We thus consider changes in income inequality and economic well-being within these two groups of income recipients in Canada over the period of 2000-2020.

The estimation samples for the study consist of those records on the PUMF files for individuals age 18 or over who did not attend school (either full-time or part-time) in the income year and whose total income that year was at least $1000. Total income consists of wages and salaries, net self-employment income, investment income, retirement pensions, and other money income (e.g., disability benefits). Summary statistics on mean income and sample size of each of the estimation samples are provided in appendix Table A1. Incomes are in real 2020 Canadian dollars (adjusted by Statistics Canada's annual CPI series, Table 18-10-0005-01). As can be seen, all the sample sizes are quite large and vary between 254,607 and 345,002 observations.

Tables 1a and 1b present decile mean incomes for women (1a) and men (1b) for each of the four census years. Figures in parentheses are standard errors. Given the large sample sizes, it is not surprising that these decile means are all highly statistically significant. From simple inspection, it can be seen that the real decile mean income levels generally increase over time with typically weaker increases (or even some declines) over 2000-05 and much stronger increases over 2015-20 (except at the top end of the distributions). As expected, men's decile means are considerably higher than women's.

Decile income shares (expressed as percentages) over the four censuses appear in Table 2a (for women) and Table 2b (for men), again with standard errors in parentheses. And again, the large sample sizes ensure highly statistically significant results. One notes that in 2005, 2015 and 2020 the lower nine decile shares are generally higher for women than for men, while the top share is distinctly higher for men than for women. This would suggest that, at the simple level of inspection, income inequality is more marked in the distribution of men's incomes than for women's incomes.

As to distributional changes over time, two periods are noticeable. First, between 2000 and 2005 for both women and men, the lower nine decile shares fell, while the top decile share rose quite dramatically, suggesting a marked rise in income inequality over this period. The changes were more marked for men than for women. Second, between 2015 and 2020 and again for both men and women, the lower eight or nine decile shares rose quite substantially, while the top decile shares fell quite markedly, in turn suggesting a notable reduction in income inequality over this period. As found by Beach (2016), workers' earnings — by far the largest component of total income — over the 2000-05 period were continuing their longer-run dramatic widening of earnings differentials and particularly the growing gap between middle and upper earnings levels. These changes were apparently driven by the major economic forces of automation and technological change; globalization, off-shoring, and changing international trade patterns; and deregulation and growing industrial concentration. The apparent strong reduction in income inequality over the 2015-20 period likely reflects the large federal government payouts of income support payments in the face of the 2020-21 COVID pandemic.

Cumulative income shares or Lorenz curve ordinates for the four census years are presented in Tables 3a (for women) and 3b (for men). In comparison between men and women, only in 2005 and 2015 are the Lorenz curves for women uniformly above those for men suggesting greater income inequality within the men's income distribution. For 2000 and 2020, the results are not so clear-cut. Comparing Lorenz curves across time, one sees that over 2000-05, Lorenz curves for both men's and women's incomes indeed shifted consistently down or out, more strongly so for men. In contrast, over 2015-20, the Lorenz curves again for both men and women shifted consistently up and quite strongly so for both groups of income recipients. Over the longer 2005-15 period, both Lorenz curves continued their earlier downward shift, but only weakly so. These results suggest a strong increase in overall income inequality between 2000 and 2005, followed by a continuing but weaker inequality increase through to 2015, but then a quite sharp reversal and marked reduction in inequality between 2015 and 2020. Whether this reversal will hold after the COVID government transfer payouts stop remains to be seen.

Results for generalized Lorenz curves (GLCs) appear in Tables 4a and 4b (again standard errors are in parentheses). For women's income distributions, the GLCs rise consistently across all four census years suggesting an on-going improvement in overall economic well-being. However, this is not the case for the men's income distributions. Between 2000 and 2020 as a whole, all men's GLC ordinates rose. But this did not occur over all subperiods. Between 2005 and 2015, men's GLCs rose over all deciles. But between 2000 and 2005, men's GLC ordinates fell over the lowest six deciles, changed very little over deciles 7-9, and rose strongly for the top ten percent of the distribution. Between 2015 and 2020, however, men's GLC ordinates rose strongly over the lower nine deciles, but fell for the top decile. Thus gains and losses were not uniform over the first two decades of this century, and conclusions as to changes in men's overall economic well-being are not straightforward.

9. Empirical Application: Testing for Distributional Changes

9.1 Testing for Rank Dominance

The first step in the practical empirical criterion (PEC) of Section 4 for rank or quantile mean dominance is a chi-square test of equality of the full set of decile mean vectors (with K=10 degrees of freedom). The resulting chi-square statistics are as follows:

	2000-05	2005-15	2015-20	2000-20
Women	2922.0	6481.2	8159.9	33,400.
Men	2483.9	1456.9	4570.6	7417.5

With a 99 percent confidence critical value of 23.2, one can quite definitely reject the null hypothesis of no change in the set of quantile means in all eight cases considered.

Tables 5a and 5b present the actual changes in the estimated individual decile means over four time intervals: 2000-05, 2005-15, 2015-20 and for the full period 2000-20. In the tables of this section, however, the figures in parentheses are now (absolute values of) asymptotic t-ratios of the decile mean differences. In the case of women in Table 5a, all the decile mean changes are highly statistically significant on the basis of the SMM distribution critical values of Section 4 — including the negative difference for the tenth decile between 2015 and 2020. Thus one can conclude that women experienced rank dominance improvements in their incomes and thus economic well-being over 2000-05, 2005-15, and over the full period 2000-20 under quite weak social welfare conditions (i) – (iii). But one cannot reach this conclusion over the 2015-20 period because of the one statistically significant loss of incomes for the one top decile group.

In the case of men in Table 5b, there are evidently some not statistically significant changes in decile means and indeed some highly significant negative changes. One can conclude that men have experienced statistically significant improvements in economic well-being over the ten-year interval 2005-15 and over the full twenty-year period 2000-20. But over the 2000-05 period and over the 2015-20 period there are statistically significant winners and losers. So we cannot reach a conclusion of whether economic well-being improved as a whole for male income recipients over these two periods under such weak social welfare conditions.

9.2 Testing for Lorenz Dominance

The first step in testing for Lorenz dominance involves testing for whether the vectors of Lorenz curve ordinates are jointly statistically significantly different (the test is also equivalent to testing for differences in the set of income shares). The estimated chi-square statistics are as follows:

	2000-05	2005-15	2015-20	2000-20
Women	2289.7	237.85	4009.8	4778.7
Men	2989.9	321.79	3976.6	4066.7

Again, these are all highly statistically significant, indicating that the Lorenz curves have indeed changed significantly.

Changes in the individual Lorenz curve ordinates (expressed as percentage points) are presented in Tables 6a (for women) and 6b (for men). Again, the figures in parentheses are (absolute values of) t-ratios for the decile changes. Negative values for the ordinate changes indicate the Lorenz curve is shifting down or out, while positive values correspond to a Lorenz curve moving up or inward towards the 45 degree equality diagonal in a Lorenz curve diagram. Again, almost all the estimated ordinate changes are individually highly statistically significant on the basis of the SMM distribution — see the critical values in Section 5 above.

For both women and men, there are seen to be consistent statistically significant downward shifts in their respective Lorenz curves over both 2000-05 and 2005-2015, followed by a consistent statistically significant upward shift over 2015-20. One can thus conclude that, for both women and men, overall income inequality (under inequality conditions $(i) - (iv)$) rose or worsened over periods 2000-05 and 2005-15, but then declined or improved over the following 2015-20 period. However, again for both women and men, over the 2000-20 period as a whole — see the right-hand column of results in Tables 6a and 6b — there are statistically significant positive and negative changes across the decile ordinates. Thus one cannot say conclusively whether overall income inequality improved or worsened for the period as a whole (on the basis of inequality conditions $(i) - (iv)$), only that it has indeed statistically significantly changed.

9.3 Testing for Generalized Lorenz Dominance

As with the above empirical testing criteria, the first stop involves a chi-square test of equality of two vectors of generalized Lorenz curve (GLC) ordinates. The corresponding test statistics for generalized Lorenz dominance are as follows:

	2000-05	2005-15	2015-20	2000-20
Women	29.214	64.889	81.591	334.02
Men	24.847	14.554	45.707	74.184

With a 99 percent confidence level critical value of 23.2 (and 18.3 for the 95 percent confidence level) with 10 degrees of freedom, all but one are highly statistically significant, thus allowing one to infer that overall economic well-being indeed changed over the corresponding periods. For men between 2005 and 2015, the p-value of the above test statistic is p=0.15, so that at conventional levels of confidence one cannot infer statistically significant change in economic well-being, though the relatively low p-value still suggests change likely did occur.

Tables 7a and 7b provide changes in GLC ordinates over the four periods of interest, again separately for women and men. Positive values indicate an upward shift in the generalized Lorenz curve, and negative values show a downward shift of the curve. And again, the reported t-ratios in parentheses are to be compared to SMM distribution critical values in Section 4. So, for example (with 10 and infinite degrees of freedom), the 99 percent confidence level value is 3.691 and the 95 percent value is 3.254.

Over the 2000-20 period as a whole, the GLC ordinate changes for both men and women are consistently positive and all individually highly statistically significant. So one can conclude that, for both sex groups, there was indeed an improvement in their overall economic well-being based on social welfare conditions $(i) - (iv)$ — including the principle of transfers, but not transfer sensitivity. Over the 2005-15 middle time period, a similar conclusion holds for women's economic well-being. The very weak "t-statistics" for men back up the finding of a non-significant chi-square statistic over this period.

Over the 2015-20 period, women's GLC ordinates also show consistent and highly statistically significant upward shifts, so again one can infer an increase in their general economic well-being. For men, all but one of the GLC changes are positive and the lower five ordinate changes are statistically significant, while the one negative change (at the top end) is highly non-significant. So men's overall economic well-being can also be inferred to have increased.

Finally, over the 2000-05 time interval, women's GLC ordinate changes are also all consistently positive with one — the top — statistically significant. So one can again infer an improvement in their economic well-being. Men's GLC ordinate changes over this period show a mix of positive (at the upper end of the distribution) and negative (at the lower and mid regions of the distribution) values. However, the top value is again statistically significant at conventional confidence levels while none of the negative changes is remotely significant. So once again, one can infer an improvement in men's general economic well-being over this period according to Section 6's PEC.

9.4 Testing for Transfer-Sensitive Lorenz Dominance

As has been seen, estimated changes in regular Lorenz curve ordinates over the full 2000-20 time period could not be conclusively evaluated on the basis of income inequality conditions $(i) - (iv)$. Imposing the additional stronger condition of transfer sensitivity, however, may allow for more conclusive results.

The first step in Section 7's PEC again involves general chi-square tests of no change in the vectors of Lorenz curve ordinates. These have already been calculated in section 9.2 above indicating highly statistically significant Lorenz curve changes in all cases. Step 2 results on changes in the individual Lorenz curve ordinates yielded conclusive inferences on changes in overall income inequality over the three time intervals 2000-05, 2005-15 and 2015-20. Only for the 2000-20 period as a whole were the results inconclusive. Consequently, the discussion here focuses just on the latter period.

Tables 8a and 8b present the cumulative coefficients of variation for the full sets of Lorenz curve ordinates for 2000 and 2020 along with their differences. Since there is only one crossing in the Lorenz curves for women and for men, we need to examine the tenth decile or unconditional coefficient of variation difference in the bottom right corner of the tables. If this figure is statistically significantly negative (based on a one-tailed test of the SMM distribution (with 10 and infinite degrees of freedom), one could infer a significant reduction in overall income inequality over this period. But for both women and men, the figure turns out positive, so cannot possibly lie in the rejection region of the alternative one-tailed hypothesis (and indeed the $\hat{C}$ differences are not at all significant with t-ratios on the unconditional difference of 0.365 for women and 1.127 for men). So even with the stronger condition (v) of transfer sensitivity, one still cannot reach a conclusion of Lorenz dominance for an overall reduction in inequality between 2000 and 2020 for either women or men.

Further examination of the PUMF documentation, however, leads to some concerns as to the general usefulness of the Lorenz dominance tests based on transfer sensitivity. Between 2001 and 2006, Statistics Canada made some major changes to their top-coding procedures for total income (and other major income sources) on the PUMF Individual files. The effect was to raise the nominal values of the top-coded figures reported on the PUMF quite considerably (as well as vary these by sex and geographic area). Subsequent refinements followed in later censuses (see "2021 Census Public Use Microdata File (PUMF) : Individuals File — Documentation and User Guide," Statistics Canada's cat. no. 98M 0001X (2024), pp. 155-158, and the corresponding Statistics Canada documentation for the earlier census PUMF files). Even though only about one percent of income recipients are affected, these changes have led to substantial jumps in the estimates of the overall or unconditional coefficient of variation ($\hat{C}$) from census to census:

	2000	2005	2015	2000
Women	0.79793	1.02708	1.05356	0.88389
Men	0.80861	1.39846	1.43448	1.17690

Such major changes in such a robust summary statistic from our census samples — even more remarkably for men — are quite unusual in such large samples only a few years apart.

It would thus appear that applying transfer-sensitive Lorenz dominance tests to public use microdata files (that contain top-coding) is problematic. One alternative could perhaps be to work with the (confidential) census Master Files for this test.

10. Review and Conclusions

The theoretical literature on social choice and economic welfare evaluations for some while has offered several dominance criteria for ranking key aspects of income distributions — such as rank dominance or Lorenz dominance — based on comparing curves such as quantile curves or Lorenz curves. But empirical application of these criteria has proved challenging because (i) they involve restrictive assumptions about the form of the distributions being compared or the computational burden of density estimation from available microdata, and (ii) the applications are typically not accompanied by methods of statistical inference to indicate how reliable their results are. This paper provides the statistical tools and procedures for actually implementing these dominance criteria empirically with microdata sets that can be readily obtained from statistical agencies such as Statistics Canada and the U.S. Bureau of the Census.

More specifically, this paper builds on recent statistical developments in Beach and Davidson (2024a) establishing distribution-free statistical inference methods for quantile means and income shares for a sample distribution, and applies their variance – covariance formulas in a set of practical empirical procedures for formally testing general social welfare and income inequality dominance comparisons between distributions within a conventional statistical inference framework. Four sets of dominance rules are examined: rank dominance, Lorenz curve dominance, generalized Lorenz curve dominance, and transfer-sensitive Lorenz dominance — the first and third refer to social welfare or economic well-being comparisons, the second and fourth refer to income inequality comparisons. The paper thus offers an empirical approach to evaluate distributional changes in an easily implementable distribution-free framework of statistical inference based on well-known disaggregative distributional statistics.

The process for implementing the developments of this paper involve three stages. The first stage consists of representing a dominance curve by a vector of the curve's estimated ordinates — in terms of either quantile means or income shares — for a set of specified quantile points (such as deciles or percentiles). This transforms a theoretical problem into a statistical one. The second stage involves establishing the statistical properties of this vector of sample ordinates through use of recent developments on quantile-based inferences that are distribution-free and yield variance-covariance formulas that are very straightforward to implement directly. This transform the statistical problem into an inferential one by providing a framework for basing comparisons between sample vectors on formal statistical inference and testing procedures. The third stage involves proposing specific practical empirical criteria (or PECs) — one can think of these as a type of decision tree — for using formal statistical inference tests to reach empirical conclusions about the ranking (in terms of better or worse) or dominance of the key aspects of income distributions being compared.

The tests and procedures are illustrated with Canadian census public use microdata for women's and men's incomes between 2000 and 2020. There were five principal empirical findings. First, given the large sample sizes (between 254 and 345 thousand observations over eight census files) involved in the analysis, all the estimated decile means and income shares turn out to be highly statistically significant, thus providing lots of opportunity to obtain quite strong and convincing dominance conclusions.

Second, in terms of the formal inequality dominance tests, for both women and men, it is found that overall income inequality statistically significantly increased or worsened over both 2000-05 and 2005-15 and statistically significantly decreased or improved over 2015-20. Over the full 2000-20 period, however, for both women and men, statistically significant inferences could not be reached, apart from that significant changes had indeed occurred.

Third, in terms of formal tests of welfare dominance, again for both women and men, it is found that statistically significant improvement in overall economic well being occurred over all three subperiods as well as for the 2000-20 period as a whole.

Fourth, statistical inference-based conclusions and simple descriptive – or inspection-based conclusions are found to typically be pretty similar. In only one of the sixteen cases examined in this paper did the two differ — that for the comparison of generalized Lorenz curves for men over the 2000-05 period.

Fifth, transfer-sensitive Lorenz dominance results are found to be sensitive to top-coding procedures typically done by statistical agencies on their public use microdata files. Economic well-being dominance results, however, are surprisingly strong and robust and more successful in ranking distributions than inequality dominance results.

The analysis and findings in this paper have several implications. First, they show that quite broad inferences can be drawn as to inequality and social welfare that do not rely on single aggregate or summary measures and can be much more general. Thus there should be a shift of attention from various summary measures to sets of disaggregative measures such as quantile mean incomes and quantile income shares. The analysis of the paper shows that these disaggregative distributional statistics can provide not just descriptive information on changing distributional patterns, but also normative insights and the basis for formal statistical inference for evaluating distributional changes. Indeed, future advances in ranking criteria that involve comparisons of specified ordered-income curves can also, in principle, be implemented empirically in terms of the approach forwarded in this paper.

Second, the simple and straightforward test procedures for statistical dominance presented in the paper have broad applicability, such as in the empirical analysis of earnings, say, or wealth, or even more generally in terms of, say, test statistics. While all the examples provided in this paper refer to changes over time in the income distribution for a given population group, they could also be applied to comparisons of distributions between separate groups or even comparisons over time between groups (difference in differences).

Third, public debate on the objectives of economic policy should no longer focus on simply GDP or GDP per capita or on simply efficiency vs. equity. Discussion should now also include effects on empirically evaluated economic well-being — say in the form of a generalized Lorenz curve — that incorporates both efficiency and equity concerns, and that reflects rational distributional trade-offs between winners and losers and the reliability of judging these trade-offs.

Table 1a
Women's Decile Mean Incomes
(Canada, 2020 $)

Decile	2000	2005	2015	2020
1	4212	4667	5414	7733
	(23.81)	(28.41)	(28.68)	(39.10)
2	10,584	11,441	13,304	17,423
	(36.98)	(32.93)	(37.81)	(39.99)
3	15,701	16,482	18,991	23,568
	(35.92)	(40.55)	(39.93)	(35.21)
4	19,890	21,080	24,326	29,510
	(40.82)	(34.08)	(44.64)	(54.60)
5	24,847	26,430	31,046	36,309
	(54.00)	(58.65)	(64.97)	(57.34)
6	31,248	33,275	38,857	43,636
	(68.94)	(70.51)	(74.24)	(65.33)
7	39,150	41,082	47,493	52,028
	(74.14)	(67.68)	(71.19)	(70.23)
8	48,021	50,528	58,435	62,980
	(84.14)	(90.70)	(103.09)	(91.06)
9	61,366	65,082	76,066	80,464
	(116.62)	(127.12)	(133.23)	(128.70)
10	98,150	125,426	148,312	142,289
	(213.83)	(516.60)	(595.41)	(465.86)

Source: Authors' calculations from Statistics Canada's Public Use Microdata Files for Canadian Censuses.

Standard errors in parentheses.

Table 1b
Men's Decile Mean Incomes
(Canada, 2020 $)

Decile	2000	2005	2015	2020
1	7371	7205	7593	10,622
	(43.61)	(40.81)	(42.96)	(48.10)
2	17,634	17,356	18,068	21,657
	(53.56)	(57.34)	(45.57)	(39.85)
3	25,528	25,416	26,064	29,117
	(64.28)	(62.01)	(64.87)	(56.83)
4	33,979	34,012	35,363	37,499
	(79.33)	(79.62)	(80.85)	(70.95)
5	42,526	42,371	44,506	46,061
	(77.95)	(71.78)	(75.58)	(74.41)
6	51,293	51,436	54,348	55,566
	(94.19)	(95.19)	(94.25)	(82.66)
7	60,916	62,014	65,969	66,860
	(92.69)	(102.43)	(108.08)	(94.84)
8	73,462	75,425	81,039	81,352
	(120.27)	(123.01)	(136.91)	(120.67)
9	91,841	95,689	104,119	103,869
	(157.04)	(177.83)	(169.78)	(170.77)
10	158,080	231,264	255,668	221,073
	(467.66)	(1448.4)	(1505.8)	(1066.4)

Source: Authors' calculations from Statistics Canada's Public Use Microdata Files for Canadian Censuses.

Standard errors in parentheses.

Table 2a
Decile Income Shares for Women
(percentages)

Decile	2000	2005	2015	2020
1	1.1926	1.1801	1.1712	1.5593
	(0.006388)	(0.006957)	(0.006072)	(0.007575)
2	2.9968	2.8929	2.8781	3.5131
	(0.009271)	(0.008188)	(0.008057)	(0.007668)
3	4.4456	4.1675	4.1083	4.7522
	(0.008521)	(0.009831)	(0.008734)	(0.007261)
4	5.6319	5.3299	5.2627	5.9501
	(0.008939)	(0.009585)	(0.009779)	(0.009634)
5	7.0354	6.6828	6.7163	7.3214
	(0.010684)	(0.013033)	(0.01260)	(0.010037)
6	8.8479	8.4137	8.4060	8.7986
	(0.012515)	(0.014970)	(0.014087)	(0.011002)
7	11.0854	10.3873	10.2747	10.4907
	(0.012496)	(0.015709)	(0.014977)	(0.011864)
8	13.5971	12.7759	12.6415	12.6990
	(0.013596)	(0.018591)	(0.018082)	(0.013900)
9	17.3758	16.4559	16.4556	16.2243
	(0.018770)	(0.022993)	(0.021894)	(0.017707)
10	27.7914	31.7140	32.0857	28.6913
	(0.045172)	(0.084147)	(0.082610)	(0.063352)

Source: Authors' calculations from Statistics Canada's Public Use Microdata Files for Canadian Censuses.

Standard errors in parentheses.

Table 2b
Decile Income Shares for Men
(percentages)

Decile	2000	2005	2015	2020
1	1.3101 (0.007364)	1.1218 (0.006540)	1.0961 (0.006349)	1.5767 (0.007208)
2	3.1343 (0.008509)	2.7026 (0.009887)	2.6082 (0.008116)	3.2147 (0.007355)
3	4.5373 (0.009594)	3.9577 (0.011790)	3.7624 (0.011103)	4.3221 (0.009634)
4	6.0391 (0.011083)	5.2963 (0.014928)	5.1049 (0.013949)	5.5664 (0.011651)
5	7.5585 (0.010558)	6.5979 (0.016565)	6.4246 (0.015610)	6.8373 (0.012912)
6	9.1167 (0.011859)	8.0094 (0.019735)	7.8455 (0.018502)	8.2482 (0.014666)
7	10.8267 (0.011853)	9.6567 (0.022678)	9.5227 (0.021517)	9.9247 (0.016835)
8	13.0569 (0.013871)	11.7450 (0.026547)	11.6985 (0.025371)	12.0758 (0.019674)
9	16.3236 (0.016977)	14.9006 (0.031589)	15.0297 (0.030612)	15.4183 (0.023889)
10	28.0968 (0.055288)	36.0119 (0.138202)	36.9073 (0.131164)	32.8160 (0.101537)

Source: Authors' calculations from Statistics Canada's Public Use Microdata Files for Canadian Censuses.

Standard errors in parentheses.

Table 3a
Lorenz Curve Ordinates for Women's Income
(percentages)

Decile	2000	2005	2015	2020
1	1.1926	1.1801	1.1712	1.5593
	(0.006388)	(0.006957)	(0.006072)	(0.007575)
2	4.1894	4.0730	4.0493	5.0723
	(0.01443)	(0.01392)	(0.01299)	(0.01380)
3	8.6350	8.2405	8.1576	9.8246
	(0.02108)	(0.02197)	(0.02013)	(0.01928)
4	14.2670	13.5704	13.4202	15.7747
	(0.02735)	(0.02938)	(0.02796)	(0.02657)
5	21.3024	20.2532	20.1365	23.0961
	(0.03416)	(0.03940)	(0.03785)	(0.03398)
6	30.1503	28.6669	28.5425	31.8947
	(0.04104)	(0.05051)	(0.04858)	(0.04166)
7	41.2357	39.0542	38.8172	42.3854
	(0.04629)	(0.06190)	(0.05964)	(0.04938)
8	54.8328	51.8301	51.4587	55.0844
	(0.04862)	(0.07375)	(0.07123)	(0.05704)
9	72.2086	68.2860	67.9143	71.3087
	(0.04517)	(0.08415)	(0.08261)	(0.06335)
10	100.	100.	100.	100.

Source: Authors' calculations from Statistics Canada's Public Use Microdata Files for Canadian Censuses.

Standard errors in parentheses.

Table 3b
Lorenz Curve Ordinates for Men's Income
(percentages)

Decile	2000	2005	2015	2020
1	1.3101	1.1218	1.0961	1.5767
	(0.007364)	(0.006540)	(0.006349)	(0.007208)
2	4.4444	3.8244	3.7043	4.7914
	(0.01451)	(0.01521)	(0.01332)	(0.01322)
3	8.9817	7.7822	7.4667	9.1135
	(0.02202)	(0.02545)	(0.02306)	(0.02142)
4	15.0208	13.0785	12.5716	14.6799
	(0.03026)	(0.03857)	(0.03540)	(0.03141)
5	22.5792	19.6764	18.9962	21.5171
	(0.03740)	(0.05306)	(0.04919)	(0.04243)
6	31.6959	27.6858	26.8417	29.7653
	(0.04439)	(0.07035)	(0.06550)	(0.05486)
7	42.5226	37.3425	36.3644	39.6900
	(0.05001)	(0.09015)	(0.08432)	(0.06894)
8	55.5795	49.0875	48.0629	51.7658
	(0.05457)	(0.11300)	(0.10594)	(0.08469)
9	71.9032	63.9881	63.0927	67.1840
	(0.05529)	(0.13820)	(0.13117)	(0.10154)
10	100.	100.	100.	100.

Source: Authors' calculations from Statistics Canada's Public Use Microdata Files for Canadian Censuses.

Standard errors in parentheses.

Table 4a
Generalized Lorenz Curve Ordinates for Women's Income
(Canada, 2020 $)

Decile	2000	2005	2015	2020
1	421.18	466.71	541.36	773.32
	(23.810)	(28.406)	(28.682)	(39.099)
2	1479.6	1601.9	1871.8	2515.6
	(56.697)	(57.033)	(61.531)	(72.710)
3	3049.6	3259.1	3770.8	4872.4
	(86.971)	(90.896)	(94.599)	(101.00)
4	5038.6	5367.0	6203.5	7823.4
	(119.86)	(118.21)	(130.60)	(144.57)
5	7523.3	8010.0	9308.1	11,454
	(167.07)	(165.40)	(183.37)	(190.07)
6	10,648	11,338	13,194	15,818
	(218.12)	(221.94)	(242.93)	(240.94)
7	14,563	15,446	17,943	21,021
	(276.20)	(274.60)	(297.85)	(293.82)
8	19,365	20,499	23,787	27,319
	(338.54)	(342.08)	(374.28)	(360.03)
9	25,502	27,007	31,393	36,365
	(419.31)	(430.79)	(467.56)	(449.25)
10	35,317	39,549	46,225	49,594
	(556.82)	(775.79)	(870.41)	(746.31)

Source: Authors' calculations from Statistics Canada's Public Use Microdata Files for Canadian Censuses.

Standard errors in parentheses.

Table 4b
Generalized Lorenz Curve Ordinates for Men's Income
(Canada, 2020 $)

Decile	2000	2005	2015	2020
1	737.12	720.43	759.32	1062.2
	(46.605)	(40.810)	(42.959)	(48.102)
2	2500.6	2456.0	2566.1	3227.9
	(90.077)	(90.319)	(81.682)	(80.653)
3	5053.4	4997.6	5172.5	6139.6
	(143.78)	(142.05)	(135.70)	(126.63)
4	8451.2	8398.8	8708.9	9889.5
	(209.48)	(208.12)	(203.23)	(184.60)
5	12,704	12,636	13,159	14,496
	(272.55)	(265.73)	(264.49)	(244.25)
6	17,833	17,780	18,594	20,052
	(346.20)	(340.37)	(338.87)	(309.02)
7	23,925	23,981	25,191	26,738
	(415.41)	(418.00)	(421.43)	(381.38)
8	31,271	31,523	33,295	34,873
	(501.97)	(507.39)	(522.03)	(470.13)
9	40,455	41,092	43,707	45,260
	(607.65)	(629.24)	(639.15)	(588.64)
10	56,263	64,219	69,274	67,368
	(901.63)	(1739.5)	(1801.6)	(1363.3)

Source: Authors' calculations from Statistics Canada's Public Use Microdata Files for Canadian Censuses.

Standard errors in parentheses.

Table 5a
Changes in Women's Decile Mean Incomes Between Censuses
(Canada, 2020 $)

Decile	2000-05	2005-15	2015-20	2000-20
1	455.22	746.49	2319.59	3521.30
	(12.28)	(18.49)	(47.84)	(76.92)
2	857.50	1862.90	4118.63	6839.04
	(17.32)	(37.16)	(74.84)	(125.6)
3	781.70	2508.45	4577.77	7867.93
	(14.43)	(44.08)	(85.99)	(156.4)
4	1189.75	3246.19	5183.48	9619.42
	(22.37)	(57.80)	(73.50)	(141.1)
5	1582.99	4616.21	5263.11	11,462
	(19.85)	(52.74)	(60.74)	(145.5)
6	2027.21	5581.93	4779.12	12,388
	(20.56)	(54.52)	(48.33)	(130.4)
7	1932.40	6411.17	4534.66	12,878
	(19.25)	(65.27)	(45.35)	(126.1)
8	2506.96	7907.97	4544.67	14,960
	(20.26)	(57.59)	(33.04)	(120.7)
9	3716.07	10,985	4397.59	19,098
	(21.54)	(59.65)	(23.74)	(109.9)
10	27,275	22,887	-6023.14	44,139
	(48.17)	(29.03)	(7.97)	(84.82)

Source: Authors' calculations from Statistics Canada's Public Use Microdata Files for Canadian Censuses.

Absolute values of asymptotic "t-ratios" in parentheses.

Table 5b
Changes in Men's Decile Mean Incomes Between Censuses
(Canada, 2020 $)

Decile	2000-05	2005-15	2015-20	2000-20
1	-166.89	388.81	3028.75	3250.66
	(2.79)	(6.56)	(46.96)	(50.07)
2	-278.62	712.16	3588.83	4022.38
	(3.55)	(9.72)	(59.28)	(60.25)
3	-122.00	648.10	3052.91	3589.02
	(1.25)	(7.22)	(35.40)	(41.83)
4	33.42	1351.14	2135.94	3520.50
	(0.30)	(11.91)	(19.86)	(33.08)
5	-155.02	2135.35	1554.56	3534.89
	(1.46)	(20.49)	(14.66)	(32.80)
6	142.59	2912.62	1217.96	4273.17
	(1.06)	(21.74)	(9.72)	(34.10)
7	1097.62	3954.89	891.28	5943.80
	(7.95)	(26.56)	(6.20)	(44.82)
8	1963.15	5613.68	312.93	7889.76
	(11.41)	(30.51)	(1.71)	(46.31)
9	3848.41	8429.32	-249.68	12,028
	(16.22)	(34.28)	(1.04)	(51.85)
10	73,183	24,404	-34,595	62,993
	(48.08)	(11.68)	(18.75)	(54.10)

Source: Authors' calculations from Statistics Canada's Public Use Microdata Files for Canadian Censuses.

Absolute values of asymptotic "t-ratios" in parentheses.

Table 6a
Changes in Women's Lorenz Curve Ordinates Between Censuses
(percentage points)

Decile	2000-05	2005-15	2015-20	2000-20
1	-0.0125	-0.0089	0.3881	0.3667
	(1.326)	(0.965)	(39.98)	(37.01)
2	-0.1164	-0.0237	1.0231	0.8829
	(5.806)	(1.246)	(53.98)	(44.22)
3	-0.3945	-0.0830	1.6670	1.1895
	(12.96)	(2.785)	(59.80)	(41.63)
4	-0.6966	-0.1502	2.3545	1.5078
	(17.35)	(3.702)	(61.04)	(39.54)
5	-1.0492	-0.1167	2.9596	1.7937
	(20.12)	(2.136)	(58.19)	(37.23)
6	-1.4835	-0.1243	3.3522	1.7444
	(22.79)	(1.774)	(52.38)	(29.83)
7	-2.1815	-0.2370	3.5682	1.1497
	(28.22)	(2.757)	(46.08)	(16.98)
8	-3.0027	-0.3714	3.6257	0.2516
	(33.99)	(3.622)	(39.73)	(3.36)
9	-3.9226	-0.3717	3.3944	-0.8999
	(41.07)	(3.152)	(32.61)	(11.57)

Source: Authors' calculations from Statistics Canada's Public Use Microdata Files for Canadian Censuses.

Absolute values of asymptotic "t-ratios" in parentheses.

Table 6b
Changes in Men's Lorenz Curve Ordinates Between Censuses
(percentage points)

Decile	2000-05	2005-15	2015-20	2000-20
1	-0.1883	-0.0257	0.4806	0.2666
	(19.12)	(2.822)	(50.03)	(25.87)
2	-0.6200	-0.1201	1.0871	0.3470
	(29.49)	(5.940)	(57.93)	(17.67)
3	-1.1995	-0.3154	1.6468	0.1318
	(35.65)	(9.185)	(52.33)	(4.291)
4	-1.9423	-0.5068	2.1082	-0.3409
	(39.62)	(9.681)	(44.55)	(7.816)
5	-2.9029	-0.6802	2.5209	-1.0621
	(44.72)	(9.400)	(38.81)	(18.78)
6	-4.0101	-0.8441	2.9236	-1.9306
	(48.21)	(8.782)	(34.22)	(27.36)
7	-5.1801	-0.9780	3.3255	-2.8326
	(50.25)	(7.923)	(30.53)	(33.26)
8	-6.4920	-1.0246	3.7029	-3.8138
	(51.74)	(6.615)	(27.30)	(37.85)
9	-7.9151	-0.8954	4.0914	-4.7191
	(53.17)	(4.699)	(24.67)	(40.82)

Source: Authors' calculations from Statistics Canada's Public Use Microdata Files for Canadian Censuses.

Absolute values of asymptotic "t-ratios" in parentheses.

Table 7a
Changes in Women's Generalized Lorenz Curve Ordinates Between Censuses
(Canada, 2020 $)

Decile	2000-05	2005-15	2015-20	2000-20
1	45.525	74.655	231.96	352.14
	(1.228)	(1.849)	(4.784)	(7.692)
2	131.28	260.92	643.83	1036.0
	(1.632)	(3.110)	(6.759)	(11.24)
3	209.47	511.73	1101.6	1822.8
	(1.665)	(3.901)	(7.962)	(13.68)
4	328.38	836.46	1619.9	2784.7
	(1.951)	(4.749)	(8.315)	(14.83)
5	486.70	1298.0	2146.3	3931.0
	(2.095)	(5.256)	(8.127)	(15.70)
6	689.44	1856.1	2624.2	5169.8
	(2.216)	(5.641)	(7.670)	(15.91)
7	882.56	2497.5	3077.6	6457.6
	(2.266)	(6.165)	(7.356)	(16.01)
8	1133.3	3288.2	3532.0	7953.5
	(2.355)	(6.485)	(6.801)	(16.09)
9	1504.9	4386.5	3971.8	9863.3
	(2.503)	(6.900)	(6.125)	(16.05)
10	4232.6	6675.3	3369.5	14,277
	(4.432)	(5.725)	(2.939)	(15.33)

Source: Authors' calculations from Statistics Canada's Public Use Microdata Files for Canadian Censuses.

Absolute values of asymptotic "t-ratios" in parentheses.

Table 7b
Changes in Men's Generalized Lorenz Curve Ordinates Between Censuses
(Canada, 2020 $)

Decile	2000-05	2005-15	2015-20	2000-20
1	-16.693	38.892	302.86	325.06
	(0.280)	(0.656)	(4.696)	(5.007)
2	-44.569	110.13	661.72	727.28
	(0.349)	(0.904)	(5.765)	(6.015)
3	-55.790	174.89	967.06	1086.2
	(0.276)	(0.890)	(5.210)	(5.669)
4	-52.342	310.05	1180.6	1438.3
	(0.177)	(1.066)	(4.300)	(5.151)
5	-67.878	523.50	1336.2	1791.8
	(0.178)	(1.396)	(3.711)	(4.896)
6	-53.658	814.83	1457.9	2219.0
	(0.111)	(1.697)	(3.179)	(4.782)
7	56.293	1210.2	1547.1	2813.6
	(0.096)	(2.039)	(2.722)	(4.989)
8	252.55	1771.7	1578.3	3602.5
	(0.354)	(2.434)	(2.247)	(5.238)
9	637.32	2614.4	1553.5	4805.3
	(0.729)	(2.915)	(1.788)	(5.680)
10	7955.6	5055.1	-1906.3	11,104
	(4.060)	(2.019)	(0.844)	(6.794)

Source: Authors' calculations from Statistics Canada's Public Use Microdata Files for Canadian Censuses.

Absolute values of asymptotic "t-ratios" in parentheses.

Table 8a
Cumulative Coefficients of Variations for Women's Income
2020-2020

Decile	2000	2020	Difference
1	0.48452	0.53936	0.05484
2	0.49802	0.46666	-0.03136
3	0.49152	0.43788	-0.05364
4	0.48226	0.43340	-0.04886
5	0.48863	0.44350	-0.04513
6	0.51102	0.45916	-0.05186
7	0.54271	0.47943	-0.06328
8	0.57499	0.50885	-0.06614
9	0.62280	0.56059	-0.06221
10	0.79793	0.88389	0.08596

Source: Authors' calculations from Statistics Canada's Public Use Microdata Files for Canadian Censuses.

Table 8b
Cumulative Coefficients of Variations for Men's Income
2020-2020

Decile	2000	2020	Difference
1	0.51383	0.50927	-0.00456
2	0.48141	0.42565	-0.05576
3	0.47404	0.41109	-0.06295
4	0.48383	0.42236	-0.06147
5	0.49454	0.43817	-0.05637
6	0.50570	0.45776	-0.04794
7	0.51855	0.48231	-0.03624
8	0.54005	0.51484	-0.02521
9	0.57840	0.56809	-0.01031
10	0.80861	1.17690	0.36829

Source: Authors' calculations from Statistics Canada's Public Use Microdata Files for Canadian Censuses.

Table A1
Summary Statistics
(Individual Censuses - 2020 $)

a) Women

	2000	2005	2015	2020
Mean Income -	$35,317	$39,549	$46,225	$49,594
NOBS -	256,129	274,147	313,063	345,002

b) Men

	2000	2005	2015	2020
Mean Income -	$56,263	$64,219	$69,274	$67,368
NOBS -	254,607	266,549	304,245	338,219

Source: Authors' calculations from Statistics Canada's Public Use Microdata Files for Canadian censuses.

<u>References</u>

Aaberge, R. (2000) "Characterizations of Lorenz Curves and Income Distributions", *Social Choice and Welfare* 17, 639-53.

— (2001) "Axiomatic Charaterization of the Gini Coeffieint and Lorenz Curve Orderings", *Journal of Economic Theory* 101, 115-32.

Acemoglu, D.H., D. Autor, G.H. Hanson, and B. Price (2016) "Import Competition and the Great U.S. Employment Sag of the 2000s", *Journal of Labor Economics* 34, S141-98.

Ahn, S., and J.A. Fessler (2003) "Standard Errors of Mean, Variance and Standard Deviation Estimators, available on-line at: web.eecs.umich.edu/~fessler/papers/files/tr/stderr.pdf.

Alvaredo, F., L. Chancel, T. Piketty, E. Saez, and G. Zucman (2018) "For Good Measure: Advancing Research on Well- Being Metrics Beyond GDP", *Organization for Economic Cooperation and Development*, chapter on Distributional National Accounts, 143-61.

Alvaredo, F, A.B. Atkinson, L. Chancel, T. Piketty, E Saez, and G. Zucman (2020) "Distributional National Accounts Guidelines: Methods and Concepts Used in WID.world", *World Inequality Database*, WID Working Paper No. 2016/2 (version of April 5, 2018).

Anderson, G. (1996) "Nonparametric Tests of Stochastic Dominance in Income Distributions",*Econometrica* 64 (5), 1183-93.

Atkinson, A.B. (1970) "On the Measurement of Inequality", *Journal of Economic Theory* 2, 244-63.

Autor, D.H., D. Dorn, and G.H. Hanson (2013) "The Geography of Trade and Technology Shocks in the United States", *American Economic Review* 103, 220-25.

Barrett, G.F., and S.G. Donald (2003) "Consistent Tests for Stochastic Dominance", *Econometrica* 71(1), 71-104.

Beach, C.M. (2016) "Changing Income Inequality: A Distributional Paradigm for Canada", *Canadian Journal of Economics* 49(4), 1229-92.

Beach, C.M., and J. Richmond (1985) "Joint Confidence Intervals for Income Shares and Lorenz Curves", *International Economic Review* 26(2), 439 – 50.

Beach, C.M., and R. Davidson (2024a) "Income Share Standard Errors and a Quantile Toolbox of Distributional Statistics", presented at the Canadian Economics Association annual meetings at Toronto Metropolitan University, May; forthcoming as "Quantile Means and Quantile Share Standard Errors and a Toolbox or Distributional Statistics".

— (2024b) "A Statistical Characterization of Median-Based Inequality Measures", presented at the Canadian Econometrics Study Group meetings at York University, October; Queen's University, Department of Economics Discussion Paper.

Beach, C.M., R Davidson, and G. A. Slotsve (1994) "Distribution-Free Statistical Inference for Inequality Dominance with Crossing Lorenz Curves", Queen's University, Department of Economics Discussion Paper No. 912, November.

Bishop, J. A., J. P. Formby, and P. D Thistle (1989) "Statistical Inference, Income Distributions and Social Welfare", in *Research on Economic Inequality*, vol. 1, ed. D.G. Slottje. Greenwich, CT: JAI Press, 49 – 82.

— (1992) “Convergence of the South and Non-South Income Distributions, 1969 – 1979”, *American Economic Review* 82(1), 262 -72.

Blackorby, C., W. Bossert, and W. Donaldson (1999) “Income Inequality Measurement: The Normative Approach”, in *Handbook of Income Inequality Measurement*, ed. J Silber. New York: Springer Science + Business Media, 133- 57.

Boadway, R. W., and N. Bruce (1984) *Welfare Economics*. Oxford: Basil Blackwell Inc.

Bosmans, K., and Z.E. Ozturk (2021) “Measures of Inequality of Opportunity: A Normative Approach”, *The Journal of Economic Inequality* 19, 213 – 37.

Chiu, W.H. (2007) “Intersecting Lorenz Curves, the Degree of Downward Inequality Aversion, and Tax Reforms”, *Social Choice and Welfare* 28(3), 375-99.

Cowell, F.A. (2011) *Measuring Inequality*, Third Edition. Oxford: Oxford University Press.

Davidson, R., and J-Y. Duclos (2000) “Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality”, *Econometrica* 68(6), 1435 – 64.

Davies, J.B., and M. Hoy (1994) “The Normative Significance of Using Third-Degree Stochastic Dominance in Comparing Income Distributions”, *Journal of Economic Theory* 64(2), 520-30.

— (1995) “Making Inequality Comparisons when Lorenz Curves Intersect”, *American Economic Review* 85(3), 980-86.

Davies, J.B., M. Hoy and Z. Lin (2022) “Revising Comparisons of Income Inequality When Lorenz Curves Intersect”, *Social Choice and Welfare* 58(1), 101-09.

Del Rio, C., C. Gradin, and O. Canto (2011) "The Measurement of Gender Wage Discrimination: The Distributional Approach Revisited", *Journal of Economic Inequality* 9, 57-86.

Drummond, D. (2021) "Viewpoint: Canada Should Establish an Equitable Growth Institute", Centre for the Study of Living Standards, *International Productivity Monitor*, No. 38 (Spring).

Duclos, J.-Y, and. A. Araar (2006) *Poverty and Equity: Measurement, Policy, and Estimation with DAD*, International Development Research Centre, Ottawa: Springer Science + Business Media.

The Economist (2023) "Incomes: The Unstuck Middle", Jan. 21, 2023, 19.

The Economist (2021a) "Free Exchange: Fleshing out the Olive", Aug. 28, 2021, 65.

— (2021b) "Free Exchange: Black Cat, White Cat, Fat Cat, Thin Cat", Oct. 2, 2021, 62.

— (2021c) "Incomes: Piketty Lines", Oct. 9, 2021, 24.

Fleurbaey, M. (2009) "Beyond GDP: The Quest for a Measure of Social Welfare", *Journal of Economic Literature* 47(4), 1029-75.

Goos, M., A. Manning, and A. Salomons (2014) "Explaining Job Polarization: Routine-Biased Technological Change and Offshoring", *American Economic Review* 104, 2509-26.

Guvenen, F., L. Pistaferri, and G.L. Violante (2022) "Global Trends in Income Inequality and Income Dynamics: New Insights from GRID", *Quantitative Economics* 13, 1321-60.

Hoffman, F., D.S. Lee, and T. Lemieux (2020) "Growing Income Inequality in the United States and Other Advanced Economics", *Journal of Economic Perspectives* 34, 52-78.

Jenkins, S.P. (1994) "Earnings Discrimination Measurement: A Distributional Approach", *Journal of Econometrics* 61, 81-102.

Jones, C.I., and P.J. Klenow (2016) "Beyond GDP? Welfare Across Countries and Time", *American Economic Review* 106(9), 2426-57.

Jorgenson, D. (2018) "Production and Welfare: Progress in Economic Measurement", *Journal of Economic Literature* 56(3), 867-919.

Lambert, P.J. (2001) *The Distribution and Redistribution of Income*, Third Edition. Manchester: Manchester University Press.

Linton, O., E. Massoumi, and Y-G. Whang (2005) "Consistent Testing for Stochastic Dominance Under General Sampling Schemes", *Review of Economic Studies* 72, 735-65.

Maasoumi, E. (1998) "Empirical Analyses of Welfare and Inequality", Ch. 5 in *Handbook of Applied Econometrics*, Vol. II: Microeconometrics, eds. M.H. Pearson and P. Schmidt. Oxford: Blackwell Publishers, 187-226.

Maasoumi, E., and A. Heshmati (2000) "Stochastic Dominance Amongst Swedish Income Distributions", *Econometric Reviews* 19(3), 287-320.

Palmisano, F., and V. Peragine (2015) "The Distributional Incidence of Growth: A Social Welfare Approach", *Review of Income and Wealth* 61(3), 440-64.

Peragine, V. (2004) “Ranking Income Distributions According to Equality of Opportunity”, *Journal of Economic Inequality* 2, 11- 30.

Post, T. (2003) “Empirical Tests for Stochastic Dominance Efficiency”, *Journal of Finance* 68(5), 1905-31.

Rapley, J. (2024) “Recession Worries May Be Wrong Focus in Economy That Is Neither Sick Nor Strong”, The Globe and Mail, Report on Business, B2 (Jan. 20, 2024).

Rendell, M. (2024) “The Bank of Canada Prepares for a Future That Does Not Look Like the Past”, The Globe and Mail, Report on Business, B6 (Jan. 20, 2024).

Saez, E., and M.R. Veall (2007) “The Evolution of High Incomes in Canada, 1920-2000”, Ch. 6 in *Top Incomes over the Twentieth Century*, eds. A.B. Atkinson and T. Piketty. Oxford: Oxford University Press.

Salas, R., J.A. Bishop, and L.A. Zearzer (2018) “Second-Order Discrimination and Generalized Lorenz Dominance”, *Review of Income and Wealth* 64(3), 563-75.

Saposnik, R. (1981) “Rank Dominance in Income Distribution”, *Public Choice* 36(1), 147 – 51.

Scaillet, O., and N. Topaloglou (2010) “Testing for Stochastic Dominance Efficiency”, *Journal of Business and Economic Statistics* 28(1), 169-80.

Shorrocks, A. F. (1983) “Ranking Income Distributions”, *Economica* 50(197), 3-17.

Shorrocks, A.F., and J. Foster (1987) “Transfer Sensitive Inequality Measures”, *Review of Economic Studies* 54, 485-97.

Statistics Canada (2024) “2021 Census Public Use Microdata File (PUMF): Individuals File — Documentation and Use Guide”, Cat. No. 98M 0001X.

Stoline, M. R., and H. K. Ury (1979) “Tables of the Studentized Maximum Modulus Distribution and Applications to Multiple Comparisons Among Means”, *Technometrics* 21(1). 87-93.

Zucman, G., T. Piketty, and E. Saez (2018) “Distributional National Accounts: Methods and Estimates for the United States”, *Quarterly Journal of Economics* 133(2), 553- 609.