

Statistical inference with Generalized Gini indices of inequality and poverty

Garry Barrett

Department of Economics
University of New South Wales
Sydney, NSW, 2052

and

Stephen Donald*

Department of Economics
Boston University
Boston, MA

April 9, 2000

Abstract

This paper presents the asymptotic distribution of the generalized Gini indices of inequality, poverty and welfare. Unlike previous studies, our results do not require the grouping of the population into a fixed number of quantiles. Therefore our results permit statistical inference with consistent estimators of the Gini class of indices without compromising the normative properties of these indices. The methods are illustrated with an analysis of the trends in income inequality and poverty in the United States between 1978 and 1998.

*Stephen Donald thanks the Alfred P. Sloan Foundation for research support provided through a Research Fellowship.

1 Introduction

The work of Atkinson (1970), Kolm (1969) and Sen (1973) generated renewed interest in the theory of inequality measurement. Their research emphasized the duality between measures of inequality (and poverty) and social welfare functions. This insight led to a large literature which presented new indices of inequality explicitly derived from the desirable properties of the underlying social welfare function (see Lambert (1993) for a comprehensive survey). Recently there has been a growing body of research directed at determining the statistical properties of the normative measures of inequality and poverty so that they can be used for formal statistical inference. A key development in this line of research was the work of Beach and Davidson (1983) who derived the variance-covariance matrix of the vector of Lorenz curve ordinates without placing parametric restrictions on the form of the underlying income distribution. Subsequently, researchers have used the results in Beach and Davidson to derive the sampling distribution of inequality (Barrett and Pendakur, 1995; Rongve and Beach, 1997) and welfare indices (Bishop et.al., 1990) which can be expressed as functions of Lorenz curve ordinates or income quantiles.¹

However, an important limitation of this approach to deriving the sampling distribution of these Gini-related indices is that the indices are expressed as a function of a small, finite number of Lorenz curve ordinates, rather than the complete set of ordinates implied by an empirical distribution function. Consequently, the estimators of the inequality indices are inconsistent. By only considering a small number of ordinates, inequality within population quantiles is ignored and an underestimate of the true level of inequality is obtained. The inconsistency is greater the greater is the level of aggregation into quantiles (the smaller the set of ordinates) and the more transfer-sensitive is the measure (the more inequality averse the index). This line of research is therefore based on a compromise between using a desirable measure of inequality and being able to undertake formal statistical inference.² An exception is the work of Bishop et.al. (1997)

¹Another class of normative inequality and welfare indices can be expressed as functions of the negative and fractional raw moments of the income distribution. This class includes the Atkinson and generalised entropy indices: Thistle (1990) presents the sampling distribution of these indices.

²The methods proposed by Davidson and Duclos (1997) for conducting statistical inference with

who used results from the theory of U-statistics to derive the sampling properties of a consistent estimator of the Sen poverty index. Based on an alternative approach, the results presented in this paper generalize the result of Bishop et.al.(1997) to all members of the generalized Gini family of inequality, poverty and welfare indices. Our results show that the desirable normative properties of quantile based inequality, poverty and welfare indices need not be compromised in order to undertake statistical inference.

In this paper we examine the asymptotic distribution theory for a variety of inequality, poverty and welfare measures that can be written as statistical functionals³ whose argument is either the Lorenz Curve (LC) or Generalized Lorenz Curve (GLC). After presenting preliminary results concerning the standardized empirical LC or GLC, which are treated as stochastic processes, we use results concerning the delta method for statistical functionals to derive the asymptotic properties of the various inequality, poverty and welfare indices. Our approach to estimating standard errors is, like that of Beach and Davidson (1983), distribution free so that our methods, like theirs, can be considered to be nonparametric since we impose no structure on the underlying income distribution beyond some quite weak regularity conditions. Our approach to estimating standard errors differs from that of Beach and Davidson (1983) in that we use the influence functions for the various objects. As a by product we provide an alternative means for computing the variance covariance matrix for a vector of LC or GLC ordinates. The approach based on the influence function has considerable computational advantages when one considers consistent estimation of the asymptotic variance of the Gini-based inequality, poverty and welfare indices. Unlike previous research, we do not require that only a small set of Lorenz ordinates be used to compute the indices and we show that indeed our indices are consistent – all information concerning the distribution of income is used in our estimates.

The paper is organized as follows. The next section briefly reviews the set of inequality, poverty and welfare indices that can be written as functionals of the LC or GLC.

quantile-based indices of redistribution, progressivity and horizontal inequity also compromises estimator consistency in order to construct standard errors of the estimates.

³A statistical functional is a parameter which can be expressed as a function of a function. The theory of statistical inference for statistical functionals dates back to von Mises (1947) who developed a version of the delta method suitable to the task of deriving their asymptotic distribution.

In section 3 the sampling distribution of these measures are derived, and computational algorithms are briefly presented in section 4. In section 5 the methods are illustrated by examining changes in income inequality and poverty in the United States between 1978 and 1998. Concluding comments are provided in the final section.

2 Measures of Inequality and Poverty

2.1 Preliminaries

The objective is to undertake statistical inference with the normative indices of inequality and poverty which are generalizations of the Gini coefficient. Let Y denote the random variable income, and let the population c.d.f. be given by $F(y)$, which is assumed to be continuous and differentiable to at least second order. Also let $Q(p) = F^{-1}(p)$ denote the p th quantile of income. The LC for the distribution is then given by

$$\begin{aligned} L(p) &= \frac{1}{\mu} \int_0^{Q(p)} y \cdot dF(y) \\ &= \frac{1}{\mu} \int_0^p Q(t) dt \end{aligned}$$

where μ is the mean level of income. The Lorenz curve represents the proportion of total income going to the bottom p proportion of the population. Closely related to the LC is the GLC,

$$\begin{aligned} G(p) &= pE(Y|Y \leq Q(p)) \\ &= \int_0^{Q(p)} y \cdot dF(y) \\ &= \int_0^p Q(t) dt \end{aligned}$$

which is often used to compare income distributions and when the income distributions have unequal means and the differing mean is taken into account when judging social welfare (Shorrocks, 1983). These curves are popular tools in the literature on inequality and welfare measurement and, as shown below, a variety of indices of inequality and poverty can be represented as functionals of either the GLC or the LC.

2.2 S-Gini indices (Donaldson and Weymark,1980,1983; Kakwani, 1980; Yitzhaki, 1983)

The S-Gini relative indices of inequality are given by

$$\begin{aligned}
 I_R^\delta &= 1 - \frac{1}{\mu} \int_{-\infty}^{\infty} \delta \cdot [1 - F(y)]^{\delta-1} f(y) y dy \\
 &= 1 - \frac{1}{\mu} \int_0^1 \delta (1-p)^{\delta-1} Q(p) dp \\
 &= 1 - \delta(\delta-1) \int_0^1 (1-p)^{\delta-2} L(p) d(p)
 \end{aligned} \tag{1}$$

where the last expression is valid only when $1 < \delta < \infty$. In the case where $\delta = 1$, the index is identically zero. The index is scale free (homogeneous of degree 0 in y) and S-Concave (see Blackorby and Donaldson, 1978). The S-Gini indices are ‘ethically flexible’ in that $\delta \geq 1$ is the inequality aversion parameter, set by the researcher, which determines the social weight attached to different points in the distribution. For values of $\delta \in (1, 2)$, the indices place relatively greater weight on individuals ranked at the top of the income distribution. When $\delta = 2$ the index corresponds to the popular Gini coefficient, and as δ increases towards ∞ progressively more weight is placed on Lorenz ordinates at the lower end of the distribution. In the limit, as $\delta \rightarrow \infty$, all the social weight is focused on the income share of the poorest individual and the index is equal to the maximin index of inequality.

Geometrically, the Gini coefficient is equal to (twice) the area between the Lorenz curve and the line of equality. More generally, the difference in the value of an S-Gini index for two income distributions corresponds to the weighted integral of the area between the Lorenz curve and the line of equality, with the weight determined by δ .

The S-Gini absolute index of inequality is given by

$$\begin{aligned}
 I_A^\delta &= \mu - \int_{-\infty}^{\infty} \delta \cdot [1 - F(y)]^{\delta-1} f(y) y dy \\
 &= \mu - \int_0^1 \delta (1-p)^{\delta-1} Q(p) dp
 \end{aligned} \tag{2}$$

$$= \mu - \delta(\delta - 1) \int_0^1 (1 - p)^{\delta-2} G(p) dp$$

where $G(p) = \mu.L(p) = \int_0^{Q(p)} y dF(y)$, represents the Generalized Lorenz Curve. Similar to the relative index, the last expression is valid only when $1 < \delta < \infty$ and the index is identical to the mean when $\delta = 1$. Absolute inequality indices are a function of absolute income differences, or income differentials, rather than income shares, and have been interpreted as measures of relative deprivation (Yitzhaki, 1979). The S-Gini absolute indices are S-Concave, invariant to equal additions to income (translation free) and invariant to arbitrary replications of the income vector (see Donaldson and Weymark, 1980; Blackorby and Donaldson, 1980a).

The S-Gini inequality indices imply, and are implied by, the same class of ordinally equivalent SWFs. One cardinal representation of the welfare index underlying the S-Gini inequality indices is

$$\begin{aligned} W^\delta &= \int_{-\infty}^{\infty} \delta.[1 - F(y)]^{\delta-1} f(y) y dy \\ &= \int_0^1 \delta(1 - p)^{\delta-1} Q(p) dp \\ &= \delta(\delta - 1) \int_0^1 (1 - p)^{\delta-2} G(p) dp \end{aligned} \tag{3}$$

The specific case of $\delta = 2$ corresponds to the ‘Sen welfare index’. The sampling distribution of this particular index was analyzed by Bishop et.al.(1989): the propositions presented in the following section generalize their result to consistent estimators of all members of the Gini family of welfare indices.

2.3 E-Gini indices (Chakravarty, 1988)

Another set of generalized Gini indices are the E-Ginis presented by Chakravarty (1988). The E-Gini relative index is

$$I_R^\alpha = 2 \left[\int_0^1 (p - L(p))^\alpha dp \right]^{\frac{1}{\alpha}} \tag{4}$$

where $\alpha \geq 1$ is the inequality aversion parameter. As α increases the index is increasingly sensitive to transfers at the lower end of the income distribution. When $\alpha=1$ the index corresponds to the Gini coefficient, and as $\alpha \rightarrow \infty$ the index approaches twice the Schutz index (the maximum distance between the LC and line of equality).

The S-Gini and E-Gini inequality indices have similar normative properties. The key distinction is that only the E-Ginis (for $\alpha > 1$) satisfy the ‘principle of diminishing transfers’. The S-Ginis are linear in incomes, and the social weight attached to individuals is a function of their rank in the ordered distribution. The effect of a transfer on the level of measured inequality will be a function of the ranks of the individuals’ involved in the transfer. However, the E-Gini are a CES function of the difference between the LC and the line of equality, and the effect of a transfer on measured inequality will be sensitive to the differences in individuals’ income shares (see Chakravarty, 1988, for a proof and fuller discussion of this property).

The E-Gini absolute index of inequality is given by

$$\begin{aligned} I_A^\alpha &= 2\mu([\int_0^1 (p - L(p))^\alpha dp]^\frac{1}{\alpha}) \\ &= 2[\int_0^1 (\mu \cdot p - G(p))^\alpha dp]^\frac{1}{\alpha} \end{aligned} \tag{5}$$

Underlying the E-Gini relative and absolute inequality indices is a common class of SWFs. One cardinalization of the E-Gini welfare index is

$$\begin{aligned} W^\alpha &= 2\mu[1 - \int_0^1 (p - L(p))^\alpha dp]^\frac{1}{\alpha} \\ &= 2(\mu - [\int_0^1 (\mu \cdot p - G(p))^\alpha dp]^\frac{1}{\alpha}) \end{aligned} \tag{6}$$

2.4 Gini Based Poverty Indices

Indices of inequality and welfare are closely related to measures of poverty (see Sen, 1976). Following Blackorby and Donaldson (1980b), normative poverty indices can be expressed as a composite function of the proportion of the population below the poverty line (the

head-count ratio), the average income of the poor and the level of income inequality among the poor:⁴

$$\begin{aligned} P &= F(z) \left(1 - \frac{\mu(z)(1 - I_R(y|y < z))}{z} \right) \\ &= F(z) \left(\frac{z - W(y|y < z)}{z} \right) \end{aligned} \quad (7)$$

where z is the income level representing the ‘poverty line’, $W(y|z)$ represents the social welfare of the poor which consists of individuals with income values less than z and $\mu(z)$ is the average income of the poor. Equivalently, the generalized Gini poverty indices can be expressed as a composite function of the proportion of the population below the poverty line ($F(z)$), and the normalised Gini welfare indices defined over the poor. Note that since these welfare indices are defined over the poor, then, for the S-Gini welfare index, when $\delta > 1$,

$$W^\delta(y|z) = \int_{-\infty}^z \delta \left[1 - \frac{F(y)}{F(z)} \right]^{\delta-1} \frac{f(y)}{F(z)} y dy \quad (8)$$

$$= \int_0^1 \delta (1-p)^{\delta-1} Q(pF(z)) dp \quad (9)$$

$$= \frac{1}{F(z)} \delta(\delta-1) \int_0^1 (1-p)^{\delta-2} G(pF(z)) dp$$

and when $\delta = 1$,

$$\begin{aligned} W^\delta(y|z) &= E(Y|Y \leq z) \\ &= \frac{E(Y \cdot 1(Y \leq z))}{F(z)} \end{aligned}$$

Consequently the poverty index based on the S-Gini has the form,

$$P^\delta = F(z) - \frac{1}{z} \delta(\delta-1) \int_0^1 (1-p)^{\delta-2} G(pF(z)) dp$$

⁴Blackorby and Donaldson (1980b) refer to this set of poverty indices as relative poverty indices since they are invariant to a common positive scaling of all incomes and the poverty lines. Blackorby and Donaldson (1980b) also propose a set of absolute poverty indices which are a function of the number of poor and the absolute shortfall of the conditional welfare function from the poverty line. The estimator of the generalized Gini absolute poverty indices is equal to the relative poverty indices discussed in this section multiplied by the number of poor, $N(z)$, and the poverty line z . Similarly, the standard errors for these indices will be equal to the standard errors for the relative indices multiplied by $N(z)z$.

when $\delta > 1$, and,

$$P^\delta = \frac{E((z - Y) \cdot 1(Y \leq z))}{z}$$

when $\delta = 1$. The S-Gini poverty index in the case of $\delta = 1$ is equal to the head-count ratio times the average income shortfall of the poor divided by the poverty line. As discussed below, this particular index can be simply estimated using a scaled average, making estimation and inference trivial. Note also that when $\delta = 2$ the S-Gini poverty index corresponds to the well-known Sen index of poverty; the sampling distribution of this specific index was analysed by Bishop et.al. (1997).

When the welfare index underlying the E-Gini indices is adopted the conditional social welfare of the poor is given by,

$$W^\alpha = 2(\mu(z) - \frac{1}{F(z)} [\int_0^1 (\mu(z) \cdot pF(z) - G(pF(z)))^\alpha dp]^{\frac{1}{\alpha}}) \quad (10)$$

where, $\mu(z) = E(Y|Y \leq z)$ so that the E-Gini poverty based index is given by the function,

$$P^\alpha = F(z) - \frac{2}{z}G(F(z)) + \frac{1}{z} [\int_0^1 (pG(F(z)) - G(pF(z)))^\alpha dp]^{\frac{1}{\alpha}}$$

where we have used the fact that $G(F(z)) = \mu(z)F(z)$.

3 Asymptotic Distribution of the Generalized Gini Estimators

3.1 Preliminaries

We begin by defining some notation, providing preliminary results and then giving a brief overview of the approach taken in this paper. Following Rongve and Beach (1997) we assume that income, Y , is distributed on the interval $[y_l, y_u]$ with $0 < y_l < y_u < \infty$. We denote the cumulative distribution function of Y by $F(\cdot)$ and the probability density function of Y by $f(\cdot)$. We make the following assumption regarding the distribution of income. Unless stated otherwise, all sup's and inf's will be over the set $[y_l, y_u]$.

Assumption 1 Assume that $F : [y_l, y_u] \rightarrow [0, 1]$ is twice continuously differentiable with $f(y) = F'(y)$, where $0 < \inf f(y) < \sup f(y) < \infty$.

For a random sample of incomes drawn from F , denoted by $\{Y_i\}_{i=1}^N$, we denote by $\hat{F}(y)$ the empirical distribution function of income given by,

$$\hat{F}(y) = \frac{1}{N} \sum_{i=1}^N 1(Y_i \leq y)$$

where $1(A)$ denotes the indicator function for the event A . We first state a well known preliminary result concerning the statistical properties of the Empirical Distribution Function. The notation “ \Rightarrow ” used here and elsewhere (unless stated otherwise) refer to weak convergence in the set of functions that are specified. The space $D(X)$ is the notation for the space of cadlag functions⁵ on the set $X \subset R$. Therefore the function, $\sqrt{N}(\hat{F}(y) - F(y)) \in D[y_l, y_u]$. We also use the notation \mathcal{B} to refer to the Brownian Bridge process on $[0, 1]$ which is a mean zero Gaussian process⁶ with covariance kernel given by,

$$Cov(\mathcal{B}(p)\mathcal{B}(q)) = \min\{p, q\} - pq$$

Lemma 1: *Given Assumption 1, the following results hold,*

(i) $\sup |\hat{F}(y) - F(y)| \xrightarrow{a.s.} 0$

(ii) $\sqrt{N}(\hat{F} - F) \Rightarrow \mathcal{B} \circ F$ in $D[y_l, y_u]$

The first two results are well known results for the empirical distribution function. Note that the notation $\mathcal{B} \circ F$ is the composition of \mathcal{B} with F and hence refers to the process defined on $[y_l, y_u]$ whose value at $y \in [y_l, y_u]$ is $\mathcal{B}(F(y))$.

The objects that are considered in this paper are statistical functionals in the sense that they depend on the function F and not just on the values of F at a fixed number of values for y . Thus the objects of interest can be treated as being of the form $T(F)$

⁵A function is cadlag if it is continuous from the right with limits from the left.

⁶Note that $\mathcal{B}(p) = \mathcal{W}(p) - p\mathcal{W}(1)$ where $\mathcal{W}(p)$ is the usual Wiener process on $[0, 1]$.

where $T : \mathcal{F} \subset D[y_l, y_u] \rightarrow \mathcal{V}$ where \mathcal{F} is a set of functions to which F belongs (a subset of $D[y_l, y_u]$) and \mathcal{V} is some other set which could be a set of functions (if the functional defines a function) or a set of real numbers (if the functional is real valued).

The key to deriving results for the normalized functional $\sqrt{N}(T(\hat{F}) - T(F))$ is showing that the map T is Hadamard differentiable at F . The strict definition of this given by Van der Vaart and Wellner (1994) is as follows.

Definition 1: *The map $T : \mathcal{F} \subset D[y_l, y_u] \rightarrow \mathcal{V}$ is Hadamard differentiable at F tangentially to the set \mathcal{F}_0 if there is a continuous linear map $T'_F : D[y_l, y_u] \rightarrow \mathcal{V}$ such that,*

$$\frac{T(F + t_n h_n) - T(F)}{t_n} - T'_F(h_n) \rightarrow 0$$

for all converging sequences $t_n \rightarrow 0$, $h_n \rightarrow h \in \mathcal{F}_0$ with $F + t_n h_n \in \mathcal{F}$.

Typically in this definition the set \mathcal{F}_0 will be the set to which $\mathcal{B} \circ F$ takes its values which is the set of continuous functions on $[y_l, y_u]$. Then once one has established Hadamard differentiability the functional delta method (see Theorem 3.9.4 of Van der Vaart and Wellner, 1994) implies the following result holds:

$$\sqrt{N}(T(\hat{F}) - T(F)) \Rightarrow T'_F(\mathcal{B} \circ F)$$

This result turns out to be a useful way of characterizing the limiting distribution of a number of the objects considered in this paper. As a practical matter the Hadamard derivative of $T(F)$ can be found by the following calculation:

$$T'_F(H) = \left. \frac{d}{dt} T(F + tH) \right|_{t=0}$$

When the functional $T(F)$ is real valued and Hadamard differentiable then the above delta method implies that the limiting random variable is normally distributed. In order to make inferences one must estimate the variance for which a useful tool is the influence function or influence curve which gives the effect of an observation on an estimate. In many estimation problems the variance of an estimator is equal to the expectation of the squared influence function and this can be estimated by taking the sample average of the

squared influence functions. Note that for the estimator $\hat{F}(y)$ the influence function is given by,

$$\phi_i(y; F) = 1(Y_i \leq y) - F(y)$$

and that,

$$V(\hat{F}(y)) = E(\phi_i(y; F)^2) = F(y)(1 - F(y))$$

is the (asymptotic) variance of $\hat{F}(y)$. Moreover one can estimate the variance of $\hat{F}(y)$ by taking,

$$\hat{V}(\hat{F}(y)) = \frac{1}{N} \sum_{i=1}^N \hat{\phi}_i(y; F)^2$$

where $\hat{\phi}_i(y; F) = 1(Y_i \leq y) - \hat{F}(y)$ is the estimated influence function.

Fernholz (1983, Theorem 4.4.2) has shown that one can compute asymptotic variances using influence functions when $T(F)$ is a real valued Hadamard differentiable mapping from $[y_l, y_u]$ to R with non-zero Hadamard derivative. Because the calculation of an influence curve uses basic rules of calculus, this fact provides a convenient route for finding estimates of asymptotic variances. For the functional $T(F)$ the influence curve is,

$$\phi_i(\cdot; T) = \frac{d}{dt} T(F + t(\delta_{Y_i} - F))|_{t=0}$$

where the function $\delta_i(y) = 1(y \geq Y_i)$ is the distribution function of the point mass one at Y_i (see Fernholz (1983, p. 8) and Huber (1981, p.13)). The notation $\phi_i(\cdot; T)$ is to allow for the fact that the functional may be evaluated at a particular point, which is the first argument inside the parentheses, as would be the case for F evaluated at a point in $[y_l, y_u]$. If such an argument is not needed as in the case of the inequality indices then the influence function will be represented as simply $\phi_i(T)$. Then the variance of the functional can be characterized as,

$$V(T(\hat{F})) = E(\phi_i(\cdot; T)^2)$$

and can be estimated by replacing the unknowns with estimates and taking the sample average,

$$\hat{V}(T(\hat{F})) = \frac{1}{N} \sum_{i=1}^N \hat{\phi}_i(\cdot; T)^2$$

This method will be used repeatedly as a way of estimating variances of the objects of interest.

3.2 Statistical Properties of Quantiles and Lorenz Curves

In this section we present a number of preliminary results relating to the asymptotic properties of quantiles and the LC and GLC. These results will then be used subsequently to derive the asymptotic properties of the inequality and poverty indices. In dealing with quantiles we first note that Assumption 1 implies the existence of the inverse function, or quantile of income function $Q(p) = F^{-1}(p)$ which is mapping from $[0, 1] \rightarrow [y_l, y_u]$. Also, define the empirical counterpart of the quantile function, by,

$$\hat{Q}(p) = \inf\{y : \hat{F}(y) \geq p\}$$

If there are no ties among the incomes in the sample then $\hat{Q}(p)$ is a step function that takes on the values equal to the order statistics of income over intervals of length $1/N$.

The Generalized Lorenz Curve (hereafter GLC), at ordinate value p , can be defined in terms of the quantile function by,

$$G(p) = \int_0^p Q(t)dt$$

and the Lorenz Curve (LC) itself at p is given by,

$$L(p) = \frac{G(p)}{G(1)} = \frac{G(p)}{\mu}.$$

Note that for the second equality we have used the fact that $G(1) = E(Y) = \mu$, the overall mean level of income in the population. Also note that $L(p)$ gives the proportion of income possessed by those individuals with income level less than y_p where $Q(p) = y_p$. Also, note that $G(p) = pE(Y|Y \leq y_p)$ which is the proportion of the population with income below y_p times the conditional mean income given that income is below y_p .

Using our empirical quantile function we can define the empirical Generalized Lorenz curve (or function) and Lorenz curve as respectively,

$$\hat{G}(p) = \int_0^p \hat{Q}(t) dt$$

and,

$$\hat{L}(p) = \frac{\hat{G}(p)}{\hat{\mu}}.$$

where $\hat{\mu} = \hat{G}(1) = \bar{Y}$ is the sample mean of the Y_i . For the moment we have defined these functions as integrals. This will suffice for much of the statistical analysis, although it is not very helpful as far as actual computation is concerned. We discuss computational issues below, although it should be noted that since the empirical quantile function is a step function the integrals can easily be calculated as a sum.

Having defined the empirical GLC and LC we can then define the inequality indices in terms of either function. Before doing so, however, we provide a few preliminary results concerning the statistical properties of the empirical quantile function, and the GL and Lorenz curves. We first state a preliminary result concerning the statistical properties of the Quantile function.

Lemma 2: *Given Assumption 1, the following results hold,*

(i) $\sup |\hat{Q}(p) - Q(p)| \xrightarrow{a.s.} 0$

(ii) *in $l^\infty(0, 1)$ (the space of uniformly bounded real functions on $(0, 1)$) we have,*

$$\sqrt{N}(\hat{Q} - Q) \Rightarrow -\frac{\mathcal{B}}{f(Q)}$$

These results combine the strong uniform convergence result contained in Corollary 1.4.1 of Csorgo (1983) and a weak convergence result such as shown in Van der Vaart and Wellner (1996, page 387). Note that for a fixed point p the weak convergence result implies that,

$$\sqrt{N}(\hat{Q}(p) - Q(p)) \xrightarrow{d} N\left(0, \frac{p(1-p)}{f(Q(p))^2}\right)$$

which is a well known result for empirical quantiles. There is a large literature on the Quantile processes – see Csorgo (1983) for more details.

Although we will not be interested in estimating the variance of the quantiles *per se* the influence function for the quantile function will be quite useful when we deal with Lorenz curves and the inequality indices. As shown by Huber (1981, p. 56) the influence function at a point p is,

$$\phi_i(p; Q) = 1(Y_i < Q(p)) \frac{p-1}{f(Q(p))} + 1(Y_i > Q(p)) \frac{p}{f(Q(p))}$$

As is well known it is necessary to use some type of density estimation when estimating the variance of the quantile. Fortunately this will not be required for the functionals considered in this paper. We next turn to the straightforward derivation of weak convergence results for the empirical GLC and LC which can be shown to be piecewise linear and continuous. Consequently we treat these objects as elements of $C[0, 1]$, the set of all continuous functions on $[0, 1]$. Define the Gaussian stochastic process, \mathcal{G} on $[0, 1]$ to be such that for $p \in [0, 1]$,

$$\mathcal{G}(p) = - \int_0^p \frac{\mathcal{B}(t)}{f(Q(t))} dt$$

Lemma 3 *Given Assumption 1 we have that,*

(i) *for the GLC,*

$$\sup |\hat{G}(p) - G(p)| \xrightarrow{a.s.} 0$$

and in the space $C[0, 1]$,

$$\sqrt{N}(\hat{G} - G) \Rightarrow \mathcal{G}$$

(ii) *for the LC we have that,*

$$\sup |\hat{L}(p) - L(p)| \xrightarrow{a.s.} 0$$

and in the space $C[0, 1]$,

$$\sqrt{N}(\hat{L} - L) \Rightarrow \frac{\mathcal{G}}{\mu} - \frac{L}{\mu} \mathcal{G}(1) \equiv \mathcal{L}.$$

Notice that the stochastic processes \mathcal{G} and \mathcal{L} are both mean zero Gaussian stochastic processes so that the result implies that a normalized vector of LC (or GLC) ordinates is asymptotically multivariate normal as suggested by the results derived in Beach and Davidson (1983). The results in Lemma 3 are not new and date back to at least Goldie (1977), who presented a full weak convergence result for the LC process under very weak conditions. Our proof of the results of Lemma 3 is somewhat simpler than that of Goldie (1977) since we take as a starting point the result in Lemma 2 which requires slightly stronger assumptions than required by the method of Goldie (1977). Other results concerning the empirical LC process include Gail and Gastwirth (1978) who derived an asymptotic distribution result for a single ordinate of the normalized LC and Csörgő (1983) who proved that the empirical LC process could be strongly approximated by a sequence of Gaussian processes which are equal in distribution to that given above.

As shown in Beach and Davidson (1983) the covariances of the limiting stochastic process are given by, respectively, (for $p \leq q$),

$$\begin{aligned} V_G(p, q) &= E(\mathcal{G}(p)\mathcal{G}(q)) \\ &= p(\sigma_p^2 + (1 - q)(y_p - \frac{G(p)}{p})(y_q - \frac{G(q)}{q}) \\ &\quad + (y_p - \frac{G(p)}{p})(\frac{G(q)}{q} - \frac{G(p)}{p})) \end{aligned}$$

and,

$$\begin{aligned} V_L(p, q) &= E(\mathcal{L}(p)\mathcal{L}(q)) \\ &= \frac{1}{\mu^2}V^G(p, q) + \frac{L(p)L(q)}{\mu^2}\sigma_1^2 - \frac{L(p)}{\mu^2}V^G(q, 1) - \frac{L(q)}{\mu^2}V^G(p, 1) \end{aligned}$$

where $\sigma_p^2 = V(Y|Y \leq y_p)$. The variances are then simply,

$$V_G(p) = p \left(\sigma_p^2 + (1 - p)(y_p - \frac{G(p)}{p})^2 \right)$$

and,

$$V_L(p) = \frac{1}{\mu^2}V_G(p) + \frac{L(p)^2}{\mu^2}\sigma_1^2 - 2\frac{L(p)}{\mu^2}V_G(p, 1)$$

Beach and Davidson (1983) suggested variance estimates based on replacing unknowns with estimates. It turns out to be convenient for what follows to represent the variances and covariances in terms of the influence function for the respective process. The influence curves for the LC and GLC can be obtained quite simply using the influence curve for the quantile function and standard rules of calculus.

Lemma 4: *The influence curve for the GLC and LC are respectively,*

$$\phi_i(p; G) = (pQ(p) - G(p)) - 1(Y_i < Q(p))(Q(p) - Y_i)$$

and

$$\phi_i(p; L) = \frac{1}{\mu} \phi_i(p; G) - \frac{L(p)}{\mu} (Y_i - \mu)$$

Note that the influence curves can be defined quite compactly which is convenient for both presentation and computational purposes. As a final set of preliminary results we consider the consistent estimation of the variance function for the GLC and LC at a point p . Let $\hat{\phi}_i(p; G)$ and $\hat{\phi}_i(p; L)$ denote the influence curves presented in Lemma 4 with the unknowns $Q(p)$, $G(p)$, μ and $L(p)$ replaced with the consistent estimates that were defined above. Then let,

$$\hat{V}_G(p) = \frac{1}{N} \sum_{i=1}^N \hat{\phi}_i(p; G)^2$$

and,

$$\hat{V}_L(p) = \frac{1}{N} \sum_{i=1}^N \hat{\phi}_i(p; L)^2$$

Lemma 5: *Given Assumption 1 the following hold for a fixed value of p ,*

$$\hat{V}_G(p) \xrightarrow{p} V_G(p)$$

and,

$$\hat{V}_L(p) \xrightarrow{p} V_L(p)$$

3.3 Asymptotic Distribution of Inequality Indices

Now we are in a position to derive the limiting distribution of the inequality and welfare indices discussed in Section 2. As was shown in Section 2 the inequality indices can all be expressed as functionals of F, Q, G and L . This differs from other approaches to statistical analysis such as in Rongve and Beach (1997) and Davidson and Duclos (1997) which are based on the standard Delta method and hence can only be applied for a fixed number of ordinates of F, G and L . It should also be noted that our results will show that even if all of the ordinates are used, one still obtains root-N consistency and asymptotic normality.

3.3.1 S-Gini indices

For the S-Gini relative index of inequality it is most convenient to define the estimator in terms of the empirical Lorenz curve estimator,

$$\hat{I}_R^\delta = 1 - \delta(\delta - 1) \int_0^1 (1 - p)^{\delta-2} \hat{L}(p) dp$$

and to define the S-Gini absolute index of inequality estimator in terms of the GL curve estimator,

$$\hat{I}_A^\delta = \hat{\mu} - \delta(\delta - 1) \int_0^1 (1 - p)^{\delta-2} \hat{G}(p) dp$$

Finally denote the empirical version of the associated welfare index underlying the S-Ginis by,

$$\hat{W}^\delta = \delta(\delta - 1) \int_0^1 (1 - p)^{\delta-2} \hat{G}(p) dp$$

The following result provides a characterization of the limiting distribution of these three objects.

Proposition 1(A): *Given Assumption 1, the following results holds for a fixed value of δ , such that $1 < \delta < \infty$,*

(i) $\hat{I}_R^\delta \xrightarrow{p} I_R^\delta$ and

$$\sqrt{N}(\hat{I}_R^\delta - I_R^\delta) \Rightarrow -\delta(\delta - 1) \int_0^1 (1 - p)^{\delta-2} \mathcal{L}(p) dp \sim N(0, V(\hat{I}_R^\delta))$$

(ii) $\hat{I}_R^\delta \xrightarrow{p} I_R^\delta$ and

$$\sqrt{N}(\hat{I}_A^\delta - I_A^\delta) \Rightarrow \mathcal{G}(1) - \delta(\delta - 1) \int_0^1 (1-p)^{\delta-2} \mathcal{G}(p) dp \sim N(0, V(\hat{I}_A^\delta))$$

(iii) $\hat{W}^\delta \xrightarrow{p} W^\delta$ and

$$\sqrt{N}(\hat{W}^\delta - W^\delta) \Rightarrow \delta(\delta - 1) \int_0^1 (1-p)^{\delta-2} \mathcal{G}(p) dp \sim N(0, V(\hat{W}^\delta)).$$

Note that in addition to all of the estimators being consistent, they also have limiting normal distributions. This follows from the fact that the distributions are all linear functionals of Gaussian stochastic processes. This fact makes it quite straightforward to derive the influence curves for the different indices in terms of the influence curves for the GLC and LC. In particular let,

$$\begin{aligned} \phi_i(W^\delta) &= \delta(\delta - 1) \int_0^1 (1-p)^{\delta-2} \phi_i(p; G) dp \\ \phi_i(I_A^\delta) &= (Y_i - \mu) - \phi_i(W^\delta) \\ \phi_i(I_R^\delta) &= \frac{1}{\mu} \phi_i(I_A^\delta) - \frac{I_R^\delta}{\hat{\mu}} (Y_i - \hat{\mu}) \end{aligned}$$

Then the variances can be estimated using,

$$\begin{aligned} \hat{V}(\hat{W}^\delta) &= \frac{1}{N} \sum_{i=1}^N \hat{\phi}_i(W^\delta)^2 \\ \hat{V}(\hat{I}_A^\delta) &= \frac{1}{N} \sum_{i=1}^N \hat{\phi}_i(I_A^\delta)^2 \\ \hat{V}(\hat{I}_R^\delta) &= \frac{1}{N} \sum_{i=1}^N \hat{\phi}_i(I_R^\delta)^2 \end{aligned}$$

where as usual the $\hat{\phi}$ are ϕ with unknowns replaced by the relevant estimates. Appendix B indicates how these influence functions can be estimated for a particular sample.

Proposition 1(B): *Given the conditions of Proposition 1(A),*

$$\hat{V}(\hat{I}_R^\delta) \xrightarrow{p} V(\hat{I}_R^\delta),$$

$$\hat{V}(\hat{I}_A^\delta) \xrightarrow{p} V(\hat{I}_A^\delta)$$

$$\hat{V}(\hat{W}^\delta) \xrightarrow{p} V(\hat{W}^\delta)$$

3.3.2 E-Gini indices

For the E-Gini relative index estimator we have,

$$\hat{I}_R^\alpha = 2\left[\int_0^1 (p - \hat{L}(p))^\alpha dp\right]^{\frac{1}{\alpha}}$$

while it is convenient to define the E-Gini absolute index estimator in terms of the empirical GL curve as,

$$\hat{I}_A^\alpha = 2\left[\int_0^1 (\hat{\mu}\cdot p - \hat{G}(p))^\alpha dp\right]^{\frac{1}{\alpha}}$$

and finally the empirical E-Gini welfare index is given by,

$$\hat{W}^\alpha = 2(\hat{\mu} - \left[\int_0^1 (\hat{\mu}\cdot p - \hat{G}(p))^\alpha dp\right]^{\frac{1}{\alpha}})$$

Our result concerning these estimators is contained below in Proposition 2(A).

Proposition 2(A) *Given Assumption 1, and assuming that I_R^α , I_A^α and W^α are all strictly positive then the following results hold,*

(i) $\hat{I}_R^\alpha \xrightarrow{p} I_R^\alpha$ and,

$$\begin{aligned} \sqrt{N}(\hat{I}_R^\alpha - I_R^\alpha) &\Rightarrow -2^\alpha (I_R^\alpha)^{1-\alpha} \int_0^1 (p - L(p))^{\alpha-1} \mathcal{L}(p) dp \\ &\sim N(0, V(\hat{I}_R^\alpha)), \end{aligned}$$

(ii) $\hat{I}_A^\alpha \xrightarrow{p} I_A^\alpha$ and,

$$\begin{aligned} \sqrt{N}(\hat{I}_A^\alpha - I_A^\alpha) &\Rightarrow 2^\alpha (I_A^\alpha)^{1-\alpha} \int_0^1 (\mu p - G(p))^{\alpha-1} (p\mathcal{G}(1) - \mathcal{G}(p)) dp \\ &\sim N(0, V(\hat{I}_A^\alpha)), \end{aligned}$$

(iii) $\hat{W}^\alpha \xrightarrow{p} W^\alpha$ and,

$$\begin{aligned} \sqrt{N}(\hat{W}^\alpha - W^\alpha) &\Rightarrow 2\mathcal{G}(1) - 2^\alpha (I_A^\alpha)^{1-\alpha} \int_0^1 (\mu p - G(p))^{\alpha-1} (\mathcal{G}(p) - p\mathcal{G}(1)) dp \\ &\sim N(0, V(\hat{W}^\alpha)). \end{aligned}$$

Again the variances can be expressed in terms of the influence functions which are,

$$\begin{aligned}\phi_i(I_A^\alpha) &= -2^\alpha(I_A^\alpha)^{1-\alpha} \int_0^1 (\mu p - G(p))^{\alpha-1} (\phi_i(p; G) - p(Y_i - \mu)) dp \\ \phi_i(W^\alpha) &= 2(Y_i - \mu) + \phi_i(I_A^\alpha) \\ \phi_i(I_R^\alpha) &= \frac{1}{\mu} \phi_i(I_A^\alpha) - \frac{I_R^\alpha}{\hat{\mu}} (Y_i - \hat{\mu})\end{aligned}$$

Then the variances can be estimated using,

$$\begin{aligned}\hat{V}(\hat{W}^\alpha) &= \frac{1}{N} \sum_{i=1}^N \hat{\phi}_i(W^\alpha)^2 \\ \hat{V}(\hat{I}_A^\alpha) &= \frac{1}{N} \sum_{i=1}^N \hat{\phi}_i(I_A^\alpha)^2 \\ \hat{V}(\hat{I}_R^\alpha) &= \frac{1}{N} \sum_{i=1}^N \hat{\phi}_i(I_R^\alpha)^2\end{aligned}$$

where the $\hat{\phi}$ are ϕ with unknowns replaced by the relevant estimates. Appendix B indicates how these influence functions can be estimated for a particular sample.

Proposition 2(B): *Given Assumption 1,*

$$\hat{V}(\hat{I}_R^\alpha) \xrightarrow{p} V(\hat{I}_R^\alpha),$$

$$\hat{V}(\hat{I}_A^\alpha) \xrightarrow{p} V(\hat{I}_A^\alpha)$$

$$\hat{V}(\hat{W}^\alpha) \xrightarrow{p} V(\hat{W}^\alpha)$$

3.3.3 Poverty Indices

It is now straightforward to consider the estimation of the poverty indices defined in Section 2. The index of Poverty P^δ based on the S-Gini can be estimated by,

$$\hat{P}^\delta = \hat{F}(z) - \frac{1}{z} \delta(\delta - 1) \int_0^1 (1 - p)^{\delta-2} \hat{G}(p\hat{F}(z)) dp$$

when $\delta > 1$, and,

$$\hat{P}^\delta = \frac{1}{Nz} \sum_{i=1}^N (z - Y_i) \cdot 1(Y_i \leq z)$$

when $\delta = 1$. For the E-Gini based poverty index one can use,

$$\hat{P}^\alpha = \hat{F}(z) - \frac{2}{z}\hat{G}(\hat{F}(z)) + \frac{1}{z}\left[\int_0^1 (p\hat{G}(\hat{F}(z)) - \hat{G}(p\hat{F}(z)))^\alpha dp\right]^{\frac{1}{\alpha}}.$$

The following Propositions provide characterizations of the limiting distribution of these estimators of the Poverty indices.

Proposition 3(A): *Given Assumption 1, a fixed poverty level z , and $1 < \delta < \infty$, then*

$$\hat{P}^\delta \xrightarrow{p} P^\delta \text{ and}$$

$$\begin{aligned} \sqrt{N}(\hat{P}^\delta - P^\delta) &\Rightarrow \mathcal{B}(F(z)) - \frac{\delta(\delta-1)}{z} \int_0^1 (1-p)^{\delta-2} \{Q(pF(z))p\mathcal{B}(F(z)) + \mathcal{G}(pF(z))\} dp \\ &\sim N(0, V(\hat{P}^\delta)) \end{aligned}$$

The influence curves for the S-gini based poverty index is given by,

$$\begin{aligned} \phi_i(P^\delta) &= \hat{\phi}_i(z; F) \\ &\quad - \frac{\delta(\delta-1)}{z} \int_0^1 (1-p)^{\delta-2} (pQ(pF(z))\hat{\phi}_i(z; F) + \phi_i(pF(z); G)) dp \end{aligned}$$

and the variance can be estimated using,

$$\hat{V}(\hat{P}^\delta) = \frac{1}{N} \sum_{i=1}^N \hat{\phi}_i(P^\delta)^2$$

Appendix B details how these influence functions can be estimated for a particular sample.

Proposition 3(B): *Given the conditions of Theorem 1 $\hat{V}(\hat{P}^\delta) \xrightarrow{p} V(\hat{P}^\delta)$.*

In the case where $\delta = 1$ things are much simpler because of the fact that the index in this case is essentially a scaled sample average (of the shortfall of income from the poverty line normalized by the poverty line). Therefore standard results for sample averages apply so that the usual sample variance formula can be used to consistently estimate the variance of the limiting distribution of the statistic.

Proposition 4(A): *Given Assumption 1, a fixed poverty level z , and $\delta = 1$, then*

(i) $\hat{P}^\delta \xrightarrow{p} P^\delta$ and

$$\begin{aligned}\sqrt{N}(\hat{P}^\delta - P^\delta) &\Rightarrow N\left(0, \frac{1}{z^2} \left\{ E \left[(z - Y_i)^2 \cdot 1(Y_i \leq z) \right] - E \left[(z - Y_i) \cdot 1(Y_i \leq z) \right]^2 \right\}\right) \\ &\equiv N(0, V(\hat{P}^\delta))\end{aligned}$$

(ii)

$$\hat{V}(\hat{P}^\delta) = \frac{1}{Nz^2} \sum_{i=1}^N (z - Y_i)^2 \cdot 1(Y_i \leq z) - (\hat{P}^\delta)^2 \xrightarrow{p} V(\hat{P}^\delta)$$

Proposition 4(B): *Given the conditions of Theorem 1 then for $\delta = 1$, $\hat{V}(\hat{P}^\delta) \xrightarrow{p} V(\hat{P}^\delta)$.*

Finally for the Poverty index based on the E-gini index we have the following result.

Proposition 5(A): *Given Assumption 1, a fixed poverty level z , and $1 < \alpha < \infty$, then*

$\hat{P}^\alpha \xrightarrow{p} P^\alpha$ and

$$\begin{aligned}\sqrt{N}(\hat{P}^\alpha - P^\alpha) &\Rightarrow \mathcal{B}(F(z)) - \frac{2}{z}\mathcal{G}(F(z)) - 2\mathcal{B}(F(z)) \\ &\quad + \frac{1}{z}[T_E^z]_{\alpha}^{\frac{1}{\alpha}-1} \int_0^1 (pG(F(z)) - G(pF(z)))^{\alpha-1} \Upsilon(p, F(z)) dp \\ &\sim N(0, V(\hat{P}^\alpha)).\end{aligned}$$

where,

$$T_E^z = \int_0^1 (pG(F(z)) - G(pF(z)))^\alpha dp$$

and,

$$\Upsilon(p, F(z)) = p(z - Q(pF(z))) \mathcal{B}(F(z)) + p\mathcal{G}(F(z)) - \mathcal{G}(pF(z)).$$

For the E-gini based poverty index the influence curve is given by

$$\begin{aligned}\phi_i(P^\alpha) &= -\hat{\phi}_i(z; F) - \frac{2}{z}\phi_i(F(z); G) \\ &\quad + \frac{1}{z}[T_E^z]_{\alpha}^{\frac{1}{\alpha}-1} \int_0^1 (pG(F(z)) - G(pF(z)))^{\alpha-1} \phi_i(p, F(z); \Upsilon) dp\end{aligned}$$

where,

$$\phi_i(p, F(z); \Upsilon) = p(z - Q(pF(z))) \hat{\phi}_i(z; F) + p\phi_i(F(z); G) - \phi_i(pF(z); G)$$

Then the variance can be estimated using,

$$\hat{V}(\hat{P}^\alpha) = \frac{1}{N} \sum_{i=1}^N \hat{\phi}_i(P^\alpha)^2$$

where $\hat{\phi}$ is ϕ with unknowns replaced by the relevant estimates. Appendix B indicates how these influence functions can be estimated for a particular sample.

Proposition 5(B): *Given the conditions of Theorem 1 $\hat{V}(\hat{P}^\alpha) \xrightarrow{p} V(\hat{P}^\alpha)$.*

Note that these results establish the root-N consistency and asymptotic normality of the poverty indices, although the formulae for the variances of the limiting normal distributions in all cases except the S-Gini $\delta = 1$ case, although tractable, are quite long and involved.

4 Computational Issues

4.1 Generalized Gini Indices

In this section we consider the issue of computing the estimates of the inequality indices. Before considering this we first consider the computation of the GLC and LC. Note that for any particular sample size the estimator of the $F(y)$, $\hat{F}(y)$ is a step function. Denote the (ordered) distinct sample values for the Y_i by the notation y_j so that,

$$y_l \leq y_1 < y_2 < \dots < y_{\dot{N}} \leq y_u$$

where \dot{N} is the number of distinct values in the sample and it may be the case that $\dot{N} < N$ if there are ties in the sample. Then given this notation, the function $\hat{F}(y|x)$ is a step function with increments,

$$\hat{\pi}_j = \sum_{i=1}^N 1(Y_i = y_j)$$

occurring at each of the y_j values. Therefore the estimator $\hat{F}(y)$ takes on the value $\sum_{l=1}^j \hat{\pi}_l = \hat{p}_j$ on the interval $[y_j, y_{j+1})$. In terms of this we can then define the quantile

function using similar arguments. In terms of the y_j we can define the quantile estimator as follows,

$$\hat{Q}(p|x) = \sum_{j=1}^{\hat{N}} 1(\hat{p}_{j-1} < p \leq \hat{p}_j) y_j$$

where we define $\hat{F}(y_0) = \hat{p}_0 = 0$. Thus over the range of p where $p \in (\hat{p}_{j-1}, \hat{p}_j]$ (which has length $\hat{\pi}_j$) the quantile function takes on the value y_j . Thus for $p \leq \hat{\pi}_1$ the quantile function is y_1 for $\hat{\pi}_1 < p \leq \hat{\pi}_1 + \hat{\pi}_2$ the quantile function is y_2 and so on.

Given the above representation it is then a straightforward exercise to calculate the integrals that define the GLC and LC. In particular for p such that $p \in [\hat{p}_{j-1}, \hat{p}_j)$ we have that,

$$\begin{aligned} \hat{G}(p) &= (p - \hat{p}_{j-1}) y_j + \sum_{l=1}^{j-1} \hat{\pi}_l y_l \\ &= p y_j + a_j \end{aligned}$$

where,

$$a_j = \sum_{l=1}^{j-1} \hat{\pi}_l (y_l - y_j)$$

and define $a_1 = 0$ so that when $p < \hat{p}_1$ the function $\hat{G}(p) = p y_1$. Thus, $\hat{G}(p)$ is piecewise linear and continuous. For simplicity we assume that $y_{\hat{N}} = y_u$ so that by construction for $p = 1$ we have that,

$$\hat{G}(1) = \sum_{j=1}^{\hat{N}} \hat{\pi}_j y_j = \hat{\mu} = \bar{Y}$$

which is the sample mean of the Y_i . Given the estimate of GLC it is then straightforward to see that,

$$\hat{L}(p) = \frac{\hat{G}(p)}{\hat{G}(1)}$$

since this is how we have defined our LC estimator above. Note that $\hat{G}(0) = \hat{L}(0) = 0$ and $\hat{L}(1) = 1$. Note that the LC estimate is also piecewise linear and continuous on $[0, 1]$.

In computing the S-gini indices it is most convenient to use the representation in terms of quantiles. Use is made of the fact that,

$$\int_0^1 = \sum_{j=1}^{\hat{N}} \int_{\hat{p}_{j-1}}^{\hat{p}_j}$$

where the relevant functions have been defined over the intervals $(\hat{p}_{j-1}, \hat{p}_j]$. Based on the representation of the empirical quantiles given above the estimates for the S-gini indices require the calculation of \hat{W}^δ which is done as follows,

$$\begin{aligned}\hat{W}^\delta &= \int_0^1 \delta(1-p)^{\delta-1} \hat{Q}(p) dp \\ &= \sum_{j=1}^{\hat{N}} y_j \int_{\hat{p}_{j-1}}^{\hat{p}_j} \delta(1-p)^{\delta-1} dp \\ &= \sum_{j=1}^{\hat{N}} y_j \left\{ (1-\hat{p}_{j-1})^\delta - (1-\hat{p}_j)^\delta \right\}\end{aligned}$$

Therefore we can compute the S-gini based indices as follows,

$$\begin{aligned}\hat{I}_R^\delta &= 1 - \frac{\hat{W}^\delta}{\hat{\mu}}, \\ \hat{I}_A^\delta &= \hat{\mu} - \hat{W}^\delta.\end{aligned}$$

Next, for the E-gini indices it is necessary to calculate the following quantity,

$$\hat{T}_E = \int_0^1 (\hat{\mu} \cdot p - \hat{G}(p))^\alpha dp$$

To do this we use the fact that, over the interval $[\hat{p}_{j-1}, \hat{p}_j)$

$$\begin{aligned}\hat{\mu} \cdot p - \hat{G}(p) &= \hat{\mu} \cdot p - \left\{ (p - \hat{p}_{j-1}) y_j + \sum_{l=1}^{j-1} \hat{\pi}_l y_l \right\} \\ &= (\hat{\mu} - y_j) p + \sum_{l=1}^{j-1} \hat{\pi}_l (y_j - y_l) \\ &= b_j p - a_j\end{aligned}$$

where $b_j = (\hat{\mu} - y_j)$, which is a linear function of p . Therefore,

$$\hat{T}_E = \frac{1}{\alpha + 1} \sum_{j=1}^{\hat{N}} \left(1(\hat{\mu} \neq y_j) \frac{1}{b_j} \{ (b_j \hat{p}_j - a_j)^{\alpha+1} - (b_j \hat{p}_{j-1} - a_j)^{\alpha+1} \} - 1(\hat{\mu} = y_j) a_j (\hat{p}_j - \hat{p}_{j-1}) \right)$$

where the second term deals with any cases where $b_j = \hat{\mu} - y_j = 0$ in which case,

$$\hat{\mu} \cdot p - \hat{G}(p) = -a_j$$

so that the general integral does not apply.⁷ With these calculations we can then compute the indices by,

$$\hat{I}_A^\alpha = 2[\hat{T}_E]^\frac{1}{\alpha}$$

$$\hat{I}_R^\alpha = \frac{\hat{I}_A^\alpha}{\hat{\mu}}$$

and,

$$\hat{W}^\alpha = 2\hat{\mu} - \hat{I}_A^\alpha$$

4.2 Poverty Indices

For the poverty indices the computation is very similar to the computation of the Gini inequality indices. The computation of \hat{P}^δ in the case where $\delta > 1$ and \hat{P}^α will be considered since the case of \hat{P}^δ with $\delta = 1$ is a simple average. For the S-gini based index we use the second representation for the S-gini based conditional Welfare index given in (8) and estimate,

$$\begin{aligned} \hat{P}^\delta &= \hat{F}(z) - \frac{\hat{F}(z)}{z} \int_0^1 \delta(1-p)^{\delta-1} \hat{Q}(p\hat{F}(z)) dp \\ &= \hat{F}(z) - \frac{1}{z\hat{F}(z)^{\delta-1}} \int_0^{\hat{F}(z)} \delta(\hat{F}(z) - p)^{\delta-1} \hat{Q}(p) dp \end{aligned}$$

where the second line follows after a change of variables. In order to estimate this quantity we define $N(z)$ to be the value for which $\hat{F}(z) = p_{N(z)}$ ⁸. Then similar to the estimation of the S-Gini inequality indices we can compute the poverty index by,

$$\hat{P}^\delta = \hat{F}(z) - \frac{1}{z\hat{F}(z)^{\delta-1}} \sum_{j=1}^{N(z)} y_j \left\{ (\hat{F}(z) - \hat{p}_{j-1})^\delta - (\hat{F}(z) - \hat{p}_j)^\delta \right\}$$

Similar arguments for the E-Gini based index suggest estimating the index in the form,

$$\hat{P}^\alpha = \hat{F}(z) - \frac{2}{z} \hat{G}(\hat{F}(z)) + \frac{1}{z} \left(\frac{1}{\hat{F}(z)} \int_0^{\hat{F}(z)} (p\hat{\mu}(z) - \hat{G}(p))^\alpha dp \right)^\frac{1}{\alpha}.$$

⁷Note the the sums are separated only for display purposes and could be combined.

⁸In other words $N(z)$ is the number of observations that are less than or equal to z . Also note that $\hat{F}(z) = N(z)/N$.

where,

$$\hat{\mu}(z) = \frac{\hat{G}(\hat{F}(z))}{\hat{F}(z)}$$

and,

$$\begin{aligned} \hat{T}_E^z &= \int_0^1 (p\hat{G}(\hat{F}(z)) - \hat{G}(p\hat{F}(z)))^\alpha dp \\ &= \frac{1}{\hat{F}(z)} \int_0^{\hat{F}(z)} (p\hat{\mu}(z) - \hat{G}(p))^\alpha dp \\ &= \frac{1}{\alpha + 1} \sum_{j=1}^{N(z)} \left(1(\hat{\mu}(z) \neq y_j) \frac{1}{b_j} \{ (b_j \hat{p}_j - a_j)^{\alpha+1} - (b_j \hat{p}_{j-1} - a_j)^{\alpha+1} \} - 1(\hat{\mu}(z) = y_j) a_j (\hat{p}_j - \hat{p}_{j-1}) \right) \end{aligned}$$

using a change of variable and arguments similar to those used to derive \hat{T}_E . Then we have that,

$$\hat{P}^\alpha = \hat{F}(z) - \frac{2\hat{G}(\hat{F}(z))}{z} + \frac{1}{z} \left(\hat{T}_E^z \right)^\frac{1}{\alpha}$$

where we recall that $\hat{F}(z) = \hat{p}_{N(z)}$.

5 Empirical Application

5.1 Trends in Income Inequality and Poverty in the US

The theoretical results derived in Section 3, and the computational algorithms presented in the previous section and Appendix B, enable formal statistical inference to be conducted using consistent estimates of the generalized Gini indices. It is straightforward to use the estimate of an inequality, poverty or welfare index, \hat{I} , and the sampling variance, $Var(\hat{I})$, to construct appropriate test statistics. The test of the null hypothesis that the level of index is equal to a predetermined value, I_0 , is based on the statistic

$$z_1 = \frac{I_0 - \hat{I}}{\sqrt{\frac{Var(\hat{I})}{n}}} = \frac{I_0 - \hat{I}}{se(\hat{I})}$$

which is asymptotically distributed standard normal. When comparing independent samples (with sample sizes N_1 and N_2 respectively) from two income distributions and

testing the null hypothesis that they are ‘equally unequal’ (or contain the same level of poverty or social welfare) the appropriate test statistic is

$$z_2 = \frac{\hat{I}_2 - \hat{I}_1}{\sqrt{\frac{Var(\hat{I}_2)}{N_2} + \frac{Var(\hat{I}_1)}{N_1}}} = \frac{\hat{I}_2 - \hat{I}_1}{\sqrt{se(\hat{I}_2)^2 + se(\hat{I}_1)^2}}$$

where \hat{I}_1 and \hat{I}_2 are the estimates of the indices for the two samples. This statistic is asymptotically distributed as a standard normal random variable under the null hypothesis.

The inferential methods are illustrated by considering the changes in distribution of income in the United States over the past two decades. The data are from the Current Population Survey (CPS) March Demographic files for 1978, 1988 and 1998. Annual gross family income⁹ from the set of family records is examined. Nominal values are inflated to 1998 dollars using the GDP implicit price deflator. To take account of economies of scale in consumption and differences in family needs, gross family income is divided by the adult equivalent scale implicit in the Bureau of the Census Poverty Thresholds. The Poverty Thresholds for 1998 are also adopted as the poverty lines. The CPS data include a record weight to allow for the non-random sample frame. The family weights are used in the analysis after being multiplied by the number of family members. This procedure assigns every member in the family the same level of equivalent income, and provides a sample representative of the population of individuals.¹⁰

Basic descriptive statistics for the samples are provided in Table 1. There was considerable income growth over the past two decades: average real individual equivalent income increased 17 percent from 1978 to 1988 and by more than 18 percent between 1988 and 1998. In this context of strong income growth, we now examine the trends in inequality and poverty.

Table 1.

⁹This is total family income pre-tax but post cash transfers.

¹⁰Several sample exclusions were imposed. First, records for families which reported non-zero farm income were dropped due to the difficulties of imputing non-market income. Secondly, families with elderly members were dropped from the analysis. Therefore, the samples examined are representative of the nonfarm, nonelderly population.

Descriptive Statistics for the Distribution of Equivalent Income (US\$1998)

Sample	Mean	Median	Obs.
1978	21548	18729	51222
1988	25778	20962	56939
1998	31312	23157	54285

The point estimates (and asymptotic standard errors) of the generalized Gini indices of inequality indices are presented in Table 2.¹¹ Several members of the S-Gini and E-Gini family of indices are presented in order to capture a broad range of normative positions. The first panel contains the relative inequality indices. All the relative Gini indices imply a significant increase in inequality between 1978 and 1988. The generalized Gini relative inequality indices increased by approximately 10 percent, which is consistent with the findings of the substantial empirical literature on changes in the US distribution of earnings and income over the 1980's (Levy and Murnane, 1992; Gottschalk and Smeeding, 1997). The trend of rising inequality continued throughout the 1990's, and in some respects accelerated. The least inequality averse S-Gini relative indices ($\delta=1.25$) increased by 20 percent, the Gini coefficient ($\delta=2.00$) increased by 11 percent and the more inequality averse S-Ginis rose by a smaller fraction over the period 1988-98. The E-Gini relative indices uniformly increased by approximately 12 percent. The S-Ginis with $\delta \in (1, 2)$ are most sensitive to transfers at the top of the income distribution, and together the pattern of change in the S-Gini and E-Gini indices reveal that it was the income shares of those at the very top of the income distribution that grew the most over the 1990's.

Another dimension of the change in income inequality is revealed by the generalized Gini absolute indices. Recall that the absolute inequality indices are sensitive to income differentials unlike the relative indices which are sensitive to differences in income shares. As shown in the second panel of Table 2, all members of the S-Gini and E-Gini family of absolute inequality indices show a significant rise in income inequality between 1978 and 1988 of over 30 percent. Although both sets of absolute indices reveal that the trend

¹¹A Fortran program that calculates the generalised Gini inequality, poverty and welfare indices, and their corresponding standard errors, is available from the authors upon request.

toward greater inequality continued, the pattern was somewhat different over the 1990's. The least inequality averse absolute S-Ginis show the greatest rise in inequality between 1988 and 1998, exceeding the increase in inequality over the 1980's. In contrast, the more inequality averse S-Ginis and all the E-Ginis report an increase in inequality over the 1990's that was comparable to the rise in inequality over the 1980's. Overall, it is apparent that the growth in income inequality that occurred during the 1980's continued into the 1990's; however unlike the 1980's, the 1990's has seen the growth in income shares and income differentials concentrated among individuals at the very top of the distribution.

The final panel of Table 2 presents estimates of the welfare indices underlying the generalized Gini inequality indices. The period 1978-98 witnessed strong growth in average income combined with significant increases in inequality. Given these countervailing trends, it is an empirical question whether social welfare, which incorporates a concern for both 'efficiency' and 'equity', rose over this period. The estimates of the generalized Gini welfare indices, and associated standard errors, show that aggregate social welfare progressively improved over this period. Even the E-Gini welfare indices most sensitive to changes at the very bottom of the distribution show an unambiguous rise in social welfare between 1978-88 and 1988-98.

Table 2.

Estimates of Gini Inequality and Welfare Indices (and standard errors).

Year	S-Gini Relative Indices				E-Gini Relative Indices			
	δ				α			
	<i>1.25</i>	<i>2.00</i>	<i>2.50</i>	<i>3.50</i>	<i>1.00</i>	<i>1.75</i>	<i>3.00</i>	<i>10.0</i>
1978	0.1384 (0.0006)	0.3584 (0.0011)	0.4395 (0.0011)	0.5414 (0.0012)	0.3584 (0.0011)	0.3807 (0.0011)	0.4043 (0.0012)	0.4530 (0.0013)
1988	0.1554 (0.0007)	0.3982 (0.0011)	0.4858 (0.0011)	0.5940 (0.0012)	0.3982 (0.0011)	0.4233 (0.0011)	0.4499 (0.0012)	0.5045 (0.0014)
1998	0.1872 (0.0011)	0.4415 (0.0016)	0.5251 (0.0016)	0.6251 (0.0015)	0.4415 (0.0016)	0.4686 (0.0017)	0.4971 (0.0019)	0.5558 (0.0021)
	S-Gini Absolute Indices				E-Gini Absolute Indices			
	δ				α			
	<i>1.25</i>	<i>2.00</i>	<i>2.50</i>	<i>3.50</i>	<i>1.00</i>	<i>1.75</i>	<i>3.00</i>	<i>10.0</i>
1978	2982.79 (18.51)	7723.39 (35.55)	9471.53 (39.88)	11667.29 (44.68)	7723.39 (35.55)	8202.75 (37.36)	8710.99 (39.35)	9761.25 (43.68)
1988	4047.45 (24.31)	10370.94 (47.25)	12652.12 (53.15)	15470.46 (59.48)	10370.94 (47.25)	11025.92 (49.95)	11717.47 (52.85)	13140.05 (59.02)
1998	6033.48 (56.74)	14228.56 (101.79)	16921.02 (109.71)	20142.59 (116.82)	14228.56 (101.79)	15101.02 (108.15)	16019.31 (114.77)	17911.14 (127.42)
	S-Gini Welfare Indices				E-Gini Welfare Indices			
	δ				α			
	<i>1.25</i>	<i>2.00</i>	<i>2.50</i>	<i>3.50</i>	<i>1.00</i>	<i>1.75</i>	<i>3.00</i>	<i>10.0</i>
1978	18565 (48.76)	13825 (39.19)	12077 (36.87)	9881 (34.32)	35373 (95.19)	34894 (94.21)	34386 (93.21)	33335 (91.20)
1988	21999 (62.60)	15675 (48.05)	13394 (44.24)	10575 (39.86)	41721 (121.16)	41066 (119.57)	40374 (117.95)	38952 (114.73)
1998	26191 (89.92)	17996 (58.05)	15303 (52.09)	12082 (46.21)	50220 (180.79)	49348 (175.90)	48430 (170.94)	46538 (161.85)

Estimates of the generalized Gini poverty indices are presented in Table 3. The proportion of the population below the poverty line in 1978 was 16.86 percent, and the head-count ratio did not change significantly between 1978 and 1988. Although the head-count ratio is popular with policy analysts and researchers (eg. Hanratty and Blank, 1992; Slesnick, 1993), it is a very restrictive summary measure of poverty since it is not sensitive to either the depth of poverty or the distribution of income among the poor population (Sen, 1976). The S-Gini poverty index for $\delta = 1$ is equal to the headcount ratio multiplied by the average poverty gap (normalized by the poverty line). This poverty index increased by 0.0093 (z -statistic=3.434) from 1978-88 revealing that the average income shortfall of the poor increased over the 1980s. The S-Gini poverty indices for $\delta > 1$ and the E-Gini poverty indices are a composite function of the headcount ratio, average poverty gap and the level of income inequality among the poor. All these indices show a statistically significant increase in the level of poverty over the 1980's.

The period from 1988 to 1998 witnessed a significant and substantial decline in the poverty head-count by 0.0175 (z -statistic=3.865) and by 1998 the fraction of the population in poverty was significantly less than that in 1978. In addition, the S-Gini poverty index for $\delta = 1$, plus the distributionally sensitive S-Gini poverty indices show a significant decline in poverty over the 1990s. Although the point estimates of the E-Gini poverty indices indicate a fall in poverty between 1988 and 1998, these changes are not significant. Over the entire 1978-98 period, the S-Gini poverty indices show a statistically significant improvement in poverty. However, despite the decline in the headcount ratio and average poverty gap over this period, the E-Gini poverty indices indicate no significant change in poverty over the past two decades.

Table 3

Estimates of the Gini Poverty Indices (and standard errors).

Year	S-Gini Poverty Indices				E-Gini Poverty Indices			
	δ				α			
	<i>1.00</i>	<i>1.50</i>	<i>2.00</i>	<i>3.00</i>	<i>1.00</i>	<i>1.75</i>	<i>3.00</i>	<i>10.0</i>
1978	0.0610 (0.0017)	0.0744 (0.0008)	0.0839 (0.0008)	0.0968 (0.0010)	0.1290 (0.0045)	0.1337 (0.0047)	0.1386 (0.0050)	0.1485 (0.0058)
1988	0.0703 (0.0018)	0.0852 (0.0008)	0.0954 (0.0008)	0.1089 (0.0009)	0.1458 (0.0054)	0.1512 (0.0056)	0.1569 (0.0060)	0.1683 (0.0068)
1998	0.0552 (0.0016)	0.0673 (0.0007)	0.0759 (0.0008)	0.0874 (0.0009)	0.1322 (0.0079)	0.1378 (0.0082)	0.1437 (0.0087)	0.1557 (0.0098)

Notes: The poverty head-count ratios (and standard error) were 0.1686 (0.0017), 0.1663 (0.0016) and 0.1347 (0.0015) for 1978, 1988 and 1998 respectively.

6 Conclusion

The generalized Gini indices of inequality, poverty and welfare have many desirable normative properties. In this paper we have derived the statistical properties of consistent estimators of these indices. In contrast to previous contributions in this area, the methods presented in this paper do not require the grouping of sample observations into income quantiles and hence the desirable normative properties of the Gini indices need not be compromised when performing formal hypothesis testing.¹² The sampling variances of the Gini indices are derived by defining the indices as statistical functionals of Lorenz or Generalized Lorenz curves, and then considering the influence function for these estimators. As a by-product of this approach we present an alternative derivation of the variance-covariance matrix of the empirical LC and GLC estimators. The theoretical results presented in this paper unify and extend previous work on the sampling properties of specific members of the generalized Gini class of indices.

The methods for consistent statistical inference with the Gini indices are illustrated by considering the changes in the distribution of income in the United States over the 1980's

¹²The results concerning the asymptotic distributions of the indices provide some justification for the use of bootstrap methods for inference. It is not clear, however, whether there is any computational advantages of such an approach and, moreover, there is no guarantee that improvements in inferences will be obtained unless computationally costly pivotal statistics are used.

and 1990's. It was found that over the period income inequality continually increased and that while the rise in inequality over the 1980's occurred across the income distribution, the increase in inequality during the 1990's was associated with more concentrated income growth among individuals at the top of the distribution. Over the same period, average real income growth was sufficiently strong to outweigh the rise in inequality and aggregate social welfare increased. The Gini poverty indices revealed an increase in poverty over the 1980's followed by a subsequent decline in the 1990s. The level of poverty in 1998 was found to be no worse than that experienced in 1978, and some indices indicate a significant improvement in poverty over the two decades.

Appendix A: Proofs of Results

Proof of Lemma 3: (i) Note that ,

$$\hat{G}(p) - G(p) = \int_0^p (\hat{Q}(t) - Q(t))dt.$$

Consequently we have that,

$$\begin{aligned} \sup_p |\hat{G}(p) - G(p)| &= \sup_p \left| \int_0^p (\hat{Q}(t) - Q(t))dt \right| \\ &\leq \sup_p \int_0^p \sup_{t \in [0,p]} |\hat{Q}(t) - Q(t)| dt \\ &\leq \sup_{t \in [0,1]} |\hat{Q}(t) - Q(t)| \sup_p \int_0^p dt \\ &= \sup_{t \in [0,1]} |\hat{Q}(t) - Q(t)| \\ &= o_p(1) \end{aligned}$$

using the result in Lemma 2. For the next result note that $G = T(Q)$ is a linear functional of Q such that,

$$T(Q)(p) = \int_0^p Q(t)dt$$

and is therefore Hadamard differentiable at Q tangentially to $C[0,1]$ with Hadamard derivative T' with

$$T'(\tilde{Q})(p) = \int_0^p \tilde{Q}(t)dt$$

for $\tilde{Q} \in C[0,1]$. Consequently,

$$\sqrt{N}(\hat{G} - G) \Rightarrow \mathcal{G}$$

in $l^\infty(0,1)$ where,

$$\mathcal{G}(p) = - \int_0^p \frac{\mathcal{B}(t)}{f(Q(t))} dt.$$

(ii) For the LC we have that since $\hat{L}(p) = \hat{G}(p)/\hat{\mu}$ then,

$$\begin{aligned} \sup_p |\hat{L}(p) - L(p)| &\leq \frac{1}{\hat{\mu}} \sup_p |\hat{G}(p) - G(p)| + \frac{L(p)}{\hat{\mu}} |\mu - \hat{\mu}| \\ &\xrightarrow{p} 0 \end{aligned}$$

by result (i) and $\hat{\mu} \xrightarrow{p} \mu$ with $0 < \mu < \infty$ which follows from Assumption 1. Similarly,

$$\begin{aligned}\sqrt{N}(\hat{L} - L) &= \frac{1}{\hat{\mu}}\sqrt{N}(\hat{G} - G) - \frac{L}{\hat{\mu}}\sqrt{N}(\hat{\mu} - \mu) \\ &\Rightarrow \frac{\mathcal{G}}{\mu} - \frac{L}{\mu}\mathcal{G}(1)\end{aligned}$$

using the Slutsky Theorem (ST) the Continuous Mapping Theorem (CMT) and the fact that,

$$\sqrt{N}(\hat{\mu} - \mu) \Rightarrow \mathcal{G}(1)$$

since $\hat{\mu} = \hat{G}(1)$. **Q.E.D.**

Proof of Lemma 4: (i) Given the definition of $\phi_i(p; Q)$ and the fact that $G(p) = \int_0^p Q(t)dt$ then it follows using a change of variable from t to y that,

$$\begin{aligned}\phi_i(p; G) &= \int_0^p \phi_i(t; Q)dt \\ &= \int_0^p \left(1(Y_i < Q(t)) \frac{t-1}{f(Q(t))} + 1(Y_i > Q(t)) \frac{t}{f(Q(t))} \right) dt \\ &= - \int_{y_i}^{y_p} 1(Y_i < y) (1 - F(y)) dy + \int_{y_i}^{y_p} 1(Y_i > y) F(y) dy \\ &= -1(Y_i < y_p) \left(\int_{Y_i}^{y_p} (1 - F(y)) dy - \int_{y_i}^{Y_i} F(y) dy \right) + 1(Y_i > y_p) \int_{y_i}^{y_p} F(y) dy \\ &= (py_p - G(p)) - 1(Y_i < y_p)(y_p - Y_i)\end{aligned}$$

after integration by parts is used to show that,

$$\begin{aligned}\int_{y_i}^{y_p} F(y) dy &= yF(y)|_{y_i}^{y_p} - \int_{y_i}^{y_p} yf(y) dy \\ &= y_p F(y_p) - G(p)\end{aligned}$$

Then take the definition of $G(p)$ and the facts that $F(y_p) = p$ and $y_p = Q(p)$ and the result follows.

(ii) For the LC we use the fact that,

$$\hat{L}(p) = \frac{\hat{G}(p)}{\hat{G}(1)}$$

so that by the product rule for differentiation,

$$\begin{aligned}\phi_i(p; L) &= \frac{1}{G(1)}\phi_i(p; G) - \frac{G(p)}{G(1)^2}\phi_i(1; G) \\ &= \frac{1}{\mu}\phi_i(p; G) - \frac{L(p)}{\mu}(Y_i - \mu)\end{aligned}$$

using the fact that $G(1) = \mu$. **Q.E.D.**

Proof of Lemma 5: (i) By the fact shown in Fernholz (1983) that

$$\frac{1}{N} \sum_{i=1}^N \phi_i(p; G)^2 \xrightarrow{p} E(\phi_i(p; G)^2) = p \left(\sigma_p^2 + (1-p) \left(y_p - \frac{G(p)}{p} \right)^2 \right)$$

all we must do is show that the sufficient conditions in Lemma A.1 hold. To do this note that,

$$\begin{aligned} |\hat{\phi}_i(p; G) - \phi_i(p; G)| &\leq p|\hat{Q}(p) - Q(p)| + |\hat{G}(p) - G(p)| + 1(Y_i < Q(p))|\hat{Q}(p) - Q(p)| \\ &\quad + Q(p)|1(Y_i < \hat{Q}(p)) - 1(Y_i < Q(p))| + Y_i|1(Y_i < \hat{Q}(p)) - 1(Y_i < Q(p))| \end{aligned}$$

The facts that $0 < p < 1$ and $1(Y_i < Q(p)) \leq 1$ and the results shown in Lemma 2 and 3,

$$\begin{aligned} \hat{Q}(p) - Q(p) &= o_p(1) \\ \hat{G}(p) - G(p) &= o_p(1) \end{aligned}$$

imply that conditions (i) and (ii) are satisfied. To show (iii) and (iv) we note that $Q(p) < \infty$ and $Y_i < \infty$ by Assumption 1 and that because,

$$\begin{aligned} |1(Y_i < \hat{Q}(p)) - 1(Y_i < Q(p))|^4 &= |1(Y_i < \hat{Q}(p)) - 1(Y_i < Q(p))|^2 \\ &= |1(Y_i < \hat{Q}(p)) - 1(Y_i < Q(p))| \end{aligned}$$

it suffices to show that,

$$\frac{1}{N} \sum_{i=1}^N |1(Y_i < \hat{Q}(p)) - 1(Y_i < Q(p))| = o_p(1).$$

To show this we note that by Lemma 2 for any $\eta > 0$ there exists a finite $\bar{N}(\eta)$ such that $|\hat{Q}(p) - Q(p)| < \eta$ for all $N \geq \bar{N}(\eta)$ with probability one. Then note that for $N \geq \bar{N}(\eta)$ the event $|1(Y_i < \hat{Q}(p)) - 1(Y_i < Q(p))| = 1$ implies that $Y_i \in (Q(p) - \eta, Q(p) + \eta)$ so that,

$$\frac{1}{N} \sum_{i=1}^N |1(Y_i < \hat{Q}(p)) - 1(Y_i < Q(p))| \leq \frac{1}{N} \sum_{i=1}^N 1(Y_i \in (Q(p) - \eta, Q(p) + \eta)) \quad (11)$$

Let $Z_i(\eta) = 1(Y_i \in (Q(p) - \eta, Q(p) + \eta))$ and note that

$$\begin{aligned} E(Z_i(\eta)) &= P(Y_i \in (Q(p) - \eta, Q(p) + \eta)) \\ &\leq 2\eta \sup_y f(y) \end{aligned}$$

Thus by the law of large numbers we have that for any fixed $\epsilon > 0$,

$$P\left(\left|\frac{1}{N} \sum_{i=1}^N Z_i(\eta)\right| > 2\eta \sup_y f(y) + \frac{\epsilon}{2}\right) \rightarrow 0.$$

Now take any arbitrarily small fixed $\epsilon > 0$ and let $\eta = \epsilon \left(4 \sup_y f(y)\right)^{-1} > 0$ apply the arguments above to get,

$$\begin{aligned} P\left(\left|\frac{1}{N} \sum_{i=1}^N Z_i(\eta)\right| > 2\eta \sup_y f(y) + \frac{\epsilon}{2}\right) &= P\left(\left|\frac{1}{N} \sum_{i=1}^N Z_i(\eta)\right| > \epsilon\right) \\ &\rightarrow 0 \end{aligned}$$

so that,

$$P\left(\frac{1}{N} \sum_{i=1}^N |1(Y_i < \hat{Q}(p)) - 1(Y_i < Q(p))| > \epsilon\right) \rightarrow 0$$

because of (11). The result follows since ϵ was chosen arbitrarily.

For the LC similar arguments apply with,

$$\phi_i(p; L) = \frac{1}{\hat{\mu}} \hat{\phi}_i(p; G) - \frac{\hat{L}(p)}{\hat{\mu}} (Y_i - \hat{\mu})$$

where in this case we have,

$$\begin{aligned} |\hat{\phi}_i(p; L) - \phi_i(p; L)| &\leq \frac{1}{\hat{\mu}} |\hat{\phi}_i(p; G) - \phi_i(p; G)| + \frac{\phi_i(p; G)}{\hat{\mu}\mu} |\mu - \hat{\mu}| + \frac{\hat{L}(p)}{\hat{\mu}} |\hat{\mu} - \mu| \\ &\quad + \left(\frac{\hat{L}(p)}{\hat{\mu}} - \frac{L(p)}{\mu}\right) (Y_i - \mu) \end{aligned}$$

and the result follows immediately upon application of Lemma A.1. **Q.E.D.**

Lemma A1: For an influence curve ϕ_i with estimate $\hat{\phi}_i$ with

$$|\hat{\phi}_i - \phi_i| \leq \Delta_1 \hat{A}_1 + \Delta_{2i} \hat{A}_2 + \Delta_3 \hat{A}_{3i} + \Delta_{4i} \hat{A}_{4i}$$

then sufficient conditions for,

$$\frac{1}{N} \sum_{i=1}^N \hat{\phi}_i^2 - \frac{1}{N} \sum_{i=1}^N \phi_i^2 = o_p(1)$$

are that the following hold,

- (i) $\hat{A}_1 \xrightarrow{p} 0$ and $|\Delta_1|$ is stochastically bounded;
- (ii) $\hat{A}_2 \xrightarrow{p} 0$ and $\left| \frac{1}{N} \sum_{i=1}^N \Delta_{2i}^2 \right|$ is stochastically bounded;
- (iii) $\frac{1}{N} \sum_{i=1}^N \hat{A}_{3i}^2 \xrightarrow{p} 0$ and $|\Delta_3|$ is stochastically bounded;
- (iv) $\frac{1}{N} \sum_{i=1}^N \hat{A}_{4i}^2 \xrightarrow{p} 0$ and $\frac{1}{N} \sum_{i=1}^N \Delta_{4i}^4$ is stochastically bounded;
- (v) $\frac{1}{N} \sum_{i=1}^N \phi_i^2$ is stochastically bounded.

Proof: This follows from the fact that,

$$\begin{aligned} \left| \frac{1}{N} \sum_{i=1}^N \hat{\phi}_i^2 - \frac{1}{N} \sum_{i=1}^N \phi_i^2 \right| &= \left| \frac{1}{N} \sum_{i=1}^N (\hat{\phi}_i - \phi_i)^2 + \frac{2}{N} \sum_{i=1}^N \phi_i (\hat{\phi}_i - \phi_i) \right| \\ &\leq \frac{1}{N} \sum_{i=1}^N (\hat{\phi}_i - \phi_i)^2 \\ &\quad + 2 \left(\frac{1}{N} \sum_{i=1}^N \phi_i^2 \right)^{1/2} \left(\frac{1}{N} \sum_{i=1}^N (\hat{\phi}_i - \phi_i)^2 \right)^{1/2} \end{aligned}$$

and the conditions given with repeated use of the Cauchy Schwarz inequality. **Q.E.D.**

Proof of Proposition 1(A): Each of these results follow from Lemma 3 because of the linearity of the S-gini indices in the LC and GLC. In particular for (i),

$$\begin{aligned} |\hat{I}_R^\delta - I_R^\delta| &\leq \delta \sup_p |\hat{L}(p) - L(p)| \int_0^1 (\delta - 1)(1 - p)^{\delta-2} dp \\ &= \delta \sup_p |\hat{L}(p) - L(p)| = o_p(1) \end{aligned}$$

using Lemma 3(ii). Also,

$$\begin{aligned} \sqrt{N}(\hat{I}_R^\delta - I_R^\delta) &= -\delta(\delta - 1) \int_0^1 (1 - p)^{\delta-2} \sqrt{N}(\hat{L}(p) - L(p)) dp \\ &\Rightarrow -\delta(\delta - 1) \int_0^1 (1 - p)^{\delta-2} \mathcal{L}(p) dp \sim N(0, V(\hat{I}_R^\delta)) \end{aligned}$$

by the CMT and using the result in Lemma 3(ii). Exactly analogous arguments yield the results in (ii) and (iii) using the result contained in Lemma 3(ii). **Q.E.D.**

Proof of Proposition 1(B): We verify the conditions of Lemma A.1. Consider first

$$\hat{\phi}_i(W^\delta) = \delta(\delta - 1) \int_0^1 (1-p)^{\delta-2} \hat{\phi}_i(p; G) dp$$

and write,

$$\hat{\phi}_i(W^\delta) - \phi_i(W^\delta) = \sum_{j=1}^4 T_j$$

where,

$$\begin{aligned} T_1 &= \delta(\delta - 1) \int_0^1 (1-p)^{\delta-2} p(\hat{Q}(p) - Q(p)) dp \\ T_2 &= -\delta(\delta - 1) \int_0^1 (1-p)^{\delta-2} (\hat{G}(p) - G(p)) dp \\ T_3 &= -\delta(\delta - 1) \int_0^1 (1-p)^{\delta-2} (1(Y_i < \hat{Q}(p))\hat{Q}(p) - 1(Y_i < Q(p))Q(p)) dp \\ T_4 &= Y_i \delta(\delta - 1) \int_0^1 (1-p)^{\delta-2} (1(Y_i < \hat{Q}(p)) - 1(Y_i < Q(p))) dp \end{aligned}$$

For the first term we have,

$$\begin{aligned} |T_1| &\leq \sup_p |\hat{Q}(p) - Q(p)| \delta(\delta - 1) \int_0^1 (1-p)^{\delta-2} p dp \\ &= \sup_p |\hat{Q}(p) - Q(p)| \\ &= o_p(1) \end{aligned}$$

using Lemma 2(i). Similarly using Lemma 3(i),

$$\begin{aligned} |T_2| &\leq \sup_p |\hat{G}(p) - G(p)| \delta(\delta - 1) \int_0^1 (1-p)^{\delta-2} p dp \\ &= o_p(1) \end{aligned}$$

Next,

$$\begin{aligned} |T_3| &= \delta(\delta - 1) \left| \int_{\hat{F}(Y_i)}^1 (1-p)^{\delta-2} \hat{Q}(p) dp - \int_{F(Y_i)}^1 (1-p)^{\delta-2} Q(p) dp \right| \\ &\leq \delta(\delta - 1) \left| \int_{\hat{F}(Y_i)}^1 (1-p)^{\delta-2} \hat{Q}(p) dp - \int_{\hat{F}(Y_i)}^1 (1-p)^{\delta-2} Q(p) dp \right| \end{aligned}$$

$$\begin{aligned}
& +\delta(\delta-1)\left|\int_{\hat{F}(Y_i)}^1(1-p)^{\delta-2}Q(p)dp-\int_{F(Y_i)}^1(1-p)^{\delta-2}Q(p)dp\right| \\
\leq & \sup_p|\hat{Q}(p)-Q(p)|\delta(\delta-1)\int_0^1(1-p)^{\delta-2}dp \\
& +\left|\delta(\delta-1)\int_{\hat{F}(Y_i)}^{F(Y_i)}(1-p)^{\delta-2}Q(p)dp\right| \\
\leq & \sup_p|\hat{Q}(p)-Q(p)|\delta+\sup_p|Q(p)|\sup_y|(1-\hat{F}(y))^{\delta-1}-(1-F(y))^{\delta-1}| \\
= & o_p(1)
\end{aligned}$$

where the last line follows from Assumption 1, and the results in Lemma 2(i) and Lemma 1(i) using the fact that the function $g(x) = x^{\delta-1}$ is uniformly continuous on $[0, 1]$ for $\delta > 1$.

Finally,

$$\begin{aligned}
|T_4| & \leq Y_i\delta(\delta-1)\left|\int_0^1(1-p)^{\delta-2}(1(Y_i < \hat{Q}(p))-1(Y_i < Q(p)))dp\right| \\
& \leq Y_i\delta\left|\int_{\hat{F}(Y_i)}^{F(Y_i)}(\delta-1)(1-p)^{\delta-2}dp\right| \\
& \leq Y_i\sup_y|(1-\hat{F}(y))^{\delta-1}-(1-F(y))^{\delta-1}|.
\end{aligned}$$

This term satisfies (ii) of Lemma A.1 so that because all the other terms satisfy (i) of Lemma A.1 we have the result in (iii). The results in (i) and (ii) follow similarly. For (ii),

$$|\hat{\phi}_i(I_A^\delta) - \phi_i(I_A^\delta)| \leq |\hat{\mu} - \mu| + |\hat{\phi}_i(W^\delta) - \phi_i(W^\delta)|$$

so that the result in (i) follows from (iii) and Lemma A.1 using the fact that $\hat{\mu} - \mu = o_p(1)$.

For (i),

$$\begin{aligned}
|\hat{\phi}_i(\hat{I}_R^\delta) - \phi_i(I_R^\delta)| & \leq \frac{1}{\hat{\mu}}|\hat{\phi}_i(I_A^\delta) - \phi_i(I_A^\delta)| + |\phi_i(I_A^\delta)|\left|\frac{1}{\hat{\mu}} - \frac{1}{\mu}\right| \\
& + \left|\frac{\hat{I}_R^\delta}{\hat{\mu}}\right||\hat{\mu} - \mu| + |Y_i - \mu|\left|\frac{\hat{I}_R^\delta}{\hat{\mu}} - \frac{I_R^\delta}{\mu}\right|
\end{aligned}$$

so that the result follows from (ii) and Lemma A.1 using $\hat{\mu} - \mu = o_p(1)$ the result of Proposition 1(A)(i) and Assumption 1 which implies that $E(|Y_i - \mu|^2) < \infty$. **Q.E.D.**

Proof of Proposition 2(A): Consistency of all three indices follows simply from the CMT. More specifically consider first the estimator,

$$\hat{I}_R^\alpha = 2\left[\int_0^1(p - \hat{L}(p))^\alpha dp\right]^{\frac{1}{\alpha}}$$

and consider the component $(p - \hat{L}(p))^\alpha$. Note that $p - \hat{L}(p) \in [0, 1]$. Now consider the function $g(z) = z^\alpha$ for $z \in [0, 1]$ with $\alpha \geq 1$ fixed. It is easy to show that for $z_1, z_2 \in [0, 1]$, $|g(z_1) - g(z_2)| \leq \alpha|z_1 - z_2|$ so that,

$$|(p - \hat{L}(p))^\alpha - (p - L(p))^\alpha| \leq \alpha|\hat{L}(p) - L(p)|$$

so that,

$$\begin{aligned} \left| \int_0^1 (p - \hat{L}(p))^\alpha dp - \int_0^1 (p - L(p))^\alpha dp \right| &\leq \int_0^1 |(p - \hat{L}(p))^\alpha - (p - L(p))^\alpha| dp \\ &\leq \alpha \sup |\hat{L}(p) - L(p)| \\ &= o_p(1) \end{aligned}$$

using Lemma 3(ii) so that

$$\int_0^1 (p - \hat{L}(p))^\alpha dp \xrightarrow{p} \int_0^1 (p - L(p))^\alpha dp$$

Then it follows by the CMT that

$$\hat{I}_R^\alpha = 2 \left[\int_0^1 (p - \hat{L}(p))^\alpha dp \right]^{\frac{1}{\alpha}} \xrightarrow{p} 2 \left[\int_0^1 (p - L(p))^\alpha dp \right]^{\frac{1}{\alpha}} = I_R^\alpha$$

For the other indices we use the CMT and ST and the facts that,

$$\begin{aligned} \hat{I}_A^\alpha &= \hat{\mu} \hat{I}_R^\alpha \xrightarrow{p} \mu I_R^\alpha = I_A^\alpha \\ \hat{W}^\alpha &= 2\hat{\mu} - \hat{I}_A^\alpha \xrightarrow{p} 2\mu - I_A^\alpha = W^\alpha \end{aligned}$$

In order to derive the limiting distribution results we first focus on \hat{I}_R^α and derive the limiting distribution of the component $\int_0^1 (p - \hat{L}(p))^\alpha dp$ which is a linear functional of $(p - \hat{L}(p))^\alpha$. It suffices to consider the behavior on the interval $(0, 1)$ because $1 - L(1) = 1 - \hat{L}(1) = L(0) = L(1) = 0$. Let C^* be the set of functions that are continuous mappings from $(0, 1)$ to R . Clearly \hat{L} and L belong to C^* . Now consider the mapping $\rho : C^* \rightarrow C^*$ defined by,

$$\rho(L)(p) = p - L(p)$$

Clearly the map ρ is Hadamard Differentiable at L tangentially to C^* with Hadamard

derivative $\rho'(h) = -h$ since for any sequence $t_n \rightarrow 0$, $h_n \rightarrow h \in C^*$ with $L + t_n h_n \in C^*$ we have that,

$$\frac{\rho(L + t_n h_n) - \rho(L)}{t_n} = -h_n \rightarrow -h$$

Define the mapping $g : (0, 1) \rightarrow R$ by,

$$g(x) = x^\alpha$$

This mapping is differentiable at $x \in [0, 1]$ with derivative given by $\alpha x^{\alpha-1}$ which is uniformly continuous on $[0, 1]$ and hence uniformly continuous and bounded on $(0, 1)$. Since for $p \in (0, 1)$ we have that, $0 < p - L(p) < 1$ then by Lemma 3.9.25 of Van der Vaart and Wellner (1996) the mapping $\psi(\rho)(p) = g(p - L(p))$ considered as a mapping from C^* to C^* (where $\rho \in C^*$, so that $\rho(p) = p - L(p)$) is Hadamard differentiable with derivative evaluated at $\gamma \in C^*$,

$$\psi'_\rho(\gamma) = g'(p - L(p))\gamma(p).$$

Consequently the functional delta method gives the result,

$$\sqrt{N}(\psi(\hat{\rho}) - \psi(\rho)) \Rightarrow \psi'(\mathcal{L}) \equiv -\alpha \rho^{\alpha-1} \mathcal{L}$$

which for a fixed value of p gives,

$$\sqrt{N} \left((p - \hat{L}(p))^\alpha - (p - L(p))^\alpha \right) \Rightarrow -\alpha (p - L(p))^{\alpha-1} \mathcal{L}(p)$$

Then it is trivial to see that by linearity of the integral (and hence Hadamard differentiability),

$$\sqrt{N} \left(\int_0^1 (p - \hat{L}(p))^\alpha dp - \int_0^1 (p - L(p))^\alpha dp \right) \Rightarrow \int_0^1 -\alpha (p - L(p))^{\alpha-1} \mathcal{L}(p) dp$$

which is a scalar Gaussian process. To get the final result we note that,

$$\hat{I}_R^\alpha = 2 \left[\int_0^1 (p - \hat{L}(p))^\alpha dp \right]^{\frac{1}{\alpha}} = 2 \hat{A}^{1/\alpha}$$

where we have shown already that,

$$\sqrt{N} (\hat{A} - A) \Rightarrow \int_0^1 -\alpha (p - L(p))^{\alpha-1} \mathcal{L}(p) dp$$

Using this fact a simple application of the standard delta method for scalars gives the result (for some \tilde{A} lying between \hat{A} and A),

$$\begin{aligned}\sqrt{N}(\hat{I}_R^\alpha - I_R^\alpha) &= \frac{2}{\alpha} \tilde{A}^{(1/\alpha)-1} \sqrt{N} (\hat{A} - A) \\ &\Rightarrow -2A^{(1/\alpha)-1} \int_0^1 (p - L(p))^{\alpha-1} \mathcal{L}(p) dp\end{aligned}$$

and the result in (i) follows upon noting that

$$2A^{(1/\alpha)-1} = 2^\alpha 2^{1-\alpha} (A^{1/\alpha})^{1-\alpha} = 2^\alpha (2A^{1/\alpha})^{1-\alpha} = 2^\alpha (I_R^\alpha)^{1-\alpha}$$

For (ii) we note that

$$\begin{aligned}\sqrt{N}(\hat{I}_A^\alpha - I_A^\alpha) &= \hat{\mu} \sqrt{N}(\hat{I}_R^\alpha - I_R^\alpha) + I_R^\alpha \sqrt{N}(\hat{\mu} - \mu) \\ &\Rightarrow -\mu 2A^{(1/\alpha)-1} \int_0^1 (p - L(p))^{\alpha-1} \mathcal{L}(p) dp + I_R^\alpha \mathcal{G}(1) \\ &\equiv 2^\alpha (I_A^\alpha)^{1-\alpha} \int_0^1 (\mu p - G(p))^{\alpha-1} (p \mathcal{G}(1) - \mathcal{G}(p)) dp\end{aligned}$$

using the definition of \mathcal{L} after some manipulations. The result for

$$\hat{W}^\alpha = 2\hat{\mu} - \hat{I}_A^\alpha$$

follows similarly. **Q.E.D.**

Proof of Proposition 2(B): Write,

$$\begin{aligned}|\hat{\phi}_i(I_A^\alpha) - \phi_i(I_A^\alpha)| &\leq 2^\alpha (\hat{I}_A^\alpha)^{1-\alpha} \sum_{j=1}^7 |T_j| \\ &\quad + 2^\alpha ((\hat{I}_A^\alpha)^{1-\alpha} - (I_A^\alpha)^{1-\alpha}) \phi_i^\alpha\end{aligned}$$

where,

$$\begin{aligned}T_1 &= \int_0^1 ((\hat{\mu}p - \hat{G}(p))^{\alpha-1} p \hat{Q}(p) - (\mu p - G(p))^{\alpha-1} p Q(p)) dp \\ T_2 &= \int_0^1 ((\hat{\mu}p - \hat{G}(p))^{\alpha-1} \hat{G}(p) - (\mu p - G(p))^{\alpha-1} G(p)) dp \\ T_3 &= \int_0^1 ((\hat{\mu}p - \hat{G}(p))^{\alpha-1} \hat{Q}(p) - (\mu p - G(p))^{\alpha-1} Q(p)) 1(Y_i < \hat{Q}(p)) dp \\ T_4 &= \int_0^1 ((\mu p - G(p))^{\alpha-1} Q(p) (1(Y_i < \hat{Q}(p)) - 1(Y_i < Q(p)))) dp\end{aligned}$$

$$\begin{aligned}
T_5 &= Y_i \int_0^1 \left((\hat{\mu}p - \hat{G}(p))^{\alpha-1} - (\mu p - G(p))^{\alpha-1} \right) 1(Y_i < \hat{Q}(p)) dp \\
T_6 &= Y_i \int_0^1 \left((\mu p - G(p))^{\alpha-1} (1(Y_i < \hat{Q}(p)) - 1(Y_i < Q(p))) \right) dp \\
T_7 &= \int_0^1 \left((\hat{\mu}p - \hat{G}(p))^{\alpha-1} p \hat{\mu} - (\mu p - G(p))^{\alpha-1} p \mu \right) dp \\
\phi_i^\alpha &= \int_0^1 (\mu p - G(p))^{\alpha-1} (\phi_i(p; G) - p(Y_i - \mu)) dp
\end{aligned}$$

Because of the fact that $2^\alpha (\hat{I}_A^\alpha)^{1-\alpha}$ is stochastically bounded we must show that $|T_j| = o_p(1)$ so that the conditions of Lemma A.1 are satisfied. To deal with T_1 use the inequality,

$$|\hat{a}\hat{b} - ab| = |\hat{a} - a| |\hat{b} - b| + |a| |\hat{b} - b| + |b| |\hat{a} - a|$$

and the following facts,

$$\begin{aligned}
\sup_p |Q(p)| &< \infty \\
\sup_p |(\mu p - G(p))^{\alpha-1}| &\leq \mu < \infty \\
\sup_p |p(\hat{Q}(p) - Q(p))| &\leq \sup_p |(\hat{Q}(p) - Q(p))| = o_p(1)
\end{aligned}$$

which follow from Assumption 1 and Lemma 2(i) and the fact that,

$$\sup_p |(\hat{\mu}p - \hat{G}(p))^{\alpha-1} - (\mu p - G(p))^{\alpha-1}| = o_p(1)$$

This last result follows from Lemma 3(i) and the fact that the function $g(x) = x^{\alpha-1}$ is such that for $1 < \alpha \leq 2$ and $x', x'' \in [0, 1]$,

$$|g(x') - g(x'')| \leq |x' - x''|^{\alpha-1}$$

while for $\alpha > 2$,

$$|g(x') - g(x'')| \leq (\alpha - 1) |x' - x''|$$

Then combining these results we have that,

$$\begin{aligned}
|T_1| &\leq \int_0^1 \sup_p \left| (\hat{\mu}p - \hat{G}(p))^{\alpha-1} p \hat{Q}(p) - (\mu p - G(p))^{\alpha-1} p Q(p) \right| dp \\
&\leq \sup_p \left| (\hat{\mu}p - \hat{G}(p))^{\alpha-1} p \hat{Q}(p) - (\mu p - G(p))^{\alpha-1} p Q(p) \right| \\
&= o_p(1)
\end{aligned}$$

The same arguments apply to the term T_2 and T_7 using the results in Lemma 3. For T_3 and T_5 the same arguments apply once we note that,

$$\sup_p \mathbf{1}(Y_i < \hat{Q}(p)) \leq 1.$$

Next $|T_5| = Y_i o_p(1)$ follows using the fact that,

$$\sup_p |(\mu p - G(p))^{\alpha-1} Q(p)| < \infty$$

and,

$$\begin{aligned} \left| \int_0^1 (\mathbf{1}(Y_i < \hat{Q}(p)) - \mathbf{1}(Y_i < Q(p))) dp \right| &= |\hat{F}(Y_i) - F(Y_i)| \\ &\leq \sup_y |\hat{F}(y) - F(y)| \\ &= o_p(1) \end{aligned}$$

which follows from Lemma 1(i). Finally by Proposition 2(A)(ii) we have that, $(\hat{I}_A^\alpha)^{1-\alpha} - (\hat{I}_A^\alpha)^{1-\alpha} = o_p(1)$ and since

$$|\phi_i^\alpha| \leq C_1 + C_2 Y_i$$

we have that condition (ii) of Lemma A.1 is satisfied by the term $2^\alpha((\hat{I}_A^\alpha)^{1-\alpha} - (\hat{I}_A^\alpha)^{1-\alpha})\phi_i^\alpha$. The results for the other indices follow in a manner that is similar to the proof of (i) and (ii) of Proposition 1(B). **Q.E.D.**

Proof of Proposition 3(A): Given,

$$\hat{P}^\delta = \hat{F}(z) - \frac{1}{z} \delta(\delta - 1) \int_0^1 (1-p)^{\delta-2} \hat{G}(p\hat{F}(z)) dp$$

and the fact that by Lemma 1,

$$\sqrt{N}(\hat{F}(z) - F(z)) \Rightarrow \mathcal{B}(F(z))$$

it remains to show that,

$$\sqrt{N} \int_0^1 (1-p)^{\delta-2} (\hat{G}(p\hat{F}(z)) - G(pF(z))) dp \Rightarrow \int_0^1 (1-p)^{\delta-2} \{Q(pF(z))p\mathcal{B}(F(z)) + \mathcal{G}(pF(z))\} dp.$$

To do this we first write,

$$\begin{aligned}\sqrt{N}(\hat{G}(p\hat{F}(z)) - G(pF(z))) &= \sqrt{N}(\hat{G}(pF(z)) - G(pF(z))) + \sqrt{N}(G(p\hat{F}(z)) - G(pF(z))) \\ &\quad + \sqrt{N}((\hat{G}(p\hat{F}(z)) - G(p\hat{F}(z))) - (\hat{G}(pF(z)) - G(pF(z))))\end{aligned}$$

For the last term we have,

$$\begin{aligned}&\sup_p \left| \sqrt{N} \left((\hat{G}(p\hat{F}(z)) - G(p\hat{F}(z))) - (\hat{G}(pF(z)) - G(pF(z))) \right) \right| \\ &= \sup_p \left| \sqrt{N} \int_{pF(z)}^{p\hat{F}(z)} (\hat{Q}(t) - Q(t)) dt \right| \\ &\leq \sup_p \left| \hat{Q}(p) - Q(p) \right| \sqrt{N} \sup_p |p\hat{F}(z) - pF(z)| \\ &\leq \sup_p \left| \hat{Q}(p) - Q(p) \right| \sqrt{N} |\hat{F}(z) - F(z)| \\ &= o_p(1)\end{aligned}$$

by Lemma 1 and Lemma 3(i) when $\delta = 1$. Let $\theta(F(z)) \in C^*$ be such that $\theta(F(z))(p) = pF(z)$ and $\hat{\theta}(F(z))(p) = p\hat{F}(z)$ we can apply the delta method to the term $\sqrt{N}(G(\hat{\theta}) - G(\theta))$. Note first that,

$$\sqrt{N}(\hat{\theta}(F(z)) - \theta(F(z))) \Rightarrow \theta(\mathcal{B}(F(z)))$$

where for fixed p we have,

$$\sqrt{N}(\hat{\theta}(F(z))(p) - \theta(F(z))(p)) \Rightarrow \theta(\mathcal{B}(F(z))(p) = p\mathcal{B}(F(z))).$$

Then since on the interval $(0, 1)$ the function $G(t)$ has a derivative $Q(t)$ which is continuous on $[0, 1]$ and hence bounded and uniformly continuous on $(0, 1)$ then since for $p \in (0, 1)$, $0 < \theta(F(z))(p) < 1$ then Lemma 3.9.25 of Van der Vaart and Wellner (1996) implies that the map $G(\theta)$ has a Hadamard derivative given by $Q(\theta(p))\gamma(p)$ when evaluated at the function γ that is continuous on $(0, 1)$. Hence the delta method implies that,

$$\sqrt{N}(G(\hat{\theta}(F(z))) - G(\theta(F(z)))) \Rightarrow Q(\theta)\theta(\mathcal{B}(F(z)))$$

which for a fixed value of p gives,

$$\sqrt{N}(G(p\hat{F}(z)) - G(pF(z))) \Rightarrow Q(pF(z))p\mathcal{B}(F(z))$$

For the term $\sqrt{N}(\hat{G}(pF(z)) - G(pF(z)))$ the result in Lemma 3(i) implies that,

$$\sqrt{N}(\hat{G}(\theta(F(z))) - G(\theta(F(z)))) \Rightarrow \mathcal{G}(\theta)$$

so that for fixed p we have that,

$$\sqrt{N}(\hat{G}(pF(z)) - G(pF(z))) \Rightarrow \mathcal{G}(pF(z))$$

Therefore we have that,

$$\sqrt{N}(\hat{G}(\hat{\theta}(F(z))) - G(\theta(F(z)))) \Rightarrow Q(\theta(F(z)))\theta(\mathcal{B}(F(z))) + \mathcal{G}(\theta(F(z)))$$

where for fixed p we have,

$$\sqrt{N}(\hat{G}(p\hat{F}(z)) - G(pF(z))) \Rightarrow Q(pF(z))p\mathcal{B}(F(z)) + \mathcal{G}(pF(z))$$

The result then follows similarly to the proof of Proposition 1(A). **Q.E.D.**

Proof of Proposition 3(B): Note that,

$$|\hat{\phi}_i(P^\delta) - \phi_i(P^\delta)| \leq \frac{1}{z} \sum_{j=1}^5 T_j$$

where,

$$\begin{aligned} T_1 &= \left| \mathbf{1}(Y_i \leq z) \delta(\delta - 1) \int_0^1 (1-p)^{\delta-2} p (\hat{Q}(p\hat{F}(z)) - Q(pF(z))) dp \right| \\ T_2 &= \left| \delta(\delta - 1) \int_0^1 (1-p)^{\delta-2} p (\hat{F}(z) \hat{Q}(p\hat{F}(z)) - F(z) Q(pF(z))) dp \right| \\ T_3 &= \left| \delta(\delta - 1) \int_0^1 (1-p)^{\delta-2} (\hat{G}(p\hat{F}(z)) - G(pF(z))) dp \right| \\ T_4 &= \left| \delta(\delta - 1) \int_0^1 (1-p)^{\delta-2} \mathbf{1}(Y_i < \hat{Q}(p\hat{F}(z))) (\hat{Q}(p\hat{F}(z)) - Q(pF(z))) dp \right| \\ T_5 &= \left| \delta(\delta - 1) \int_0^1 (1-p)^{\delta-2} Q(pF(z)) (\mathbf{1}(Y_i < \hat{Q}(p\hat{F}(z))) - \mathbf{1}(Y_i < Q(pF(z)))) dp \right| \\ T_6 &= \left| Y_i \delta(\delta - 1) \int_0^1 (1-p)^{\delta-2} (\mathbf{1}(Y_i < \hat{Q}(p\hat{F}(z))) - \mathbf{1}(Y_i < Q(pF(z)))) dp \right| \end{aligned}$$

For the first term note that,

$$\hat{Q}(p\hat{F}(z)) - Q(pF(z)) = (\hat{Q}(p\hat{F}(z)) - Q(p\hat{F}(z))) - (Q(p\hat{F}(z)) - Q(pF(z)))$$

so that,

$$\sup_p |\hat{Q}(p\hat{F}(z)) - Q(pF(z))| \leq \sup_p |\hat{Q}(p) - Q(p)| + \sup_p |Q(p\hat{F}(z)) - Q(pF(z))|$$

By Lemma 2(i) we have that,

$$\sup_p |\hat{Q}(p) - Q(p)| = o_p(1)$$

Since Q is uniformly continuous on $[0, 1]$ then for any $\epsilon > 0$ there is a $\delta > 0$ such that if $p', p'' \in [0, 1]$ with $|p' - p''| < \delta$ then $|Q(p') - Q(p'')| < \epsilon$. By Lemma 1(i) we have that $\hat{F}(z) \xrightarrow{p} F(z)$ so that with probability approaching 1, $|p\hat{F}(z) - pF(z)| < \delta$ which implies that with probability approaching 1,

$$\sup_p |Q(p\hat{F}(z)) - Q(pF(z))| < \epsilon$$

and since ϵ is arbitrary we have that,

$$\sup_p |Q(p\hat{F}(z)) - Q(pF(z))| \xrightarrow{p} 0.$$

Given these facts,

$$\begin{aligned} T_1 &\leq 1(Y_i \leq z) \sup_p |\hat{Q}(p\hat{F}(z)) - Q(pF(z))| \delta(\delta - 1) \int_0^1 (1-p)^{\delta-2} p dp \\ &= 1(Y_i \leq z) \sup_p |\hat{Q}(p\hat{F}(z)) - Q(pF(z))| \end{aligned}$$

which satisfies (ii) of Lemma A.1. Similar arguments can be used for the terms T_2 and T_3 . Next, $T_4 \leq T_1$ since $1(Y_i < \hat{Q}(p\hat{F}(z)))$ is less than $1(Y_i < z)$ so that T_4 satisfies (ii) of Lemma A.1. For T_5 we have that since $Q(pF(z)) \leq z$ for $p \leq 1$ then

$$\begin{aligned} T_5 &\leq z \left| \delta(\delta - 1) \int_{\hat{F}(Y_i)/\hat{F}(z)}^{F(Y_i)/F(z)} (1-p)^{\delta-2} dp \right| \\ &\leq z\delta \left| \left(1 - \frac{\hat{F}(Y_i)}{\hat{F}(z)}\right)^{\delta-1} - \left(1 - \frac{F(Y_i)}{F(z)}\right)^{\delta-1} \right| \\ &\leq z\delta \sup_y \left| \left(1 - \frac{\hat{F}(y)}{\hat{F}(z)}\right)^{\delta-1} - \left(1 - \frac{F(y)}{F(z)}\right)^{\delta-1} \right| \\ &= o_p(1) \end{aligned}$$

using the fact that z and δ are fixed and arguments similar to those used in the proof of Proposition 1(B). Therefore the term T_5 satisfies (ii) of Lemma A.1. The same argument applies to the final term T_6 . **Q.E.D.**

Proof of Proposition 4: Because (recalling that $\delta = 1$),

$$E\left(\frac{1}{z}(z - Y_i) \cdot 1(Y_i \leq z)\right) = P^\delta$$

and,

$$\begin{aligned} 0 &\leq \frac{1}{z}(z - Y_i) \cdot 1(Y_i \leq z) \\ &\leq \frac{1}{z}(z - y_l) \leq 1 \end{aligned}$$

then we have by the Strong Law of Large Numbers that, $\hat{P}^\delta \xrightarrow{a.s.} P^\delta$, so that $\hat{P}^\delta \xrightarrow{p} P^\delta$. Also by the Lindeberg-Levy Central Limit Theorem we have that,

$$\sqrt{N}(\hat{P}^\delta - P^\delta) \xrightarrow{d} N(0, V(\hat{P}^\delta))$$

where,

$$V(\hat{P}^\delta) = E\left(\frac{1}{z^2}(z - Y_i)^2 \cdot 1(Y_i \leq z)\right) - (P^\delta)^2.$$

Finally the LLN implies that,

$$\frac{1}{Nz^2} \sum_{i=1}^N (z - Y_i)^2 \cdot 1(Y_i \leq z) \xrightarrow{p} E\left(\frac{1}{z^2}(z - Y_i)^2 \cdot 1(Y_i \leq z)\right)$$

so that we have,

$$\hat{V}(\hat{P}^\delta) \xrightarrow{p} V(\hat{P}^\delta).$$

Q.E.D.

Proof of Proposition 5(A): The results of Lemma 1 and Proposition 3(A) imply the following,

$$\begin{aligned} \sqrt{N}(\hat{F}(z) - F(z)) &\Rightarrow \mathcal{B}(F(z)) \\ \sqrt{N}(\hat{G}(\hat{F}(z)) - G(F(z))) &\Rightarrow z\mathcal{B}(F(z)) + \mathcal{G}(F(z)) \\ \sqrt{N}(\hat{G}(\hat{\theta}(F(z))) - G(\theta(F(z)))) &\Rightarrow Q(\theta)\theta(\mathcal{B}(F(z)) + \mathcal{G}(\theta)) \end{aligned}$$

recalling the definition of θ used in the proof of Proposition 3(A). Then the last result above implies that for a fixed p ,

$$\sqrt{N}(\hat{G}(p\hat{F}(z)) - G(pF(z))) \Rightarrow Q(pF(z))p\mathcal{B}(F(z)) + \mathcal{G}(pF(z))$$

Then the result of the Proposition follows similar to the result in Proposition 2(A) after noting that,

$$\sqrt{N} \left((\hat{\theta}(G(F(z))) - \hat{G}(\hat{\theta}(F(z)))) - (\theta(G(F(z))) - G(\theta(F(z)))) \right) \Rightarrow \Upsilon$$

$$\Upsilon \equiv \theta(z\mathcal{B}(F(z)) + \mathcal{G}(F(z))) - Q(\theta)\theta(\mathcal{B}(F(z)) + \mathcal{G}(\theta))$$

so that for fixed p we have that,

$$\sqrt{N} \left((p\hat{G}(\hat{F}(z)) - \hat{G}(p\hat{F}(z))) - (pG(F(z)) - G(pF(z))) \right) \Rightarrow \Upsilon(p)$$

where,

$$\Upsilon(p) \equiv p(z - Q(pF(z)))\mathcal{B}(F(z)) + p\mathcal{G}(F(z)) - \mathcal{G}(pF(z)).$$

Q.E.D.

Proof of Proposition 5(B): This result follows in a similar fashion to the proof of Proposition 2(B) with adjustments similar to those used for the S-gini poverty index in Proposition 3(B). **Q.E.D.**

Appendix B: Computation of Influence Curves

Throughout the Appendix we use some of the facts discussed in Section 5. Recall that we are using the shorthand $\hat{p}_j = \hat{F}(y_j)$ and that $\hat{\pi}_j = \hat{p}_j - \hat{p}_{j-1}$. Let $\hat{p}_0 = 0$. Then interval $(\hat{p}_{j-1}, \hat{p}_j]$, $\hat{Q}(p) = y_j$. Also on the same interval,

$$\hat{G}(p) = (p - \hat{p}_{j-1})y_j + \sum_{l=1}^{j-1} \hat{\pi}_l y_l = y_j p + a_j$$

where by convention $a_1 = 0$ so that on the interval $(0, \hat{p}_1]$ we have that $\hat{G}(p) = py_1$. As in the calculations performed in Section 4.3 we use the fact that,

$$\int_0^1 = \sum_{j=1}^{\hat{N}} \int_{\hat{p}_{j-1}}^{\hat{p}_j}$$

along with the definitions given above to calculate the estimates of the influence curves for the indices.

Influence Curves for S-gini indices:

The key component of the influence curve for the S-gini related indices is the term,

$$\delta(\delta - 1) \int_0^1 (1 - p)^{\delta-2} \hat{\phi}_i(p; G) dp = \sum_{l=1}^4 \hat{I}_l^S$$

Then we are required to compute the following;

$$\begin{aligned} \hat{I}_1^S &= \delta(\delta - 1) \int_0^1 (1 - p)^{\delta-2} p \hat{Q}(p) dp \\ \hat{I}_2^S &= -\delta(\delta - 1) \int_0^1 (1 - p)^{\delta-2} \hat{G}(p) dp \\ \hat{I}_3^S &= -\delta(\delta - 1) \int_0^1 (1 - p)^{\delta-2} \hat{Q}(p) 1(Y_i < \hat{Q}(p)) dp \\ \hat{I}_4^S &= \delta(\delta - 1) \int_0^1 (1 - p)^{\delta-2} Y \cdot 1(Y_i < \hat{Q}(p)) dp \end{aligned}$$

each of which is considered in turn:

(i)

$$\hat{I}_1^S = \sum_{j=1}^{\hat{N}} y_j \hat{d}_j^1$$

where,

$$\hat{d}_j^1 = -\delta \left(\hat{p}_j (1 - \hat{p}_j)^{\delta-1} - \hat{p}_{j-1} (1 - \hat{p}_{j-1})^{\delta-1} \right) - \left((1 - \hat{p}_j)^\delta - (1 - \hat{p}_{j-1})^\delta \right)$$

(ii)

$$\hat{I}_2^S = - \sum_{j=1}^N \left(\hat{d}_j^2 + \hat{d}_j^3 y_j \right)$$

where,

$$\hat{d}_j^2 = -\delta \left((1 - \hat{p}_j)^{\delta-1} \hat{G}(\hat{p}_j) - (1 - \hat{p}_{j-1})^{\delta-1} \hat{G}(\hat{p}_{j-1}) \right)$$

$$\hat{d}_j^3 = -\left((1 - \hat{p}_j)^\delta - (1 - \hat{p}_{j-1})^\delta \right)$$

with

$$\hat{G}(\hat{p}_j) = \sum_{l=1}^j \hat{\pi}_l y_l = \hat{p}_j y_j + a_j = \hat{p}_j y_{j+1} + a_{j+1}$$

(iii)

$$\hat{I}_3^S = - \sum_{j=1}^N \mathbf{1}(Y_i < y_j) y_j \hat{d}_j^4$$

$$\hat{d}_j^4 = -\delta \left((1 - \hat{p}_j)^{\delta-1} - (1 - \hat{p}_{j-1})^{\delta-1} \right)$$

(iv)

$$\hat{I}_4^S = Y_i \sum_{j=1}^N \mathbf{1}(Y_i < y_j) \hat{d}_j^4$$

Then we can compute the S-gini influence curves using:

$$\begin{aligned} \hat{\phi}_i(\hat{W}^\delta) &= \sum_{l=1}^4 \hat{I}_l^S \\ \hat{\phi}_i(\hat{I}_A^\delta) &= (Y_i - \hat{\mu}) - \hat{\phi}_i(W^\delta) \\ \hat{\phi}_i(\hat{I}_R^\delta) &= -\frac{1}{\hat{\mu}} \hat{\phi}_i(W^\delta) - \frac{1}{\hat{\mu}} (Y_i - \hat{\mu})(1 - \hat{I}_R^\delta) \end{aligned}$$

Influence curves for E-gini indices:

For the E-gini indices the key component has the form,

$$\frac{1}{\alpha} \int_0^1 \alpha(\mu \cdot p - G(p))^{\alpha-1} (p(Y_i - \hat{\mu}) - \hat{\phi}_i(p; G)) dp = \frac{1}{\alpha} \sum_{l=1}^4 \hat{I}_l^E$$

where

$$\begin{aligned} \hat{I}_1^E &= \int_0^1 \alpha(\hat{\mu} \cdot p - \hat{G}(p))^{\alpha-1} p(Y_i - \hat{\mu}) dp \\ \hat{I}_2^E &= - \int_0^1 \alpha(\hat{\mu} \cdot p - \hat{G}(p))^{\alpha-1} p \hat{Q}(p) dp \\ \hat{I}_3^E &= \int_0^1 \alpha(\hat{\mu} \cdot p - \hat{G}(p))^{\alpha-1} \hat{G}(p) dp \\ \hat{I}_4^E &= \int_0^1 \alpha(\hat{\mu} \cdot p - \hat{G}(p))^{\alpha-1} (\hat{Q}(p) - Y_i) 1(Y_i < \hat{Q}(p)) dp \end{aligned}$$

(i)

$$\hat{I}_1^E = (Y_i - \hat{\mu}) \sum_{j=1}^{\hat{N}} \left\{ (\hat{e}_j^1 - \hat{e}_j^2) 1(\hat{\mu} \neq y_j) + \hat{e}_j^3 1(\hat{\mu} = y_j) \right\}$$

where,

$$\begin{aligned} \hat{e}_j^1 &= \frac{1}{b_j} ((b_j \hat{p}_j - a_j)^\alpha \hat{p}_j - (b_j \hat{p}_{j-1} - a_j)^\alpha \hat{p}_{j-1}) \\ \hat{e}_j^2 &= \frac{1}{b_j^2 (\alpha + 1)} \left((b_j \hat{p}_j - a_j)^{\alpha+1} - (b_j \hat{p}_{j-1} - a_j)^{\alpha+1} \right) \\ \hat{e}_j^3 &= \alpha (-a_j)^{\alpha-1} \left(\frac{\hat{p}_j^2}{2} - \frac{\hat{p}_{j-1}^2}{2} \right) \end{aligned}$$

with, $b_j = \hat{\mu} - y_j$ and $a_j = \sum_{l=1}^{j-1} \hat{\pi}_l (y_l - y_j)$

(ii)

$$\hat{I}_2^E = - \sum_{j=1}^{\hat{N}} y_j \left\{ (\hat{e}_j^1 - \hat{e}_j^2) 1(\hat{\mu} \neq y_j) + \hat{e}_j^3 1(\hat{\mu} = y_j) \right\}$$

(iii)

$$\begin{aligned}\hat{I}_3^E &= \sum_{j=1}^{\dot{N}} y_j \left\{ (\hat{e}_j^1 - \hat{e}_j^2) 1(\hat{\mu} \neq y_j) + \hat{e}_j^3 1(\hat{\mu} = y_j) \right\} \\ &\quad + \sum_{j=1}^{\dot{N}} a_j \left\{ \hat{e}_j^4 1(\hat{\mu} \neq y_j) + \hat{e}_j^5 1(\hat{\mu} = y_j) \right\}\end{aligned}$$

where,

$$\begin{aligned}\hat{e}_j^4 &= \frac{1}{b_j} \left((b_j \hat{p}_j - a_j)^\alpha - (b_j \hat{p}_{j-1} - a_j)^\alpha \right) \\ \hat{e}_j^5 &= \alpha (-a_j)^{\alpha-1} (\hat{p}_j - \hat{p}_{j-1})\end{aligned}$$

(iv)

$$\hat{I}_4^E = \sum_{j=1}^{\dot{N}} 1(Y_i < y_j) (y_j - Y_i) \left\{ \hat{e}_j^4 1(\hat{\mu} \neq y_j) + \hat{e}_j^5 1(\hat{\mu} = y_j) \right\}$$

These four terms can be combined and simplified as follows:

$$\sum_{l=1}^4 \hat{I}_l^E = \sum_{j=1}^{\dot{N}} \hat{g}_j^1 \hat{h}_j^1 + \sum_{j=1}^{\dot{N}} \hat{g}_j^2 \hat{h}_j^2$$

where,

$$\begin{aligned}\hat{g}_j^1 &= (\hat{e}_j^1 - \hat{e}_j^2) 1(\hat{\mu} \neq y_j) + \hat{e}_j^3 1(\hat{\mu} = y_j) \\ \hat{g}_j^2 &= \hat{e}_j^4 1(\hat{\mu} \neq y_j) + \hat{e}_j^5 1(\hat{\mu} = y_j) \\ \hat{h}_j^1 &= (Y_i - \hat{\mu}) \\ \hat{h}_j^2 &= (1(Y_i < y_j) (y_j - Y_i) + a_j)\end{aligned}$$

Then using these results we have the influence curves for the E-gini indices as:

$$\begin{aligned}\hat{\phi}_i(I_A^\alpha) &= \frac{2^\alpha}{\alpha} (\hat{I}_A^\alpha)^{1-\alpha} \sum_{l=1}^4 \hat{I}_l^E(T_1) \\ \hat{\phi}_i(W^\alpha) &= 2(Y_i - \hat{\mu}) - \hat{\phi}_i(I_A^\alpha) \\ \hat{\phi}_i(I_R^\alpha) &= \frac{1}{\hat{\mu}} \phi_i(I_A^\alpha) - \frac{I_R^\alpha}{\hat{\mu}} (Y_i - \hat{\mu})\end{aligned}$$

Influence Curves for Poverty Indices

As was the case with the calculation of the indices themselves we use a change of variables and the fact that,

$$\int_0^{\hat{F}(z)} = \sum_{j=1}^{N(z)} \int_{\hat{p}_{j-1}}^{\hat{p}_j}$$

For the S-gini related index we must compute,

$$\begin{aligned} \hat{\phi}_i(P^\delta) &= \hat{\phi}_i(z; F) - \frac{\delta(\delta-1)}{z} \int_0^1 (1-p)^{\delta-2} \left(p\hat{Q}(p\hat{F}(z))\hat{\phi}_i(z; F) + \hat{\phi}_i(p\hat{F}(z); G) \right) dp \\ &= \hat{\phi}_i(z; F) - \frac{\hat{\phi}_i(z; F)}{z} \delta(\delta-1) \int_0^1 (1-p)^{\delta-2} p\hat{Q}(p\hat{F}(z)) dp \\ &\quad - \frac{\delta(\delta-1)}{z} \int_0^1 (1-p)^{\delta-2} \hat{\phi}_i(p\hat{F}(z); G) dp \\ &= \hat{\phi}_i(z; F) - \frac{\hat{\phi}_i(z; F)}{z} \tilde{I}_1^S - \frac{1}{z} \sum_{l=2}^5 \tilde{I}_l^S \end{aligned}$$

where,

$$\begin{aligned} \tilde{I}_1^S &= \delta(\delta-1) \int_0^1 (1-p)^{\delta-2} p\hat{Q}(p\hat{F}(z)) dp \\ \tilde{I}_2^S &= \delta(\delta-1) \int_0^1 (1-p)^{\delta-2} p\hat{F}(z)Q(p\hat{F}(z)) dp \\ \tilde{I}_3^S &= -\delta(\delta-1) \int_0^1 (1-p)^{\delta-2} \hat{G}(p\hat{F}(z)) dp \\ \tilde{I}_4^S &= -\delta(\delta-1) \int_0^1 (1-p)^{\delta-2} Q(p\hat{F}(z))1(Y_i < \hat{Q}(p\hat{F}(z))) dp \\ \tilde{I}_5^S &= \delta(\delta-1) \int_0^1 (1-p)^{\delta-2} Y_i.1(Y_i < \hat{Q}(p\hat{F}(z))) dp \end{aligned}$$

We consider each term in turn and use a change of variables:

(i)

$$\begin{aligned} \tilde{I}_1^S &= \frac{\delta(\delta-1)}{\hat{F}(z)^\delta} \int_0^{\hat{F}(z)} (\hat{F}(z) - p)^{\delta-2} p\hat{Q}(p) dp \\ &= \frac{1}{\hat{F}(z)^\delta} \sum_{j=1}^{N(z)} y_j \tilde{d}_j^1 \end{aligned}$$

where,

$$\tilde{d}_j^1 = -\delta \left(\hat{p}_j \left(\hat{F}(z) - \hat{p}_j \right)^{\delta-1} - \hat{p}_{j-1} \left(\hat{F}(z) - \hat{p}_{j-1} \right)^{\delta-1} \right) - \left(\left(\hat{F}(z) - \hat{p}_j \right)^\delta - \left(\hat{F}(z) - \hat{p}_{j-1} \right)^\delta \right)$$

(ii) Similarly,

$$\begin{aligned}\tilde{I}_2^S &= \frac{\hat{F}(z)\delta(\delta-1)}{\hat{F}(z)^\delta} \int_0^{\hat{F}(z)} (\hat{F}(z)-p)^{\delta-2} p \hat{Q}(p) dp \\ &= \hat{F}(z) \tilde{I}_1^S\end{aligned}$$

(iii) Next

$$\begin{aligned}\tilde{I}_3^S &= -\frac{\delta(\delta-1)}{\hat{F}(z)^{\delta-1}} \int_0^{\hat{F}(z)} (\hat{F}(z)-p)^{\delta-2} \hat{G}(p) dp \\ &= -\frac{1}{\hat{F}(z)^{\delta-1}} \sum_{j=1}^{N(z)} (\tilde{d}_j^2 + \tilde{d}_j^3 y_j)\end{aligned}$$

where,

$$\tilde{d}_j^2 = -\delta \left((\hat{F}(z) - \hat{p}_j)^{\delta-1} \hat{G}(\hat{p}_j) - (\hat{F}(z) - \hat{p}_{j-1})^{\delta-1} \hat{G}(\hat{p}_{j-1}) \right)$$

$$\tilde{d}_j^3 = -\left((\hat{F}(z) - \hat{p}_j)^\delta - (\hat{F}(z) - \hat{p}_{j-1})^\delta \right)$$

with

$$\hat{G}(\hat{p}_j) = \sum_{l=1}^{j-1} \hat{\pi}_l y_l$$

(iv)

$$\tilde{I}_4^S = -\frac{1}{\hat{F}(z)^{\delta-1}} \sum_{j=1}^{N(z)} 1(Y_i < y_j) y_j d_j^4$$

where,

$$\tilde{d}_j^4 = -\delta \left((\hat{F}(z) - \hat{p}_j)^{\delta-1} - (\hat{F}(z) - \hat{p}_{j-1})^{\delta-1} \right)$$

(v)

$$\tilde{I}_5^S = \frac{Y_i}{\hat{F}(z)^{\delta-1}} \sum_{j=1}^{N(z)} 1(Y_i < y_j) \tilde{d}_j^4$$

For the E-gini based poverty index the influence curve is given by

$$\begin{aligned}\hat{\phi}_i(P^\alpha) &= -1(Y_i \leq z) - \frac{2}{z}\hat{\phi}_i(F(z); G) \\ &\quad + \frac{1}{z}[\hat{T}_E^z]^\frac{1}{\alpha}-1 \int_0^1 (p\hat{G}(\hat{F}(z)) - \hat{G}(p\hat{F}(z)))^{\alpha-1} \phi_i(p\hat{F}(z); \Upsilon) dp \\ &= -1(Y_i \leq z) + \tilde{I}_1^E + \frac{1}{z\alpha} [\hat{T}_E^z]^\frac{1}{\alpha}-1 \sum_{l=2}^7 \tilde{I}_l^E\end{aligned}$$

where \hat{T}_E^z was already calculated in Section 6.2,

$$\begin{aligned}\tilde{I}_1^E &= -\frac{2}{z} \left((\hat{F}(z)z - a_{N(z)+1}) + 1(Y_i < z)(z - Y_i) \right) \\ \tilde{I}_2^E &= z \int_0^1 \alpha(p\hat{G}(\hat{F}(z)) - \hat{G}(p\hat{F}(z)))^{\alpha-1} p dp \\ \tilde{I}_3^E &= -\int_0^1 \alpha(p\hat{G}(\hat{F}(z)) - \hat{G}(p\hat{F}(z)))^{\alpha-1} p \hat{Q}(p\hat{F}(z)) dp \\ \tilde{I}_4^E &= \left(\hat{F}(z)(z - \hat{\mu}(z)) - 1(Y_i < z)(z - Y_i) \right) \int_0^1 \alpha(p\hat{G}(\hat{F}(z)) - \hat{G}(p\hat{F}(z)))^{\alpha-1} p dp \\ \tilde{I}_5^E &= -\int_0^1 \alpha(p\hat{G}(\hat{F}(z)) - \hat{G}(p\hat{F}(z)))^{\alpha-1} p \hat{F}(z) \hat{Q}(p\hat{F}(z)) dp \\ \tilde{I}_6^E &= \int_0^1 \alpha(p\hat{G}(\hat{F}(z)) - \hat{G}(p\hat{F}(z)))^{\alpha-1} \hat{G}(p\hat{F}(z)) dp \\ \tilde{I}_7^E &= \int_0^1 \alpha(p\hat{G}(\hat{F}(z)) - \hat{G}(p\hat{F}(z)))^{\alpha-1} 1(Y_i < \hat{Q}(p\hat{F}(z))) (\hat{Q}(p\hat{F}(z)) - Y_i) dp\end{aligned}$$

(i) The term \tilde{I}_1^E has already been given.

(ii)

$$\tilde{I}_2^E = \frac{z}{\hat{F}(z)^2} \sum_{j=1}^{N(z)} \left\{ (\tilde{e}_j^1 - \tilde{e}_j^2) 1(\hat{\mu}(z) \neq y_j) + \tilde{e}_j^3 1(\hat{\mu}(z) = y_j) \right\}$$

where,

$$\begin{aligned}\tilde{e}_j^1 &= \frac{1}{b_j} \left((\tilde{b}_j \hat{p}_j - a_j)^\alpha \hat{p}_j - (\tilde{b}_j \hat{p}_{j-1} - a_j)^\alpha \hat{p}_{j-1} \right) \\ \tilde{e}_j^2 &= \frac{1}{b_j^2 (\alpha + 1)} \left((\tilde{b}_j \hat{p}_j - a_j)^{\alpha+1} - (\tilde{b}_j \hat{p}_{j-1} - a_j)^{\alpha+1} \right) \\ \tilde{e}_j^3 &= \alpha (-a_j)^{\alpha-1} \left(\frac{\hat{p}_j^2}{2} - \frac{\hat{p}_{j-1}^2}{2} \right)\end{aligned}$$

with, $\tilde{b}_j = \hat{\mu}(z) - y_j$ and $a_j = \sum_{l=1}^{j-1} \hat{\pi}_l (y_l - y_j)$

(iii)

$$\tilde{I}_3^E = -\frac{1}{\hat{F}(z)^2} \sum_{j=1}^{N(z)} y_j \left\{ (\tilde{e}_j^1 - \tilde{e}_j^2) 1(\hat{\mu}(z) \neq y_j) + \tilde{e}_j^3 1(\hat{\mu}(z) = y_j) \right\}$$

(iv)

$$\tilde{I}_4^E = \frac{1}{\hat{F}(z)^2} \left(\hat{F}(z)(z - \hat{\mu}(z)) - 1(Y_i < z)(z - Y_i) \right) \sum_{j=1}^{N(z)} \left\{ (\tilde{e}_j^1 - \tilde{e}_j^2) 1(\hat{\mu}(z) \neq y_j) + \tilde{e}_j^3 1(\hat{\mu}(z) = y_j) \right\}$$

(v)

$$\tilde{I}_5^E = -\frac{\hat{F}(z)}{\hat{F}(z)^2} \sum_{j=1}^{N(z)} y_j \left\{ (\tilde{e}_j^1 - \tilde{e}_j^2) 1(\hat{\mu}(z) \neq y_j) + \tilde{e}_j^3 1(\hat{\mu}(z) = y_j) \right\}$$

(vi)

$$\begin{aligned} \tilde{I}_6^E &= \frac{1}{\hat{F}(z)} \sum_{j=1}^{N(z)} y_j \left\{ (\tilde{e}_j^1 - \tilde{e}_j^2) 1(\hat{\mu}(z) \neq y_j) + \tilde{e}_j^3 1(\hat{\mu}(z) = y_j) \right\} \\ &\quad - \frac{1}{\hat{F}(z)} \sum_{j=1}^{N(z)} a_j \left\{ \tilde{e}_j^4 1(\hat{\mu}(z) \neq y_j) + \tilde{e}_j^5 1(\hat{\mu}(z) = y_j) \right\} \end{aligned}$$

where,

$$\begin{aligned} \tilde{e}_j^4 &= \frac{1}{\tilde{b}_j} \left((\tilde{b}_j \hat{p}_j - a_j)^\alpha - (\tilde{b}_j \hat{p}_{j-1} - a_j)^\alpha \right) \\ \tilde{e}_j^5 &= \alpha (-a_j)^{\alpha-1} (\hat{p}_j - \hat{p}_{j-1}) \end{aligned}$$

(vii)

$$\tilde{I}_7^E = \sum_{j=1}^{N(z)} 1(Y_i < y_j)(y_j - Y_i) \left\{ \tilde{e}_j^4 1(\hat{\mu}(z) \neq y_j) + \tilde{e}_j^5 1(\hat{\mu}(z) = y_j) \right\}$$

These terms can be combined as follows,

$$\sum_{l=2}^7 \tilde{I}_l^E = \sum_{j=1}^{N(z)} \tilde{g}_j^1 \tilde{h}_j^1 + \sum_{j=1}^{N(z)} \tilde{g}_j^2 \tilde{h}_j^2$$

where,

$$\begin{aligned}\tilde{g}_j^1 &= (\tilde{e}_j^1 - \tilde{e}_j^2)1(\hat{\mu}(z) \neq y_j) + \tilde{e}_j^3 1(\hat{\mu}(z) = y_j) \\ \tilde{g}_j^2 &= \tilde{e}_j^4 1(\hat{\mu}(z) \neq y_j) + \tilde{e}_j^5 1(\hat{\mu}(z) = y_j) \\ \tilde{h}_j^1 &= \frac{1}{\hat{F}(z)^2} (z - y_j + \hat{F}(z)(z - \hat{\mu}(z)) - 1(Y_i < z)(z - Y_i)) \\ \tilde{h}_j^2 &= \frac{\hat{F}(z)}{\hat{F}(z)^2} (1(Y_i < y_j)(y_j - Y_i) + a_j)\end{aligned}$$

References

- [1] Atkinson,A.B. (1970) “On the measurement of inequality” *Journal of Economic Theory*, 2, 244-63.
- [2] Barrett,G.F and K.Pendakur (1995) “The asymptotic distribution of the generalized Gini indices of inequality” *Canadian Journal of Economics*, 28, 1042-55.
- [3] Beach,C.M., and R.Davidson (1983),“Distribution-free statistical inference with Lorenz curves and income shares” *Review of Economic Studies*, 50, 723-35.
- [4] Bishop,J.A., S.Chakraborti and P.D.Thistle (1989),“An asymptotically distribution-free test for Sen’s welfare index” *Oxford Bulletin of Economics and Statistics*, 52, 105-13.
- [5] Bishop,J.A., J.P.Formby and B.Zheng (1997),“Statistical Inference and the Sen Index of Poverty” *International Economic Review*, 38(2), 381-387.
- [6] Blackorby, C., and D.Donaldson (1978) “Measures of relative equality and their meaning in terms of social welfare” *Journal of Economic Theory*, 14, 665-85.
- [7] —, (1980a) “A theoretical treatment of indices of absolute inequality” *International Economic Review*, 21, 107-36.
- [8] —, (1980b) “Ethical indices for the measurement of poverty” *Econometrica*, 48, 1053-60.
- [9] Chakravarty,S.R. (1988) “Extended Gini indices of inequality” *International Economic Review*, 29, 147-56.
- [10] Csörgó, M (1983) *Quantile Processes with Applications*, (Regional Conference Series on Applied Mathematics), S.I.A.M. Philadelphia.
- [11] Csörgó, M., S. Csörgó and L. Horváth (1986) *An Asymptotic Theory for Empirical Reliability and Concentration Processes*, Springer-Verlag Lecture Notes in Statistics, New York.

- [12] Gottschalk, P. and T.M. Smeeding (1997) "Cross-National Comparisons of Earnings and Income Inequality" *Journal of Economic Literature*, 35(2), 633-687.
- [13] Davidson, R. and J-Y. Duclos (1997) "Statistical Inference for the Measurement of the Incidence of Taxes and Transfers" *Econometrica*, 65, 1453-66.
- [14] Donaldson, D. and J.A. Weymark (1980) "A single parameter generalization of the Gini indices of inequality" *Journal of Economic Theory*, 22, 67-86.
- [15] —, (1983), "Ethically flexible Gini indices for income distributions in the Continuum" *Journal of Economic Theory*, 29, 353-58.
- [16] Fernholz, L. T. (1983) *von Mises Calculus for Statistical Functionals*, Springer-Verlag Lecture Notes in Statistics, New York.
- [17] Foster, J.E., J. Greer and E. Thorbecke (1984) "A class of decomposable poverty measures" *Econometrica*, 52, 761-66.
- [18] Gail, M. H. and J. L. Gastwirth (1978) "A Scale-Free Goodness-of-Fit Test for the Exponential Distribution Based on the Lorenz Curve" *Journal of the American Statistical Association*, **73**, 787-793.
- [19] Goldie, C. M. (1977) "Convergence Theorems for Empirical Lorenz Curves and Their Inverses", *Advances in Applied Probability*, **9**, 765-791.
- [20] Hanratty, M. J. and R. M. Blank (1992) "Down and out in North America: Recent trends in Poverty Rates in the United States and Canada" *Quarterly Journal of Economics*, 107(1), 233-54.
- [21] Huber, P. J. (1981) *Robust Statistics*, John Wiley and Sons, New York.
- [22] Kakwani, N. (1980) "On a class of poverty measures" *Econometrica*, 48, 427-46.
- [23] Kolm, S.C. (1969) "The optimal production of social justice" in J. Margolis and H. Guitton eds., *Public Economics*, Macmillan, London.

- [24] Lambert, P.J. (1993) *The Distribution and Redistribution of Income: A Mathematical Analysis*, Manchester University Press, Manchester, 2nd edition.
- [25] Levy, F. and R.J.Murnane (1992) "U.S. Earnings Levels and Earnings Inequality: A Review of Recent Trends and Proposed Explanations" *Journal of Economic Literature*, 30, 1333-81.
- [26] Rongve,I. and C.M.Beach (1997) "Estimation and inference for normative inequality indices of distributions in the continuum" *International Economic Review*, 38(1), 83-96.
- [27] Sen, A. (1973) *On Economic Inequality*, Claredon Press, Oxford.
- [28] Sen,A. (1976) "Poverty: an ordinal approach to measurement" *Econometrica*, 44, 219-31.
- [29] Shorrocks, A.F. (1983) "Ranking income distributions" *Economica*, 50, 3-17.
- [30] Slesnick, D. (1993) 'Gaining Ground: Poverty in the Postwar United States' *Journal of Political Economy*, 101, 1-38.
- [31] Van der Vaart, A. W. and J. A. Wellner (1996) *Weak Convergence and Empirical Processes: With Applications to Statistics*, Springer-Verlag New York.
- [32] Yitzhaki, S. (1979) "Relative deprivation and the Gini coefficient" *Quarterly Journal of Economics*, 93, 321-4.
- [33] Yitzhaki, S. (1983) "On an extension of the Gini inequality index" *International Economic Review*, 24, 617-28.