# Statistical Inference for the Measurement of the Incidence of Taxes and Transfers

by

**Russell Davidson**

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

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

and

**Jean-Yves Duclos**

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

**Abstract**

We establish the asymptotic sampling distribution of general functions of quantile-based estimators computed from samples that are not necessarily independent. The results provide the statistical framework within which to assess the progressivity of taxes and benefits, their horizontal inequity, and the change in the inequality of income which they cause. By the same token, these findings characterise the sampling distribution of a number of popular indices of progressivity, horizontal inequity, and redistribution. They can also be used to assess welfare and inequality changes using panel data, and to assess poverty when it depends on estimated population quantiles. We illustrate these results using micro data on the incidence of taxes and benefits in Canada.

**Keywords** Income inequality indices, Tax progressivity, Horizontal equity, Poverty indices, Distribution-free statistical inference.

This research was supported, in part, by grants from the Social Sciences and Humanities Research Council of Canada, and from the Human Capital and Mobility Program of the European Union (contract #ERBCHRXCT940647). We are grateful to participants at the July 1995 meeting of the ESRC Econometric Study Group for valuable comments, and especially to Ian Preston. We are also indebted to two anonymous referees for suggestions that have greatly improved the paper. Remaining errors are ours alone.

January, 1997

## 1. Introduction

The last decade has seen considerable use and development of statistical theory for inferring the dominance of one distribution (of income, wealth, wages, *etc.*) over another. Various welfare criteria have been applied, such as first- and second-order stochastic dominance, Lorenz dominance, "transfer-sensitivity" dominance, and comparisons of "poverty deficit curves" (Beach and Davidson (1983), Bishop, Formby and Thistle (1992), Howes (1993), Beach, Davidson and Slotsve (1994)). All such criteria involve quite general principles of anonymity, efficiency, and equity (see, for instance, Shorrocks (1983) and Davies and Hoy (1994)). The comparisons typically seek to establish inequality and social welfare rankings using independently-drawn samples from the relevant populations.

We extend these developments to the measurement of redistribution, progressivity, and horizontal inequity. More generally, we establish the asymptotic sampling distribution of general functions of quantile-based estimators computed from samples that are not necessarily independent. Dependent samples may be due, for instance, to the correlation between gross and net income distributions, or to the correlation of incomes across time when panel data are used. The results thus provide the statistical framework within which to assess the progressivity of taxes and benefits, and the changes, in the inequality of income, or in the ranking of individuals with respect to income, which they may cause. Similarly, one can readily obtain the sampling distributions of a number of popular or recent measures of progressivity, horizontal inequity, and redistribution (see Musgrave and Thin (1948), Suits (1977), Reynolds and Smolensky (1976), Kakwani (1977), Atkinson (1979), Plotnick (1981), Pfähler (1987), Aronson et al. (1994), and Lerman and Yitzhaki (1995)). The results can also be applied to the impact on poverty indices of a tax and benefit system, or of other socio-economic phenomena, when such poverty indices depend on estimated population quantiles[1]. They furthermore encompass as special cases most of the previous statistical inference results for the measurement of inequality and social welfare. Our results are distribution-free in the sense that they do not require a specification of the population distributions from which the samples are drawn.

[1] For instance, the poverty line may be half of median income, or the weights on the incomes of the poor in the poverty index may depend on their ranking in the income distribution.

## 2. The Measurement of Progressivity, Horizontal Inequity, and Redistribution

Consider a population of households indexed by $\omega \in \Omega$. Let $X(\omega)$ denote the gross income of household $\omega$, $T(\omega)$ the tax burden of the household, and $M(\omega) = X(\omega) - T(\omega)$ its income net of taxes. We write the distribution on $\Omega$ as $F(\omega)$, so that mean gross income, $\mu_X$, is defined as

$$\mu_X = \int_\Omega X(\omega)\, dF(\omega).$$

The mean tax burden, $\mu_T$, and mean net income, $\mu_M$, are defined similarly. The Lorenz curve $L_X(p)$ for gross income is defined as

$$L_X(p) = \frac{1}{\mu_X} \int_\Omega I_{\{X(\omega)\le y(p)\}}(\omega)\; X(\omega)\; dF(\omega),$$

where the indicator function $I_{\{X(\omega)\le x\}}$ is defined by

$$I_{\{X(\omega)\le x\}}(\omega) = \begin{cases} 1 & \text{if } X(\omega) \le x; \\ 0 & \text{otherwise,} \end{cases}$$

and where $y(p)$ is defined implicitly by the relation

$$p = \int_\Omega I_{\{X(\omega)\le y(p)\}}(\omega)\; dF(\omega), \tag{1}$$

that is, $y(p)$ is the $p-$quantile of the distribution of the random variable $X$. The concentration curves $L_M$ and $L_T$ for net income and taxes are, respectively:

$$L_M(p) = \frac{1}{\mu_M} \int_\Omega I_{\{X(\omega)\le y(p)\}}(\omega)\; M(\omega)\; dF(\omega),$$

and

$$L_T(p) = \frac{1}{\mu_T} \int_\Omega I_{\{X(\omega)\le y(p)\}}(\omega)\; T(\omega)\; dF(\omega).$$

The Lorenz curve for $M$ is defined as:

$$L^*(p) = \frac{1}{\mu_M} \int_\Omega I_{\{M(\omega)\le y^*(p)\}}(\omega)\; M(\omega)\; dF(\omega),$$

where $y^*(p)$ is defined as in (1) as the $p-$quantile of the distribution of $M$. $L_X(p)$, $L_M(p)$, $L_T(p)$ and $L^*(p)$ are the basis for much of the theory of the measurement of inequality, progressivity, horizontal inequity, and redistribution. For instance, it is well known that, if and only if $L^*(p)$ dominates

$L_X(p)$, inequality under $X$ will be greater than under $M$ for all inequality measures that satisfy

$P$ : symmetry (or anonymity), mean independence, and the strict Dalton-Pigou principle of transfers.

Another well-known result is that a progressive tax which does not rerank individuals necessarily causes $L_T(p)$ to be dominated by $L_X(p)$, and $L_M(p)$ and $L^*(p)$ to dominate $L_X(p)$ (Jakobsson (1976)). Measurements of progressivity have thus naturally been based both on the distance between $L_X(p)$ and $L_T(p)$ and on the distance between $L_M(p)$ and $L_X(p)$, yielding indices based on what are called the tax redistribution (TR) and income redistribution (IR) views, respectively (see Pfähler (1987)). A tax $T$ is TR progressive if and only if $L_X$ dominates $L_T$, and a tax $T$ is IR progressive if and only if $L_M$ dominates $L_X$[2]. The farther is $L_T$ from $L_X$, or $L_X$ from $L_M$, the more TR or IR progressive is a tax[3].

Testing whether a tax $T_2$ is more TR progressive than a tax $T_1$, we would then need to check whether $L_{T_1}$ dominates $L_{T_2}$; similarly, testing whether $T_2$ is more IR progressive than $T_1$ involves inferring whether $L_{M_2}$ dominates $L_{M_1}$, where $M_i = X - T_i$, $i = 1, 2$.

The distance between $L_M(p)$ and $L^*(p)$ can be used as an indication of the presence of reranking in the redistributive process of moving from $X$ to $M$, and, under some interpretations, of horizontal inequity – see Plotnick (1982) and Feldstein (1976), p.83. It is well known that $L_M(p) \geq L^*(p)$ for all $p$, with strict inequality somewhere, if and only if there is reranking, and the greater the distance between $L_M$ and $L^*$, the greater the extent of reranking.

In order to perform statistical inference on the incidence, in the form of redistribution, progressivity, and horizontal inequity, of taxes and benefits, we will establish in the next section the joint sampling distribution of $L_X$, $L_M$, $L_T$, and $L^*$. This will enable us to

a) test whether $L_X$ dominates $L_T$, or whether $L_M$ dominates $L_X$, to determine whether $T(X)$ is TR or IR progressive;

[2] A benefit $B$ is, however, TR progressive if and only if its concentration curve $L_B$ dominates $L_X$.

[3] There is some debate as to which of the IR and TR views is more appropriate as a basis for the measurement of progressivity (for an overview of this debate, see, for instance, Lambert (1993), ch.7). In the absence of reranking and when comparing two taxes yielding the same average tax rate, the two views are equivalent and yield the same ordering (Formby et al. (1990)). In general, however, the IR view is more closely linked to the income redistribution effected by a tax.

b) test whether $L_{T_1}$ is dominated by $L_{T_2}$, or whether $L_{M_1}$ dominates $L_{M_2}$, to infer whether $T_1$ is more TR or IR progressive than $T_2$;

c) test the distance between $L_M$ and $L^*$ to assess the presence of reranking and horizontal inequity;

d) determine whether $T$ is redistributive and inequality reducing by testing whether $L^*(p)$ dominates $L_X(p)$.

By knowing the joint sampling distributions of $L_X$, $L_M$, $L_T$, and $L^*$, we can also derive the sampling distributions of several commonly used indices and measures – see Duclos (1993). When performing comparisons across time using panel data, our methods are directly applicable. For instance, two Lorenz curves $L_X(p)$ obtained at two different times can be treated as dependent Lorenz curves obtained from paired observations. Finally, quantile-based poverty comparisons and indices can be seen as special cases of the above when the focus is put upon the lower portions of the income distributions (see Atkinson (1987) and Howes (1993)).

## 3. Asymptotic Distribution of Quantile-Based Estimators

Consider two jointly distributed random variables $Y$ and $Z$, and let $F$ denote the cumulative distribution function (c.d.f.) of the marginal distribution of $Z$. We are interested in estimating expectations of $Y$ conditional on $Z$ being smaller than the $p-$quantile of its distribution, that is, expectations like $\gamma_p \equiv E(Y \mid F(Z) \leq p)$. Formally:

$$p\gamma_p \equiv pE\big(Y \mid Z \leq G(p)\big) = E\big(Y\, I_{[0,G(p)]}(Z)\big), \tag{2}$$

where $G$ is the inverse of $F$. Here the indicator function satisfies

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

If $Y \equiv Z$, we note that $p\gamma_p$ yields the generalised Lorenz curve of $Y$ (Shorrocks (1983)), for which Beach and Davidson (1983) first derived the asymptotic sampling distribution.

Consider also a second set of two jointly distributed random variables $V$ and $W$, and let $F^*$ denote the c.d.f. of the marginal distribution of $W$. Analogously to (2), we may define the conditional expectation $\delta_p$ by

$$p\delta_p \equiv pE\big(V \mid W \leq G^*(p)\big) = E\big(V\, I_{[0,G^*(p)]}(W)\big), \tag{3}$$

where $G^*$ is the inverse of $F^*$.

Suppose that $N$ independent drawings have been made from the joint distribution of $Y$ and $Z$; write them as $(Y_i, Z_i)$, $i = 1, \ldots, N$. An obvious estimator of $p\gamma_p$ is then given by

$$p\hat{\gamma}_p = N^{-1} \sum_{i=1}^{N} Y_i \, I_{[0,\hat{G}(p)]}(Z_i) \tag{4}$$

where $\hat{G}(p)$ is the sample estimate of the $p-$quantile of $Z$. We may define $p\hat{\delta}_p$ as an estimate of $p\delta_p$ from a sample of $N$ independent drawings $(V_i, W_i)$ in exactly the same way.

We now show that the estimators $\hat{\gamma}_p$ and $\hat{\delta}_p$ are root-$N$ consistent and asymptotically normal, with an asymptotic covariance matrix that can be estimated consistently without knowledge of the population distribution from which our sample was drawn. Our first result gives the asymptotic covariance between $p\hat{\gamma}_p$ and $p'\hat{\delta}_{p'}$ for arbitrary $0 \le p \le 1$ and $0 \le p' \le 1$.

**Theorem 1:** Let the population second moments of $Y$ and $V$ conditional on $Z$ and $W$ be finite, and let the first moments be continuously differentiable in $Z$ and $W$. Further, let the marginal cumulative distribution functions of $Z$ and $W$ be strictly monotonic and continuously differentiable. Then the asymptotic covariance of $p\hat{\gamma}_p$ and $p'\hat{\delta}_{p'}$, as defined by (4) and its analogue for $\delta_{p'}$, is given by

$$\begin{aligned}\lim_{N\to\infty} N \operatorname{cov}\big(p\hat{\gamma}_p, p'\hat{\delta}_{p'}\big) &= E\big(YV \, I_{[0,G(p)]}(Z) \, I_{[0,G^*(p')]}(W)\big) \\ &-E\big(Y|\, Z = G(p)\big) \, E\big(V \, I_{[0,G(p)]}(Z) \, I_{[0,G^*(p')]}(W)\big) \\ &-E\big(V|\, W = G^*(p')\big) \, E\big(Y \, I_{[0,G(p)]}(Z) \, I_{[0,G^*(p')]}(W)\big) \\ &+ E\big(Y|\, Z = G(p)\big) \, E\big(V|\, W = G^*(p')\big) \, E\big(I_{[0,G(p)]}(Z) I_{[0,G^*(p')]}(W)\big) \\ &-pp'\Big(\Big(\gamma_p - E\big(Y|\, Z = G(p)\big)\Big)\Big(\delta_{p'} - E\big(V|\, W = G^*(p')\big)\Big)\Big).\end{aligned} \tag{5}$$

**Proof:** See Appendix. ∎

**Remarks and Corollaries:**

(1) Everything in (5) can be estimated consistently in a distribution-free manner: $\gamma_p$ and $\delta_{p'}$ by $\hat{\gamma}_p$ and $\hat{\delta}_{p'}$, $G(p)$ and $G^*(p')$ by $\hat{G}(p)$ and $\hat{G}^*(p')$, that is, the sample $p$ and $p'$ quantiles of $Z$ and $W$ respectively. The unconditional expectations are readily estimated by their sample equivalents; thus for $E(YV \, I_{[0,G(p)]}(Z) \, I_{[0,G^*(p')]}(W))$, for instance, we may use the estimate

$$N^{-1} \sum_{i=1}^{N} Y_i V_i \, I_{[0,\hat{G}(p)]}(Z_i) \, I_{[0,\hat{G}^*(p')]}(W_i),$$

in which the sum is effectively over only those drawings $i$ for which $Z_i$ is less than or equal to the $p-$quantile of the sample distribution of $Z$, *and* $W_i$ is less than or equal to the $p'-$quantile of the sample distribution of $W$. The conditional expectations are a little less obvious, but various forms of kernel estimation can be used under the regularity conditions of the theorem.

(2) The expression (5) gives explicitly only the covariance of $p\hat{\gamma}_p$ and $p'\hat{\delta}_{p'}$. However, the covariance of $p\hat{\gamma}_p$ and $p'\hat{\gamma}_{p'}$ can be obtained directly from (5), by replacing $\hat{\delta}_{p'}$, $V$, $W$, and $G^*$ by $\hat{\gamma}_{p'}$, $Y$, $Z$, and $G$ respectively. This procedure leads to certain simplifications, as seen in the following Corollary.

**Corollary 1** Under the conditions of Theorem 1, the asymptotic covariance between $p\hat{\gamma}_p$ and $p'\hat{\gamma}_{p'}$, for $p \le p'$, is given by

$$\begin{aligned}\lim_{N\to\infty} N \operatorname{cov}\big(p\hat{\gamma}_p, p'\hat{\gamma}_{p'}\big) = p\bigg(&\phi_p - \gamma_p^2 \\ &+(1-p')\Big(E\big(Y|\,Z=G(p)\big)-\gamma_p\Big)\Big(E\big(Y|\,Z=G(p')\big)-\gamma_{p'}\Big) \\ &+\Big(E\big(Y|\,Z=G(p)\big)-\gamma_p\Big)\big(\gamma_{p'}-\gamma_p\big)\bigg), \end{aligned} \tag{6}$$

where

$$p\phi_p \equiv E\big(Y^2\, I_{[0,G(p)]}(Z)\big). \tag{7}$$

If $p > p'$, the asymptotic covariance is obtained from (6) by inverting the roles of $p$ and $p'$. Setting $p = p'$ yields the variance of $p\hat{\gamma}_p$.

**Proof:** Making the replacements given above in (5), noting that, for $p \le p'$,

$$I_{[0,G(p)]}(Z)\; I_{[0,G(p')]}(Z) = I_{[0,G(p)]}(Z),$$

and making use of the definition (7) yields the following expression for the asymptotic covariance of $p\hat{\gamma}_p$ and $p'\hat{\gamma}_{p'}$:

$$\begin{aligned} p\bigg(&\phi_p - \gamma_p E\big(Y|\,Z=G(p)\big) - \gamma_p E\big(Y|\,Z=G(p')\big) \\ &+E\big(Y|\,Z=G(p)\big)\; E\big(Y|\,Z=G(p')\big) \\ &-p'\Big(\gamma_p - E\big(Y|\,Z=G(p)\big)\Big)\Big(\gamma_{p'} - E\big(Y|Z=G(p')\big)\Big)\bigg). \end{aligned}$$

This can readily be seen to be equal to the right-hand side of (6), the algebraic form of which is the same as that of similar expressions in Beach and Davidson (1983). The last statement in the enunciation of the Corollary follows from the symmetry of the covariance matrix of $\hat{\boldsymbol{\gamma}}$. ∎

(3) It is immediately clear that the asymptotic covariance of $p\hat{\delta}_p$ and $p'\hat{\delta}_{p'}$ has exactly the same form as (6), with $\gamma_p$ and $\gamma_{p'}$ replaced by $\delta_p$ and $\delta_{p'}$, $Y$ and $Z$ replaced by $V$ and $W$, $G$ replaced by $G^*$, and $\phi_p$ replaced by $\psi_p \equiv E(V^2\, I_{[0,G^*(p)]}(W))$.

(4) If $Z$ and $W$ are the same variable, then the covariance of $p\hat{\gamma}_p$ and $p'\hat{\delta}_{p'}$ simplifies in the same way as the covariance in (6).

(5) In some cases, the variable $Y$ may be the same as $Z$, or may be a deterministic function of $Z$. Or $V$ may be a deterministic function of $W$. In such cases, one or more of the conditional expectations in expressions (5) or (6) become trivial to evaluate. For instance, if $\delta_p = E(X|\, X \le G(p))$, then in our general notation, the variables $Z$, $W$, and $V$ are all equal to $X$, and we would have

$$E\big(V|\, Z = G(p)\big) = E\big(X|\, X = G(p)\big) = G(p).$$

This quantity can be directly estimated as $\hat{G}(p)$. The results of Beach and Davidson (1983), and of Beach, Davidson, and Slotsve (1994) then follow as special cases of the results presented here.

(6) One may legitimately wonder to what extent the results of the above Theorems and Corollaries are affected by the presence of measurement error. It is, after all, perfectly reasonable to suppose that things like incomes, taxes, and transfers are incorrectly reported in the available data sets, whether deliberately or otherwise. This matter merits further study.

Consider a set of $K$ probabilities, $p_i$, $i = 1, \ldots, K$, such that

$$0 < p_1 < p_2 < \ldots < p_K < 1.$$

Then let the $K$-vector $\boldsymbol{\gamma}$ be given by $\boldsymbol{\gamma} \equiv [\gamma_{p_1} \ldots \gamma_{p_K}]^\top$, and similarly for $\boldsymbol{\delta}$. We may define the following measure of the distance between the expectations $\gamma_p$ and $\delta_p$:

$$\mathit{\Gamma}_p = p\Big(\frac{\gamma_p}{\gamma_1} - \frac{\delta_p}{\delta_1}\Big), \tag{8}$$

where $\gamma_1$ and $\delta_1$ are just the expectations of $Y$ and $V$ respectively. We write $\boldsymbol{\mathit{\Gamma}} \equiv [\mathit{\Gamma}_{p_1} \ldots \mathit{\Gamma}_{p_K}]^\top$, and

$$\boldsymbol{\mathit{\Theta}} = \big[p_1\gamma_{p_1}, \ldots, p_K\gamma_{p_K}, \gamma_1, p_1\delta_{p_1}, \ldots, p_K\delta_{p_K}, \delta_1\big]^\top.$$

We already have estimators with known asymptotic properties for all the components of $\boldsymbol{\mathit{\Theta}}$; let us write the $2(K+1)$-vector of these estimators as $\hat{\boldsymbol{\mathit{\Theta}}}$. The obvious estimator of $\boldsymbol{\mathit{\Gamma}}$ is

$$\hat{\boldsymbol{\mathit{\Gamma}}} = \bigg[p_1\Big(\frac{\hat{\gamma}_{p_1}}{\hat{\gamma}_1} - \frac{\hat{\delta}_{p_1}}{\hat{\delta}_1}\Big) \ldots p_K\Big(\frac{\hat{\gamma}_{p_K}}{\hat{\gamma}_1} - \frac{\hat{\delta}_{p_K}}{\hat{\delta}_1}\Big)\bigg]^\top, \tag{9}$$

whose distribution can be obtained naturally from the joint distribution of the components of the $\hat{\boldsymbol{\Theta}}$. For its covariance matrix, define the $K \times 2(K+1)$ Jacobian $\boldsymbol{J}$ of the $K$-vector $\boldsymbol{\Gamma}$ with respect to the components of $\boldsymbol{\Theta}$ as follows:

$$\boldsymbol{J} \equiv \left[\frac{\partial \Gamma_i}{\partial \theta_j}\right] = [\, \boldsymbol{S}(\boldsymbol{\gamma}) \quad \vdots \quad -\boldsymbol{S}(\boldsymbol{\delta}) \,]$$

where the $K \times (K+1)$ matrices $\boldsymbol{S}(\boldsymbol{\gamma})$ and $\boldsymbol{S}(\boldsymbol{\delta})$ are given generically by the formula

$$\boldsymbol{S}(\boldsymbol{\alpha}) = \begin{bmatrix} 1/\alpha_1 & & & \vdots & -\dfrac{p_1 \alpha_{p_1}}{\alpha_1^2} \\ & \ddots & & \vdots & \vdots \\ & & 1/\alpha_1 & \vdots & -\dfrac{p_K \alpha_{p_K}}{\alpha_1^2} \end{bmatrix}.$$

We can now use a standard result of Rao (1973, pp.388-9) to state that:

> **Theorem 2:** $N^{1/2}\big(\hat{\boldsymbol{\Gamma}} - \boldsymbol{\Gamma}\big)$, as given by (9), has a $K$-variate normal limiting distribution with mean zero and covariance matrix $\boldsymbol{J \Omega J}^\top$. Here $\boldsymbol{\Omega}$ is the asymptotic covariance matrix of the vector $\hat{\boldsymbol{\Theta}}$, as given by Theorem 1 and its corollaries.

**Proof:** Standard. ∎

We have seen that all elements of $\boldsymbol{J \Omega J}^\top$ can be estimated consistently in a distribution-free manner, that is, without specifying an *a priori* functional form for the population distribution. We can then use the results of Theorem 2 to perform statistical inference on population progressivity, horizontal inequity, and the amount of redistribution effected by various taxes and benefits. We illustrate this in the following section.

## 4. Illustration: the Canadian Tax and Benefit System

Depending on whether we are interested in the measurement of progressivity, horizontal equity, or redistribution, $\gamma_p$ and $\delta_p$ in (8) are chosen to take the following forms:

IR Progressivity:

$$\gamma_p = E\big(M \mid X \leq G(p)\big) \qquad \delta_p = E\big(X \mid X \leq G(p)\big);$$

Comparisons of IR Progressivity:

$$\gamma_p = E\big(M_1 \mid X \leq G(p)\big) \qquad \delta_p = E\big(M_2 \mid X \leq G(p)\big);$$

TR Progressivity:

$$\gamma_p = E\big(X \mid X \leq G(p)\big) \qquad \delta_p = E\big(T \mid X \leq G(p)\big);$$

Comparisons of TR Progressivity:

$$\gamma_p = E\big(T_2 \mid X \leq G(p)\big) \qquad \delta_p = E\big(T_1 \mid X \leq G(p)\big);$$

Horizontal Inequity:

$$\gamma_p = E\big(M \mid X \leq G(p)\big) \qquad \delta_p = E\big(M \mid M \leq G^*(p)\big);$$

Redistribution:

$$\gamma_p = E\big(M \mid M \leq G^*(p)\big) \quad \delta_p = E\big(X \mid X \leq G(p)\big).$$

We illustrate this using micro-data from the 1981 and 1990 Canadian Surveys of Consumer Finances. The distribution of income in Canada was subjected to a number of shocks between these two years, a feature shared with many other countries. The Canadian fiscal system was also significantly altered in that decade.

The 1981 and 1990 Surveys contain, respectively, 37,779 and 45,461 observations on the distributions of incomes, income taxes, and a number of cash transfers. Families with negative gross or net incomes were removed. We use these data to compute gross incomes and the levels of personal income taxes and benefits. Combining taxes and benefits, we obtain the net effect of the entire system.[4]. The conditional expectations of the variables $Y$ and $V$ in (5) were estimated using robust Gaussian kernel estimation (Silverman (1986), p.45). To test whether curve A dominates curve B, we reject the null hypothesis of non-dominance in favour of the alternative hypothesis of dominance only if each point of curve A is found to be statistically greater than the corresponding point on curve B at a 5% level of significance. This procedure, defended by Howes (1993), ensures that the probability of type I errors is never greater than 5%.

Figure 1 shows the IR progressivity of taxes and benefits, measured as the distance between $L_M(p)$ and $L_X(p)$ as a function of $p$, for both 1981 and 1990. The graphs show the distance for each decile, with error bars of twice the estimated asymptotic standard error (not always visible if the estimated standard error is small enough). Since the 1981 and 1990 samples are independently distributed, the asymptotic standard errors of the differences across years in the IR progressivity ordinates at each decile are straightforward to calculate from the standard errors of each ordinate.[5]

[4] Detailed information on these transfers can be found, for instance, in Health and Welfare Canada (1992).

[5] If panel data were available, the standard errors would be based on the joint distribution obtained from paired observations.

It is clear from the figure that IR progressivity was significantly greater in 1990 than in 1981 after the second decile for taxes, and for all deciles for benefits.

If we wish to compare the progressivity of taxes and benefits for a given year, the comparison is complicated by the fact that the different ordinates, being estimated from the same sample, are therefore not independent. The comparisons, with appropriate standard errors that reflect this fact, are shown in Table 1. It is clearly possible here to declare income taxes statistically less IR progressive than the combination of benefits. The difference in IR progressivity is also statistically greater in 1990 than in 1981.

Next, we consider the differences among the Lorenz curves for gross and net incomes, and the concentration curve for net income, for each of the years 1981 and 1990. This allows us to measure the extent of the redistribution effected by the tax and benefit system, and the degree of reranking of after-tax incomes relative to pre-tax incomes. Figure 2, in which $L_M(p) - L^*(p)$ is plotted as a function of $p$, contains the information on reranking. The extent of reranking is precisely estimated, and reaches its maximum at the first decile. Information on redistribution is found in Figure 3, which plots the differences between the Lorenz curves for gross and net incomes as functions of $p$. The redistributive impact of the tax and benefit system is visibly highly significant: the inequality of net incomes is unambiguously lower than that of gross incomes. Redistribution is at its highest around the fifth decile, at which point 8.3% of total income is transferred in 1990 from the richer (than the fifth decile) to the poorer part of the population. Note also that the tax and benefit system raises almost eightfold the total income share of the poorest 10% of the population.

The curves for the two years, since they are based on independent samples, can be directly compared on the basis of the error bars. We find that the distribution of gross incomes in 1981 is unambiguously and significantly more equal than in 1990. The distribution of net incomes in 1981 is, however, almost unambiguously more unequal than the distribution of net incomes in 1990 (the difference at the ninth decile is not significant). Thus the 1990 tax and transfer system almost succeeds in making net incomes in 1990 unambiguously more equal than in 1981. It is therefore not surprising to find that the redistributive change in the inequality of incomes effected by the 1990 system is significantly greater than the change achieved by the 1981 system. At each of the deciles between the fourth and the eighth, for instance, 2% more of total income is redistributed from richer to poorer under the 1990 system than under the 1981 system, with a standard error of around 0.1%. Finally, it can easily be checked that reranking in 1990 is significantly greater than in 1981 at all deciles but the ninth.

## 5. Conclusion

We have established the asymptotic sampling distribution of quantile-based estimators computed from samples which need not be independent. The results are particularly useful for the measurement of progressivity, redistribution and horizontal inequity, and for the measurement and comparisons of inequality, welfare, or poverty that make use of estimated quantiles from possibly dependent samples. They also generalise most of the previous statistical inference results for the measurement of inequality and social welfare, and provide the statistical basis for the use of a number of popular indices.

Our illustrative application using the Canadian tax and benefit system shows that personal taxes and benefits are progressive, and that the separate tax and benefit components of the 1990 system are generally more progressive than those in 1981. Taxes are clearly statistically less progressive than benefits. Gross incomes are more equal in 1981 than in 1990, but net incomes are generally more equal in 1990 than in 1981. This is consistent with the finding that redistribution is significantly greater in 1990 than in 1981, with an associated increase in the extent of reranking, particularly at lower incomes.

## Appendix

**Proof of Theorem 1:**[6] Let the joint cumulative distribution function of $Z$ and $Y$ be denoted as $H(z, y)$, that is:

$$H(z, y) = \Pr\bigl(Z \le z \text{ and } Y \le y\bigr).$$

As in the text, $F$ denotes the c.d.f. of $Z$, supposed to be strictly monotonic and continuously differentiable, with inverse $G$. The quantity $\gamma_p$ of (2) can then be characterised by

$$p\gamma_p = \int_0^{G(p)} \int_0^{\infty} y \; d^2 H(z, y),$$

where the integral from 0 to $G(p)$ applies to $z$, and that from 0 to $\infty$ to $y$. The estimate $\hat{\gamma}_p$ of (4) is similarly given by

$$p\hat{\gamma}_p = \int_0^{\hat{G}(p)} \int_0^{\infty} y \; d^2 \hat{H}(z, y), \qquad (10)$$

where $\hat{G}$ is defined as before as the sample quantile function for $Z$, and $\hat{H}$ is the empirical distribution of $Z$ and $Y$ jointly.

[6] This proof is based on that found in the Appendix of Beach, Davidson, and Slotsve (1994), but extends it considerably.

The integral in (10) can be split up into an integral (over $z$) from 0 to $G(p)$ and another from $G(p)$ to $\hat{G}(p)$. The first is easy to deal with, as it can be written as a sum of i.i.d. variables:

$$\int_0^{G(p)} \int_0^{\infty} y \, d^2\hat{H}(z,y) = N^{-1} \sum_{i=1}^{N} Y_i \, I_{[0,G(p)]}(Z_i). \tag{11}$$

The second integral is of order $N^{-1/2}$, and it can be approximated as follows:

$$\begin{aligned}
\int_{G(p)}^{\hat{G}(p)} \int_0^{\infty} y \, d^2\hat{H}(y,z) &= \int_{G(p)}^{\hat{G}(p)} \int_0^{\infty} y \, d^2H(y,z) + O(N^{-1}) \\
&= \int_{G(p)}^{\hat{G}(p)} E\big(Y|\, z\big) \, dF(z) + O(N^{-1}).
\end{aligned}$$

The first equality follows from the root-$N$ consistency of $\hat{H}$ for $H$, and the second from the definition of the conditional expectation $E(Y|\, z)$. Since we assume that this conditional expectation is a smooth function of $z$, we have for $z \in [G(p), \hat{G}(p)]$ that

$$E\big(Y|\, z\big) = E\big(Y|\, Z = G(p)\big) + O(N^{-1/2}).$$

Thus, approximately with error of order only $N^{-1}$, the second integral is

$$E\big(Y|\, Z = G(p)\big) \int_{G(p)}^{\hat{G}(p)} d\hat{F}(z),$$

(using the root-$N$ consistency of $\hat{F}$), which can be rewritten as

$$\begin{aligned}
&-E\big(Y|\, Z = G(p)\big)\big(\hat{F}(G(p)) - p\big) \\
&\qquad = pE\big(Y|\, Z = G(p)\big) - \int_0^{G(p)} E\big(Y|\, Z = G(p)\big) \, d\hat{F}(z).
\end{aligned} \tag{12}$$

The expressions in (11) and (12) can be added together to obtain an asymptotic approximation for (10). To leading order,

$$\begin{aligned}
p\hat{\gamma}_p &= pE\big(Y|\, Z = G(p)\big) + \int_0^{G(p)} \int_0^{\infty} \Big(y - E\big(Y|\, Z = G(p)\big)\Big) \, d^2\hat{H}(z,y) \\
&= pE\big(Y|\, Z = G(p)\big) \\
&\quad + N^{-1} \sum_{i=0}^{N} \Big(Y_i - E\big(Y|\, Z = G(p)\big)\Big) \, I_{[0,G(p)]}(Z_i).
\end{aligned} \tag{13}$$

The expectation of $p\hat{\gamma}_p$ can readily be calculated from (13). To leading order:

$$\begin{aligned} E(p\hat{\gamma}_p) &= pE\big(Y|\, Z = G(p)\big) + E\Big(Y - E\big(Y|\, Z = G(p)\big)\ I_{[0,G(p)]}(Z)\Big) \\ &= pE\big(Y|\, Z = G(p)\big) + p\gamma_p - pE\big(Y|\, Z = G(p)\big) \\ &= p\gamma_p; \end{aligned}$$

which demonstrates the consistency of the estimator. Similarly, the fact that (13) and its analogue for $p\hat{\hat{\delta}}_p$ are sums of independent identically distributed random variables with finite second moments leads immediately to their asymptotic normality by the central limit theorem.

The covariance structure can now be obtained by simple calculation, based once more on the structure of (13) as a sum of i.i.d. variables. We have

$$\begin{aligned} \lim_{N\to\infty} N \operatorname{cov}\big(p\hat{\gamma}_p, p'\hat{\hat{\delta}}_{p'}\big) = \\ & E\bigg(\Big(Y - E\big(Y|\, Z = G(p)\big)\Big) I_{[0,G(p)]}(Z) \\ & \qquad \Big(V - E\big(V|\, W = G^*(p')\big)\Big) I_{[0,G^*(p')]}(W)\bigg) \\ & -E\Big(Y - E\big(Y|\, Z = G(p)\big)\ I_{[0,G(p)]}(Z)\Big) \\ & \qquad E\Big(V - E\big(V|\, W = G^*(p')\big)\ I_{[0,G^*(p')]}(W)\Big). \end{aligned} \tag{14}$$

Now by the definitions of $\gamma_p$ and $\delta_{p'}$ we have

$$E\Big(Y - E\big(Y|\, Z = G(p)\big)\ I_{[0,G(p)]}(Z)\Big) = p\Big(\gamma_p - E\big(Y|\, Z = G(p)\big)\Big) \text{ and}$$
$$E\Big(V - E\big(V|\, W = G^*(p')\big)\ I_{[0,G^*(p')]}(W)\Big) = p'\Big(\delta_{p'} - E\big(V|\, W = G^*(p')\big)\Big).$$

Thus the last term on the right-hand side of (14) equals the last term on the right-hand side of (5). The first term on the right-hand side of (14), when expanded, yields the other terms on the right-hand side of (5). ∎

## References


Aronson, J.R., P. Johnson and P .J. Lambert (1994). "Redistributive Effect and Unequal Income Tax Treatment." *Economic Journal*, **104(423)**, 262–270.

Atkinson, A.B.(1979). "Horizontal Equity and the Distribution of the Tax Burden" ch. 1 in *The Economics of Taxation*, ed. H.J. Aaron and M.J. Boskin, Brookings Institution, Washington DC, 3- 18.

Atkinson, A.B. (1987). "On the Measurement of Poverty." *Econometrica*, **55(4)**, 749–764.

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

Beach, C.M., R. Davidson, and G.A. Slotsve (1994). "Distribution-Free Statistical Inference for Inequality Dominance with Crossing Lorenz Curves." Queen's University Institute for Economic Research, Discussion Paper #912.

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

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

Duclos, J.-Y. (1993). "Progressivity, Redistribution and Equity, with Application to the British Tax and Benefit System." *Public Finance /Finances Publiques*, **48(3)**, 350-365.

Feldstein, M. (1976). "On the Theory of Tax Reform." *Journal of Public Economics*, **6**, 77-104.

Formby, J.P., W.J. Smith and P.D. Thistle (1990). "The Average Tax Burden and the Welfare Implications of Global Tax Progressivity." *Public Finance Quarterly*, **18(1)**, 3-24.

Health and Welfare Canada (1992). "Inventory of Income Security Programs in Canada." Department of Supply and Services.

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

Jakobsson, U. (1976). “On the Measurement of the Degree of Progression.” *Journal of Publics Economics*, **5(1-2)**, 161-168.

Kakwani, N.C. (1977). “Measurement of Tax Progressivity: An International Comparison.” *Economic Journal*, **87**, 71- 80.

Lambert, P. (1993). *The Distribution and Redistribution of Income: a Mathematical Analysis*, 2nd edition, Manchester University Press, Manchester, 306p.

Lerman, Robert I., and S. Yitzhaki (1995). “Changing Ranks and the Inequality Impacts of Taxes and Transfers.” *National Tax Journal*, **48(1)**, 45–59.

Musgrave, R.A. and T. Thin (1948). “Income Tax Progression 1929-48.” *The Journal of Political Economy*, **56(6)**, 498–514.

Pfähler, W. (1987). “Redistributive Effects of Tax Progressivity: Evaluating a General Class of Aggregate Measures.” *Public Finance/ Finances Publiques*, **42(3)**, 1-31.

Plotnick, R. (1981). “A Measure of Horizontal Inequity.” *The Review of Economics and Statistics*, **62(2)**, 283-288.

Plotnick, R. (1982). “The Concept and Measurement of Horizontal Inequity.” *Journal of Public Economics*, **17**, 373-391.

Rao, R.C. (1973). *Linear Statistical Inference and Its Applications*, John Wiley and Sons Inc., New York, 625 pages.

Reynolds, M., and E. Smolensky (1977). *Public Expenditure, Taxes and the Distribution of Income: The United States, 1950, 1961, 1970*, New York, Academic Press.

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

Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*, London, Chapman and Hall.

Suits, D.B. (1977). “Measurement of Tax Progressivity.” *American Economic Review*, **67**, 747-52.

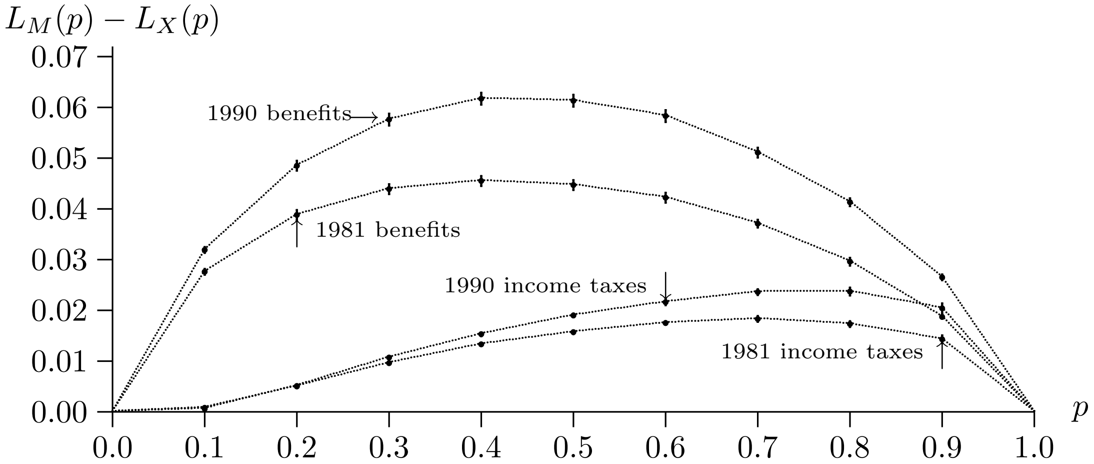


**Figure 1**

**IR Progressivity for 1981 and 1990 Taxes and Benefits**

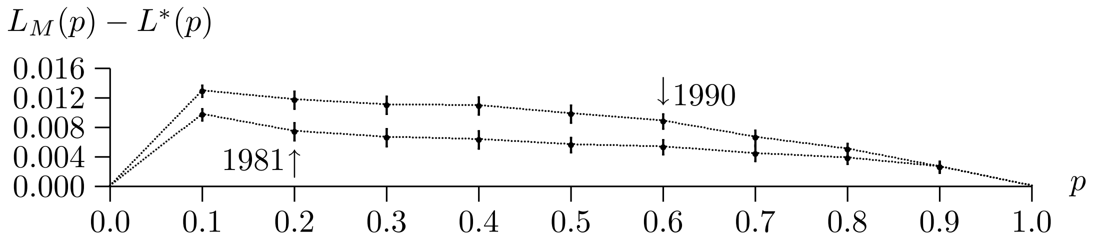


**Figure 2**

**Measure of Reranking**

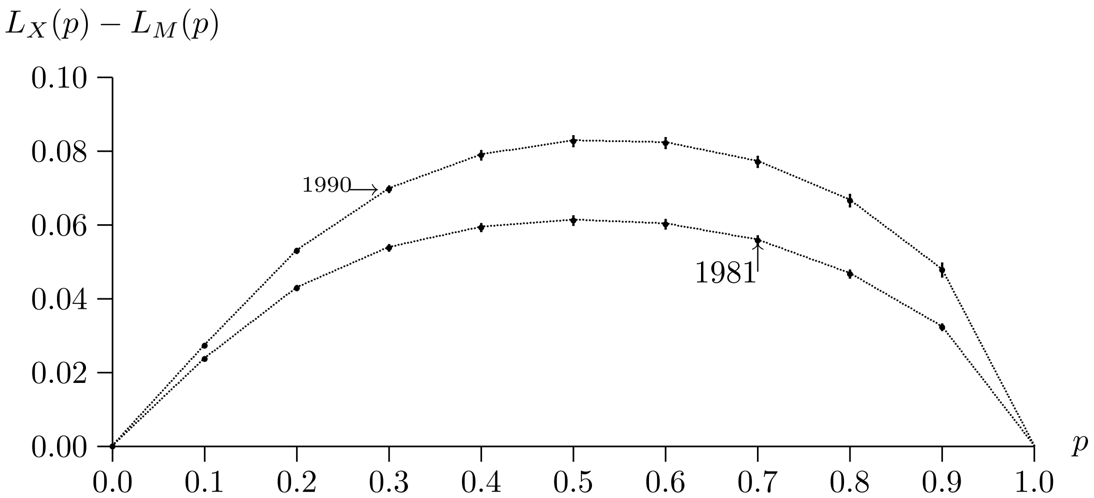


**Figure 3**

**Comparison of Lorenz Curves for Gross and Net Incomes**

**Table 1**

COMPARISONS OF IR PROGRESSIVITY
FOR 1981 AND 1990 TAXES AND BENEFITS

*asymptotic standard errors in italics*

| Deciles | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| Benefits *vs* Income taxes 1981 | 0.0268<br>*0.0006* | 0.0339<br>*0.0006* | 0.0342<br>*0.0007* | 0.0323<br>*0.0006* | 0.0290<br>*0.0006* | 0.0247<br>*0.0005* | 0.0188<br>*0.0012* | 0.0124<br>*0.0008* | 0.0044<br>*0.0004* |
| Benefits *vs* Income taxes 1990 | 0.0314<br>*0.0004* | 0.0434<br>*0.0007* | 0.0469<br>*0.0009* | 0.0464<br>*0.0008* | 0.0423<br>*0.0007* | 0.0366<br>*0.0008* | 0.0275<br>*0.0008* | 0.0176<br>*0.0007* | 0.0061<br>*0.0005* |