1997

Distribution-Free Statistical Inferences for Testing Marginal Changes in Inequality Indices

Buhong Zheng

Brian Cushing

ELSEVIER

Journal of Econometrics 101 (2001) 315–335

JOURNAL OF Econometrics

www.elsevier.nl/locate/econbase

Statistical inference for testing inequality indices with dependent samples

Buhong Zheng[a,*], Brian J. Cushing[b]

[a]*Department of Economics, University of Colorado at Denver, Campus Box 181, P.O. Box 173364, Denver, CO 80217-3364, USA*

[b]*Department of Economics, West Virginia University, Morgantown, WV, USA*

Received 1 July 1997; received in revised form 27 September 2000; accepted 2 October 2000

Abstract

This paper develops asymptotically distribution-free inference for testing inequality indices with dependent samples. It considers the interpolated Gini coefficient and the generalized entropy class, which includes several commonly used inequality indices. We first establish inference tests for changes in inequality indices with completely dependent samples (i.e., matched pairs) and then generalize the inference procedures to cases with partially dependent samples. The effects of sample dependency on standard errors of inequality changes are examined through simulation studies as well as through applications to the CPS and PSID data. © 2001 Published by Elsevier Science S.A.

JEL classification: C40; D63

Keywords: Generalized entropy indices; Gini coefficient; Theil indices; Asymptotic distribution; Statistical inference; Dependent samples

1. Introduction

When comparing inequality among different income distributions, researchers often employ summary measures of inequality such as the Gini

* Corresponding author. Tel.: +1-303-556-4413; fax: +1-303-556-3547.

E-mail address: buhong.zheng@cudenver.edu (B. Zheng).

0304-4076/01/$ - see front matter

PII: S0304-4076(00)00087-7

coefficient, coefficient of variation and the generalized entropy class of indices. Since sample data are frequently used to estimate inequality indices of populations, it is desirable to apply statistical inference procedures to test the robustness of the comparisons. Gastwirth (1974), Gail and Gastwirth (1978), Gastwirth and Gail (1985), Cowell (1989) and Thistle (1990), among others, have provided the large sample properties for several commonly used inequality indices. They show that estimates of inequality indices are asymptotically normal and, hence, conventional inference procedures can be applied straightforwardly. In income distribution studies, it is becoming standard practice to calculate standard errors for estimates of inequality indices and conduct inference tests.

Conventional inference procedures usually require that samples be independently drawn. Many frequently used income data, however, are dependent in that they have an overlap between consecutive years, thus containing information about a cross-section of individuals at two or more points in time. Examples of dependent samples include the current population survey (CPS), the survey of income and program participation (SIPP) and the panel study of income dynamics (PSID). The CPS sample rotates every 2 years, with each household surveyed in two consecutive years. Each year, about one-half of the households are dropped from the sample and are replaced by a new panel of households. The SIPP sample consists of a continuous series of national panels and the duration of each panel ranges from 2.5 years to 4 years. The PSID is a longitudinal study of U.S. individuals and families. Starting with a national sample of about 5000 households in 1968, the PSID has re-interviewed individuals from those households every year. Thus, both CPS and SIPP are partially dependent data while PSID can be regarded as completely dependent data (matched pairs).

Although the problem of sample dependency has been acknowledged in the literature, it has not been properly addressed and no method of correction has been proposed. Researchers have either used samples as if they were independently drawn, chosen a longer time span (e.g., the CPS samples are independent if they are more than 2 years apart), or used the non-matched portions of the samples. For example, Bishop et al. (1991a, b) tested annual changes in U.S. income distribution using the CPS data with conventional inference procedures. In testing German income inequality using the German socio-economic panel (GSOEP) data, Schluter (1996) used conventional inference procedures but acknowledged the problem of sample dependency and the inadequacy of the procedures. Bishop et al. (1994), Beach et al. (1997) and Dardanoni and Forcina (1997) tested inequality using the CPS data 2 or more years apart so that the samples are independent. None of these approaches directly addressed the issue of sample dependency or provided a method of correction.

This paper develops a method of correction and extends conventional inference procedures to situations where samples are dependent. We consider the (interpolated) Gini coefficient and the generalized entropy class of inequality indices, which we define in Section 2, but the methods generalize to some other inequality indices as well. In Section 3, we establish the asymptotic distributions of the estimates of changes in inequality indices when the samples are completely dependent (matched pairs). The procedures are asymptotically distribution-free in that the derived standard errors can be consistently estimated without any prior knowledge about the underlying distribution. We then modify the procedures for cases of inequality comparisons involving partially dependent samples. The covariance between two sample estimates of an inequality index is simply a fraction of the covariance between the matched subsamples. We also outline a simple two-step procedure of computation. Section 4 demonstrates the effects of sample dependency on standard errors of inequality estimates through simulation studies using two parametric bivariate income distributions. In this section, we also apply the correction method to both the CPS data (partially dependent) and the PSID data (completely dependent). We summarize and conclude the paper in Section 5.

2. Changes in inequality indices and their estimates

Consider a joint distribution between two variables $x \in (0, \infty)$ and $y \in (0, \infty)$ with a continuous cumulative distribution functions $F(x, y)$. The marginal distributions of x and y are denoted as $H(x)$ and $K(y)$. For convenience, we further assume that functions H and K are strictly monotone and the first two moments of x and y exist and are finite. Thus, for a given population share p $(0 \leqslant p \leqslant 1)$, there exist unique and finite income quantiles $\xi(p)$ and $\varsigma(p)$ such that $H(\xi(p)) = p$ and $K(\varsigma(p)) = p$.

For a given population share p, the Lorenz curve ordinates of $H(x)$ and $K(y)$ are usually defined as (Gastwirth, 1971)

$$\Phi(p) \equiv \frac{1}{\mu_x}\int_0^{\xi(p)} x\,\mathrm{d}H(x) \quad \text{and} \quad \Psi(p) \equiv \frac{1}{\mu_y}\int_0^{\varsigma(p)} y\,\mathrm{d}K(y), \tag{1}$$

where μ_x and μ_y are the mean incomes of x and y, respectively.

The Gini coefficient is one of the most popular measures of income inequality. It is twice the area between the Lorenz curve of the distribution and the diagonal line. Mathematically, the Gini coefficient of x is defined as

$$G_x = \frac{1}{2\mu_x}\int_0^{\infty}\int_0^{\infty} |x_1 - x_2|\,\mathrm{d}H(x_1)\,\mathrm{d}H(x_2). \tag{2}$$

In empirical applications, once the Lorenz curve is constructed, people often interpolate the Gini value from the Lorenz curve instead of calculating the exact

Gini coefficient. Compared with the exact Gini coefficient, interpolated Ginis are easy to estimate. This advantage becomes more evident when we compute standard errors of the Gini estimates.[1]

Suppose the Lorenz curve is characterized by a set of ordinates corresponding to the abscissas $\{p_l \mid l = 1, 2, \ldots, M+1\}$ with $p_{M+1} = 1$. Assuming $0 < p_1 < p_2 < \cdots < p_M < 1$, we have two sets of Lorenz ordinates $\{\Phi_l \mid l = 1, 2, \ldots, M+1\}$ and $\{\Psi_l \mid l = 1, 2, \ldots, M+1\}$ with $\Phi_{M+1} = \Psi_{M+1} = 1$. The Gini coefficient can be interpolated as

$$G_x^{\mathrm{I}} = \sum_{l=1}^{M} (p_l - \Phi_l)(p_{l+1} - p_{l-1}) \tag{3}$$

and the change in the interpolated Gini coefficient is

$$\Delta G^{\mathrm{I}} = G_x^{\mathrm{I}} - G_y^{\mathrm{I}} = \sum_{l=1}^{M} (\Psi_l - \Phi_l)(p_{l+1} - p_{l-1}). \tag{4}$$

ΔG^{I} converges to the change in the exact Gini as M increases.[2,3]

The generalized entropy class contains several familiar decomposable inequality indices, including two Theil's indices ($c = 0$ and 1) and the (one-half) squared coefficient of variation ($c = 2$).[4] The inequality indices of

[1] In this paper we only derive the asymptotic distribution for the interpolated Gini estimate. The asymptotic distribution for the estimate of the exact Gini coefficient with completely dependent samples can be derived, using results from U-statistics theory (Hoeffding, 1948), in a manner similar to Bishop et al. (1997, 1998). The asymptotic distribution of the exact Gini coefficient with partially dependent samples, however, is much more complicated and is left for future research.

[2] We first estimate the area below the Lorenz curve as the sum of one triangle and M trapezoids, then the Gini coefficient is obtained by subtracting twice the area from 1. A graphical illustration of this method can be found in Anand (1983, p. 312).

[3] The spacing of these abscissas may affect the value of the interpolated Gini index. Ideally, given M, one spaces these abscissas so that the difference between the exact Gini index and the interpolated Gini index is minimized. The optimal spacing rule of these abscissas in the interpolation of Gini, however, depends on the underlying income distribution. Gastwirth (1972) showed that the commonly used practice of even-spacing, i.e., $p_{l+1} - p_l = 1/(M+1)$, is optimal only for a uniform distribution. Aghevli and Mehran (1981) further investigated this issue and provided a general necessary condition for optimal spacing. For many parametric distributions, however, Aghevli and Mehran's condition does not yield closed-form solutions and an iterative procedure must be used. In a simulation study (not reported here), we perform this iteration for several plausible parametric income distributions as well as for the U.S. income data. We confirm Gastwirth's conclusion and suggest that, in general, relatively more abscissas should be placed at both the lower and upper ends of the Lorenz curve. Detailed results are available from the authors upon request. We thank a referee for alerting us to this issue and providing related references.

[4] See Cowell (1980) and Shorrocks (1980) for detailed discussions on this class of inequality indices. It is useful to note that a member of the generalized entropy class, T_x^c with $c = 2$, like the Gini coefficient, also has an interesting geometric interpretation: it is twice the area between the second-degree normalized stochastic curve of x and that of perfectly equal distribution (see Formby et al. (1999) for a detailed description).

x are

$$T_x^c = \begin{cases} \frac{1}{c(c-1)}\{\int_0^\infty [\frac{x}{\mu_x}]^c \, \mathrm{d}H(x) - 1\} & \text{if } c \neq 0, 1, \\ \int_0^\infty [\ln \mu_x - \ln x] \, \mathrm{d}H(x) & \text{if } c = 0, \\ \frac{1}{\mu_x}\int_0^\infty x[\ln x - \ln \mu_x] \, \mathrm{d}H(x) & \text{if } c = 1, \end{cases} \tag{5}$$

and the inequality indices of y, T_y^c, can be similarly defined. The change in the generalized entropy indices is

$$\Delta T^c = T_x^c - T_y^c. \tag{6}$$

Assume two simple random samples of sizes m and n; $(x_1, x_2, \ldots, x_m)$ and $(y_1, y_2, \ldots, y_n)$, are drawn from populations $H(x)$ and $K(y)$, respectively. To allow for dependency between these two samples, we further assume that parts of these samples are drawn together from the joint c.d.f. $F(x, y)$. If the samples are drawn together from $F(x, y)$ entirely, then they are completely dependent (matched pairs) and $m = n$; otherwise samples are partially dependent.

For a given population proportion p, the corresponding Lorenz ordinates of x and y can be consistently estimated as (Goldie, 1977)

$$\hat{\Phi}(p) = \frac{1}{m\bar{x}} \sum_{i=1}^{r_x(p)} x_{(i)} \quad \text{and} \quad \hat{\Psi}(p) = \frac{1}{n\bar{y}} \sum_{i=1}^{r_y(p)} y_{(i)}, \tag{7}$$

where $\bar{x}$ and $\bar{y}$ are sample means of x and y, $x_{(i)}$ and $y_{(i)}$ are the ith order statistics of $\{x_i\}$ and $\{y_i\}$, $r_x(p) = [mp]$ and $r_y(p) = [np]$. If the empirical Lorenz curves of x and y are characterized by ordinates corresponding to the M abscissas $\{p_l \mid l = 1, 2, \ldots, M\}$, then the interpolated Gini coefficients can be consistently estimated as[5]

$$\hat{G}_x^{\mathrm{I}} = \sum_{l=1}^{M} (p_l - \hat{\Phi}_l)(p_{l+1} - p_{l-1}) \tag{8}$$

and the estimated change in the interpolated Gini coefficient is

$$\Delta \hat{G}^{\mathrm{I}} = \hat{G}_x^{\mathrm{I}} - \hat{G}_y^{\mathrm{I}} = \sum_{l=1}^{M} (\hat{\Psi}_l - \hat{\Phi}_l)(p_{l+1} - p_{l-1}), \tag{9}$$

where $\hat{\Phi}_l = \hat{\Phi}(p_l)$ and $\hat{\Psi}_l = \hat{\Psi}(p_l)$.

[5] The exact Gini coefficient can be consistently estimated using formulae such as those provided by Hoeffding (1948), Kendall and Stuart (1977) or Sen (1973). Hoeffding defined the Gini estimate as $(1/2n(n-1)\bar{x})\sum_{i \neq j}|x_i - x_j|$ while both Kendall and Stuart and Sen defined it as $(1/2n^2\bar{x}) \sum_{i \neq j}|x_i - x_j|$. If $M = n - 1$ and all abscissas are evenly spaced, i.e., $p_{l+1} - p_l = 1/n$, the interpolated Gini equals the exact Gini as defined by Hoeffding (see Anand, 1983, p. 313).

Similarly, we can consistently estimate the change in generalized entropy indices as

$$\Delta\hat{T}^c = \Delta\hat{T}^c_x - \Delta\hat{T}^c_y$$
$$= \begin{cases} \frac{1}{c(c-1)}\left\{\frac{1}{n\bar{x}^c}\sum_{i=1}^n (x_i)^c - \frac{1}{n\bar{y}^c}\sum_{i=1}^n (y_i)^c\right\} & \text{if } c \neq 0,\ 1, \\ \ln\bar{x} - \ln\bar{y} - \left\{\frac{1}{n}\sum_{i=1}^n \ln x_i - \frac{1}{n}\sum_{i=1}^n \ln y_i\right\} & \text{if } c = 0, \\ \frac{1}{n\bar{x}}\sum_{i=1}^n x_i \ln x_i - \frac{1}{n\bar{y}}\sum_{i=1}^n y_i \ln y_i - \{\ln\bar{x} - \ln\bar{y}\} & \text{if } c = 1. \end{cases} \tag{10}$$

3. Statistical inference with dependent samples

In this section, we derive large sample properties of $\Delta\hat{G}^{\mathrm{I}}$ and $\Delta\hat{T}^c$, then provide a method of correction for sample dependency. For ease of presentation, we first consider the case of completely dependent samples. We then generalize the results to samples that are partially dependent.

3.1. Completely dependent samples

Assume a random (matched-pair) sample of size n, $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, is drawn independently from the population $F(x, y)$. Since $\Delta\hat{G}^{\mathrm{I}}$ is a function of vectors $(\hat{\Phi}_1, \hat{\Phi}_2, \ldots, \hat{\Phi}_M)'$ and $(\hat{\Psi}_1, \hat{\Psi}_2, \ldots, \hat{\Psi}_M)'$, as shown in (9), the large sample properties of $\Delta\hat{G}^{\mathrm{I}}$ can be derived from the joint distribution of $(\hat{\Phi}_1, \hat{\Phi}_2, \ldots, \hat{\Phi}_M, \hat{\Psi}_1, \hat{\Psi}_2, \ldots, \hat{\Psi}_M)'$. Further, $\hat{\Phi}_l$ is a function of $(1/n)\sum_{i=1}^{r_l} x_{(i)}$ and $\bar{x}$, and $\hat{\Psi}_l$ is a function of $(1/n)\sum_{i=1}^{r_l} y_{(i)}$ and $\bar{y}$, where $r_l = [np_l]$. Denoting

$$\phi_l = \int_0^{p_l} H^{-1}(t)\,\mathrm{d}t \quad \text{and} \quad \varphi_l = \int_0^{p_l} K^{-1}(t)\,\mathrm{d}t, \tag{11}$$

then $\hat{\phi}_l = (1/n)\sum_{i=1}^{r_l} x_{(i)}$ and $\hat{\varphi}_l = (1/n)\sum_{i=1}^{r_l} y_{(i)}$ are consistent estimators of ϕ_l and φ_l, respectively.

The following lemma establishes the asymptotic distribution of

$$\hat{\beta} = (\hat{\phi}_1, \hat{\phi}_2, \ldots, \hat{\phi}_M, \hat{\phi}_{M+1}, \hat{\varphi}_1, \hat{\varphi}_2, \ldots, \hat{\varphi}_M, \hat{\varphi}_{M+1})' \tag{12}$$

with $\hat{\phi}_{M+1} = \bar{x}$ and $\hat{\varphi}_{M+1} = \bar{y}$. A detailed proof of the lemma is provided in Zheng (1996);[6] it is also given in a more general result of Davidson and Duclos (1997).

[6] Zheng (1996, 1999) also proposed inference procedures for testing rank dominance, Lorenz and generalized Lorenz dominances with dependent samples.

Lemma 1. If $F(x,y)$, $H(x)$ and $K(y)$ are continuous, $H(x)$ and $K(y)$ are strictly monotone, and all elements involved also exist and are finite, then the $2(M+1)$-random vector of $\hat{\beta}$ is asymptotically normal in that $n^{1/2}(\hat{\beta}-\beta)$ has a $2(M+1)$-variate normal distribution with mean zero and covariance matrix

$$\Xi = \begin{bmatrix} \omega_{ls} & \tau_{ls} \\ \tau_{ls} & \upsilon_{ls} \end{bmatrix}, \tag{13}$$

where

$$\omega_{ls} = \int_0^{\xi_l} (x-\xi_l)(x-\xi_s)\,\mathrm{d}H(x) - \int_0^{\xi_l} (x-\xi_l)\,\mathrm{d}H(x) \times \int_0^{\xi_s} (x-\xi_s)\,\mathrm{d}H(x) \quad \textit{for } l \leqslant s, \tag{14}$$

$$\upsilon_{ls} = \int_0^{\varsigma_l} (y-\varsigma_l)(y-\varsigma_s)\,\mathrm{d}K(y) - \int_0^{\varsigma_l} (y-\varsigma_l)\,\mathrm{d}K(y) \int_0^{\varsigma_s} (y-\varsigma_s)\,\mathrm{d}K(y) \quad \textit{for } l \leqslant s \tag{15}$$

and

$$\tau_{ls} = \int_0^{\xi_l}\int_0^{\varsigma_s} (x-\xi_l)(y-\varsigma_s)\,\mathrm{d}F(x,y) - \int_0^{\xi_l} (x-\xi_l)\,\mathrm{d}H(x) \times \int_0^{\varsigma_s} (y-\varsigma_s)\,\mathrm{d}K(y). \tag{16}$$

Here ξ_l and ς_l are, respectively, the income quantiles of $H(x)$ and $K(y)$ corresponding to population proportion p_l, i.e., $\xi_l = \xi(p_l)$ and $\zeta_l = \zeta(p_l)$, $l, s = 1, 2, \ldots, M+1$.

To obtain the asymptotic distribution of $\Delta\hat{G}^{\mathrm{I}}$, we use the well-known delta-method (e.g., Rao, 1965, p. 321) on limiting distributions of differentiable functions of random variables. Recalling that $\Delta\hat{G}^{\mathrm{I}} = \sum_{l=1}^{M}(\hat{\Psi}_l - \hat{\Phi}_l)(p_{l+1}-p_{l-1}) = \sum_{l=1}^{M}(\hat{\phi}_l/\hat{\phi}_{M+1} - \hat{\varphi}_l/\hat{\varphi}_{M+1})(p_{l+1}-p_{l-1})$, we have the following result:

Theorem 1. Under the conditions of Lemma 1, the change in the interpolated Gini coefficient, $\Delta\hat{G}^{\mathrm{I}}$, is asymptotically normal in that $n^{1/2}(\Delta\hat{G}^{\mathrm{I}} - \Delta G^{\mathrm{I}})$ has a limiting *normal distribution with mean zero and variance* $\pi^2 = \sum_{s=1}^{M} \sum_{l=1}^{M} \pi_{ls}(p_{s+1} - p_{s-1})(p_{l+1} - p_{l-1})$ *with*

$$\begin{aligned}\pi_{ls} = {} & \frac{1}{\mu_x^2}\,\{\omega_{ls} - \Phi_l \omega_{s(M+1)} - \Phi_s \omega_{l(M+1)} + \Phi_l \Phi_s \sigma_x^2\} \\ & + \frac{1}{\mu_y^2}\{\upsilon_{ls} - \Psi_l \upsilon_{s(M+1)} - \Psi_s \upsilon_{l(M+1)} + \Psi_l \Psi_s \sigma_y^2\} \\ & - \frac{1}{\mu_x \mu_y}\,\{2\tau_{ls} - \Phi_l \tau_{(M+1)s} - \Phi_s \tau_{(M+1)l} - \Psi_l \tau_{s(M+1)} \\ & - \Psi_s \tau_{l(M+1)} + (\Phi_l \Psi_s + \Phi_s \Psi_l)\delta_{xy}\}. \end{aligned} \tag{17}$$

By replacing all π_{ls} in π^2 with each of the three parts of the right-hand side of (17), we obtain the asymptotic variances of $n^{1/2}\hat{G}_x^{\mathrm{I}}$ and $n^{1/2}\hat{G}_y^{\mathrm{I}}$ and (twice) the asymptotic covariance between $n^{1/2}\hat{G}_x^{\mathrm{I}}$ and $n^{1/2}\hat{G}_y^{\mathrm{I}}$, respectively.

Compared with the variance of the exact Gini (as given in Hoeffding (1948), Gastwirth and Gail (1985) and Cowell (1989)), the variance of the interpolated Gini is much easier to estimate. While the estimation of the variance of the exact Gini involves computation of double and triple summations, each item of π^2 can be easily estimated. For example, τ_{ls} can be directly estimated as follows:

$$\begin{aligned}\hat{\tau}_{ls} = {} & \frac{1}{n} \sum_{i=1}^{n} (x_i - x_{(r_l)})(y_i - y_{(r_s)}) I\{(x_i, y_i) \leqslant (x_{(r_l)}, y_{(r_s)})\} \\ & - \frac{1}{n} \sum_{i=1}^{r_l} (x_{(i)} - x_{(r_l)}) \frac{1}{n} \sum_{i=1}^{r_s} (y_{(i)} - y_{(r_s)}), \end{aligned} \tag{18}$$

where $x_{(r_l)}$ and $y_{(r_s)}$ are sample quantiles of $\{x_i\}$ and $\{y_i\}$ corresponding to p_l and p_s, $I(a \leqslant b)$ is an indicator variable which equals 1 if $a \leqslant b$ and zero otherwise, and $(x_i, y_i) \leqslant (x_{(r_l)}, y_{(r_s)})$ stands for the condition that $x_i \leqslant x_{(r_l)}$ and $y_i \leqslant y_{(r_s)}$ hold simultaneously. Similarly, other elements of π^2 can be consistently estimated and, hence, by Slutsky's theorem (e.g., Serfling, 1980, p. 19, Theorem 1.5.4), π^2 can be consistently estimated.

Denoting $N_x^c = (1/n)\sum_{i=1}^{n}(x_i)^c$, $N_y^c = (1/n)\sum_{i=1}^{n}(y_i)^c$ (hence $N_x^1 = \bar{x}$ and $N_y^1 = \bar{y}$), $P_x = (1/n)\sum_{i=1}^{n} \ln x_i$, $P_y = (1/n)\sum_{i=1}^{n} \ln y_i$, $Q_x = (1/n)\sum_{i=1}^{n} x_i \ln x_i$, and $Q_y = (1/n)\sum_{i=1}^{n} y_i \ln y_i$, the estimators of the changes in the generalized entropy indices given in (10) can be expressed as functions of N_x^c, N_y^c, P_x, P_y, Q_x and Q_y. N_x^c, N_y^c, P_x, P_y, Q_x and Q_y are consistent estimators of $\mu_x^c = \int_0^\infty x^c \,\mathrm{d}H(x)$, $\mu_y^c = \int_0^\infty y^c \,\mathrm{d}K(y)$, $\gamma_x = \int_0^\infty \ln x \,\mathrm{d}H(x)$, $\gamma_y = \int_0^\infty \ln y \,\mathrm{d}K(y)$, $\lambda_x = \int_0^\infty x \ln x \,\mathrm{d}H(x)$, and $\lambda_y = \int_0^\infty y \ln y \,\mathrm{d}K(y)$, respectively.

Since $(x_1, x_2, \ldots, x_n)$ are independently and identically distributed, so are their monotonic transformations and hence N_x^c, P_x and Q_x have limiting normal distributions as $n \to \infty$. Similarly, N_y^c, P_y and Q_y are also asymptotically normally distributed. The Cramér–Wold theorem further ensures that

N_x^c, N_y^c, P_x, P_y, Q_x and Q_y have a joint limiting normal distribution. The following results on the large sample properties of $\Delta\hat{T}^c$ can be easily verified:

Theorem 2. Under the assumptions of Lemma 1, $\Delta\hat{T}^c$ has a limiting normal distribution in that $n^{1/2}(\Delta\hat{T}^c - \Delta T^c)$ is asymptotically normally distributed with mean zero and variance

$$\psi^2 = \psi_x^2 + \psi_y^2 - 2\varepsilon_{xy}^c, \tag{19}$$

where ψ_x^2 and ψ_y^2 are asymptotic variances of $n^{1/2}(\hat{T}_x^c - T_x^c)$ and $n^{1/2}(\hat{T}_y^c - T_y^c)$, respectively, as given in Cowell (1989). *The asymptotic covariance term ε_{xy}^c is*

$$\varepsilon_{xy}^c = \begin{cases} \frac{1}{c^2(c-1)^2(\mu_x\mu_y)^{c+1}}\{c^2\eta(\mu_x, \mu_y)\mu_x^c\mu_y^c - c\eta(\mu_x, \mu_y^c)\mu_x^c\mu_y - c\eta(\mu_x^c, \mu_y)\mu_x\mu_y^c \\ \quad + \eta(\mu_x^c, \mu_y^c)\mu_x^c\mu_y^c\} \quad \textit{if } c \neq 0, 1, \\ \frac{1}{\mu_x\mu_y}\{\eta(\mu_x, \mu_y) - \eta(\mu_x, \gamma_y)\mu_y - \eta(\gamma_x, \mu_y)\mu_x \\ \quad + \eta(\gamma_x, \gamma_y)\mu_x\mu_y\} \quad \textit{if } c = 0, \\ \frac{1}{\mu_x\mu_y}\{(\frac{\lambda_x}{\mu_x} + 1)(\frac{\lambda_y}{\mu_y} + 1)\eta(\mu_x, \mu_y) - (\frac{\lambda_x}{\mu_x} + 1)\eta(\mu_x, \lambda_y) \\ \quad - (\frac{\lambda_y}{\mu_y} + 1)\eta(\lambda_x, \mu_y) + \eta(\lambda_x, \lambda_y) \quad \textit{if } c = 1, \end{cases} \tag{20}$$

where $\eta(\cdot,\cdot)$ denote the asymptotic covariances and can be directly derived. For example, the asymptotic covariance between $n^{1/2}(N_x^c - \mu_x^c)$ and $n^{1/2}(N_y^c - \mu_y^c)$ is

$$\eta(\mu_x^c, \mu_y^c) = \int_0^\infty \int_0^\infty x^c y^c \, \mathrm{d}F(x, y) - \mu_x^c\mu_y^c. \tag{21}$$

All elements of ε_{xy}^c can be consistently estimated and hence ε_{xy}^c and ψ^2 can also be consistently estimated.

3.2. Partially dependent samples

Assume two samples of sizes m and n, $\{x_i\}$ and $\{y_j\}$, are drawn from two adjacent years' income distributions with means μ_x and μ_y and variances σ_x^2 and σ_y^2. Further assume that the first q ($q \leqslant min\{m, n\}$) observations of the two samples are matched, i.e., $(x_1, \ldots, x_q)$ and $(y_1, \ldots, y_q)$ are paired, and $(x_{q+1}, \ldots, x_m)$ is independent of $\{y_j\}$ and $(y_{q+1}, \ldots, y_n)$ is independent of $\{x_i\}$. Generally speaking, $(x_{q+1}, \ldots, x_m)$ is not independent of $(x_1, \ldots, x_q)$ and $(y_{q+1}, \ldots, y_n)$ is not independent of $(y_1, \ldots, y_q)$. Thus, $Var(\hat{T}_x^0)$ may not equal ψ_x^2/m and $Var(\hat{T}_y^0)$ may not equal ψ_y^2/n, where ψ_x^2 and ψ_y^2 are given in (19). In the absence of precise information on the nature of this dependency, however, it may not be unreasonable to assume that $Var(\hat{T}_x^0) = \psi_x^2/m$ and $Var(\hat{T}_y^0) = \psi_y^2/n$. If $(x_1, \ldots, x_q)$ is randomly selected from population x and $(x_{q+1}, \ldots, x_m)$ is randomly selected from the same population less $(x_1, \ldots, x_q)$ then the dependency between them may be negligible if q is very small compared to the population size. This assumption amounts to saying that the partial dependency between $\{x_i\}$ and $\{y_j\}$ does not affect the calculation of the variances of $\hat{T}_x^0$ and $\hat{T}_y^0$. Since $Var(\hat{T}_x^0 - \hat{T}_y^0) = Var(\hat{T}_x^0) + Var(\hat{T}_y^0) - 2Cov(\hat{T}_x^0, \hat{T}_y^0)$, we only need to consider the covariance term $Cov(\hat{T}_x^0, \hat{T}_y^0)$.

Denoting $\hat{\alpha}_x = (1/m)\sum_{i=1}^{q} \ln x_i$, $\hat{\rho}_x = (1/m)\sum_{i=q+1}^{m} \ln x_i$, $\hat{\alpha}_y = (1/n)\sum_{j=1}^{q} \ln y_j$, and $\hat{\rho}_y = (1/n)\sum_{j=q+1}^{n} \ln y_j$, we can write $Cov(\hat{T}_x^0, \hat{T}_y^0)$ as

$$\begin{aligned} Cov(\hat{T}_x^0, \hat{T}_y^0) &= Cov(\hat{\alpha}_x + \hat{\rho}_x, \hat{\alpha}_y + \hat{\rho}_y) \\ &= Cov(\hat{\alpha}_x, \hat{\alpha}_y) + Cov(\hat{\alpha}_x, \hat{\rho}_y) + Cov(\hat{\rho}_x, \hat{\alpha}_y) + Cov(\hat{\rho}_x, \hat{\rho}_y). \end{aligned} \tag{22}$$

Since $\hat{\rho}_x$ is independent of $\{y_j\}$ (hence $\hat{\alpha}_y$ and $\hat{\rho}_y$) and $\hat{\rho}_y$ is independent of $\{x_i\}$ (hence $\hat{\alpha}_x$ and $\hat{\rho}_x$) by assumption, we have $Cov(\hat{\alpha}_x, \hat{\rho}_y) = Cov(\hat{\rho}_x, \hat{\alpha}_y) = Cov(\hat{\rho}_x, \hat{\rho}_y) = 0$ and thus

$$Cov(\hat{T}_x^0, \hat{T}_y^0) = Cov(\hat{\alpha}_x, \hat{\alpha}_y). \tag{23}$$

Noting that $\hat{\alpha}_x = (1/m)\sum_{i=1}^{q} \ln x_i = (q/m)[\frac{1}{q}\sum_{i=1}^{q} \ln x_i]$ and $\hat{\alpha}_y = (1/n)\sum_{j=1}^{q} \ln y_j = (q/n)[\frac{1}{q}\sum_{j=1}^{q} \ln y_j]$, we further have

$$Cov(\hat{T}_x^0, \hat{T}_y^0) = \frac{q}{m} \times \frac{q}{n} \times Cov\left(\frac{1}{q}\sum_{i=1}^{q} \ln x_i, \frac{1}{q}\sum_{j=1}^{q} \ln y_j\right). \tag{24}$$

$Cov((1/q)\sum_{i=1}^{q} \ln x_i, (1/q)\sum_{j=1}^{q} \ln y_j)$ can be directly calculated using the methods developed in Section 3.1. Thus, $Cov(\hat{T}_x^0, \hat{T}_y^0)$ can be computed using the following two-step procedure: *first calculate the covariance of the sample statistics of the matched sub-samples as if they were complete samples*, i.e., $Cov((1/q)\sum_{i=1}^{q} \ln x_i, (1/q)\sum_{j=1}^{q} \ln y_j)$; *then multiply the covariance by the percentages of the matched portions of the two samples* q/m *and* q/n), i.e., $q/m \times q/n \times Cov((1/q)\sum_{i=1}^{q} \ln x_i, (1/q)\sum_{j=1}^{q} \ln y_j)$. The generalization to other members of the generalized entropy class is straightforward.

The computation of the covariance term of the interpolated Gini index, $Cov(\hat{G}_x^{\mathrm{I}}, \hat{G}_y^{\mathrm{I}})$, is also straightforward. As in computing $Cov(\hat{T}_x^0, \hat{T}_y^0)$, we have to assume that the sample dependency does not affect the estimation of ω_{ls} and υ_{ls}. Under the same assumptions about the data structures of $\{x_i\}$ and $\{y_j\}$ as described above, the estimate of τ_{ls} of (16) is

$$\begin{aligned} \hat{\tau}_{ls} = \frac{q}{m} \times \frac{q}{n} \Bigg\{ & \frac{1}{q}\sum_{i=1}^{q} (x_i - x_{(r_l)})(y_i - y_{(r_s)}) I\{(x_i, y_i) \leqslant (x_{(r_l)}, y_{(r_s)})\} \\ & - \frac{1}{q}\sum_{i=1}^{\tilde{r}_l} (\tilde{x}_{(i)} - x_{(r_l)}) \frac{1}{q}\sum_{j=1}^{\tilde{r}_s} (\tilde{y}_{(j)} - y_{(r_s)}) \Bigg\}, \end{aligned} \tag{25}$$

where $\tilde{r}_l = [qp_l]$ and $\tilde{r}_s = [qp_s]$, and $\tilde{x}_{(i)}$ and $\tilde{y}_{(i)}$ are the ith order statistics of $\{x_1, \dots, x_q\}$ and $\{y_1, \dots, y_q\}$, respectively. Substitute this result into (17), the covariance term is the covariance between the matched samples multiplied by proportions q/m and q/n.

4. Sample dependency: Evidences from simulations and the U.S. data

The previous section proposed a method to take sample dependency into account in computing the standard errors of estimated inequality differences, but how serious is the problem of sample dependency in empirical applications remains unanswered. In this section, we answer this question using both simulation studies and the U.S. income and earnings data. Monte Carlo simulations are used to show the effect of sample dependency on standard errors of inequality changes while the U.S. data are used to demonstrate the effect of sample dependency on statistical inference.

4.1. Monte Carlo simulations

Since sample dependency may vary by the degree of correlation and the scope of overlapping between samples, it is necessary to evaluate the effect of sample dependency with different degrees of correlation and different scopes of overlapping. Given that income variables are likely to be positively correlated between consecutive years, we consider in this study five different degrees of correlation (0.1, 0.3, 0.5, 0.7 and 0.9) and five different scopes of overlapping (0.1, 0.3, 0.5, 0.7 and 0.9).[7] The parametric income distributions we use in this study are the Singh–Maddala distribution and the lognormal distribution. The c.d.f. of the univariate Singh – Maddala (1976) distribution is

$$H(x) = 1 - 1/[1 + (x/b)^a]^q \quad \text{with } a \geqslant 0,\ b > 0 \text{ and } q > 1/a. \tag{26}$$

We consider the Singh–Maddala distribution and the lognormal distribution because both distributions have been described as good approximations to actual income distributions (e.g., see McDonald, 1984), on the Singh–Maddala distribution and Aitchison and Brown, 1969, on the lognormal distribution). The parameters used in the Singh–Maddala distribution are those estimated by McDonald (1984) for the 1980 U.S. income distribution, i.e., $a = 1.6971$ and $q = 8.3679$. We set $b = 1$ since it is a scale variable. The lognormal distribution considered has mean 10,000 and standard deviation 5000.

[7] Negatively correlated bivariate random samples are much more difficult to generate than positively correlated samples. Fortunately, the postulation of positive correlation between income variables is supported in our empirical studies, even for PSID samples with a long time span.

The bivariate c.d.f. of these distributions could be quite complicated to describe, but dependent bivariate random samples can be easily generated. For ease of generating samples, we assume that random variables, x and y, have identical marginal distributions. This allows us to focus on the effect of varying correlation and varying scope of overlapping on the standard errors of $\Delta\hat{G}^{\mathrm{I}}$ and $\Delta\hat{T}^{c}$; the effect on statistical inference is demonstrated below using the U.S. data. The Singh–Maddala bivariate random samples are obtained by first generating bivariate uniform random numbers and then converting them using Johnson's transformation system (Johnson, 1987); the bivariate uniform random samples are generated using the method proposed by Lawrance and Lewis (1983). The bivariate lognormal random samples are obtained by first generating bivariate standard normal random numbers using the Box–Muller method (see, e.g., Lewis and Orav, 1989) and then converting bivariate standard normal random numbers into lognormal random numbers. Both the Lawrance–Lewis method and the Box–Muller method provide an easy way to control the degree of correlation between samples. The scope of overlapping is controlled by first drawing the required proportion of the samples from the joint distribution and then drawing the remaining samples from each marginal distribution independently. From each bivariate distribution (with each degree of correlation and each scope of overlapping), we extract 100 samples of size 5000 and compute separately the standard errors of $\Delta\hat{G}^{\mathrm{I}}$ and $\Delta\hat{T}^{c}$ – with and without correcting for sample dependency. The impact of sample dependency is of course reflected in the differences between these two sets of standard errors.

Table 1 reports the (average) simulation results of the 100 runs for the Singh–Maddala and the lognormal distributions, respectively. For each distribution, we conduct simulations for five inequality indices (the interpolated Gini and four generalized entropy measures corresponding to $c = 0$, 0.5, 1 and 2). The number at the intersection of each correlation level and each overlapping level indicates the average percentage increase in the standard error of $\Delta\hat{G}^{\mathrm{I}}$ or $\Delta\hat{T}^{c}$ if sample dependency is ignored. That is, each number is the average of the percentage differences, [(uncorrected standard error − corrected standard error)/corrected standard error] $\times$ 100, from the 100 runs. For example, the number 88.1 (at the intersection of 0.9 coefficient of correlation and 0.9 overlapping for the interpolated Gini in Table 1a) means that, on average, the standard error would go up by 88.1% if sample dependency is not taken into account.

An inspection of the tables suggests that sample dependency has substantial effects on standard errors if samples are strongly correlated and/or have a significant overlap. Since the test statistic of a difference such as $\Delta\hat{G}^{\mathrm{I}}$ is inversely related to the standard error of the difference, it follows that the correction for sample dependency may change the p-value of the difference. In the above example, failure to correct for sample dependency would lower the absolute value of the test statistic by 47%! Hence, technically speaking, the correction for sample dependency will undoubtedly affect statistical inference in the sense that

Table 1
Simulation studies: increases in standard errors if sample dependency is not corrected

Inequality measures	Correlation	The proportion of overlapping				
		0.1	0.3	0.5	0.7	0.9
(a) Singh–Maddala distribution[a]						
	0.1	0.0	0.1	0.4	0.9	1.3
Interpolated Gini	0.3	0.1	0.7	2.4	4.9	8.3
	0.5	0.2	1.7	5.3	11.4	20.9
	0.7	0.3	2.8	8.7	20.4	42.6
	0.9	0.4	4.2	13.3	33.3	88.1
	0.1	0.1	0.7	1.7	3.3	5.2
Generalized entropy	0.3	0.1	1.5	4.2	9.1	16.8
($c = 0$)	0.5	0.2	2.6	6.8	15.9	32.6
	0.7	0.3	3.6	10.2	24.8	58.0
	0.9	0.5	4.6	14.4	35.6	105.5
	0.1	0.0	0.2	0.9	1.8	3.5
Generalized entropy	0.3	0.2	1.2	3.4	7.2	13.4
($c = 0.5$)	0.5	0.3	2.1	6.7	13.8	27.0
	0.7	0.4	3.1	10.2	22.5	48.8
	0.9	0.5	4.2	13.5	33.5	91.5
	0.1	0.0	0.1	0.6	1.1	2.0
Generalized entropy	0.3	0.1	0.9	2.5	5.2	9.5
($c = 1$)	0.5	0.2	1.8	5.3	11.1	21.2
	0.7	0.4	2.9	9.0	19.9	41.8
	0.9	0.5	4.2	13.1	32.2	85.7
	0.1	0.0	0.0	0.2	0.4	0.7
Generalized entropy	0.3	0.1	0.5	1.3	2.8	4.9
($c = 2$)	0.5	0.1	1.2	3.4	7.2	13.0
	0.7	0.3	2.3	7.0	15.2	30.2
	0.9	0.4	3.9	12.3	29.6	74.8
(b) Lognormal distribution[b]						
	0.1	0.0	0.1	0.2	0.3	0.5
Interpolated Gini	0.3	0.1	0.7	1.7	3.2	5.2
	0.5	0.2	1.6	4.2	8.4	15.0
	0.7	0.3	2.9	8.3	17.9	35.4
	0.9	0.4	4.4	13.2	30.9	75.9
	0.1	0.0	0.2	0.4	0.6	1.1
Generalized entropy	0.3	0.1	0.8	2.3	4.2	7.2
($c = 0$)	0.5	0.2	1.8	5.3	10.4	19.1
	0.7	0.3	3.0	9.1	19.1	39.2
	0.9	0.4	4.3	13.5	31.2	79.1

Table 1 (*continued*)

Inequality measures	Correlation	The proportion of overlapping				
		0.1	0.3	0.5	0.7	0.9
	0.1	0.1	0.1	0.0	0.3	0.8
Generalized entropy	0.3	0.1	0.7	1.6	3.4	6.1
($c = 0.5$)	0.5	0.2	1.6	4.4	9.1	16.7
	0.7	0.3	2.8	8.0	17.4	35.1
	0.9	0.4	4.2	12.6	29.7	73.4
	0.1	0.0	0.1	0.1	0.3	0.7
Generalized entropy	0.3	0.1	0.6	1.5	3.0	5.3
($c = 1$)	0.5	0.1	1.6	4.1	8.1	14.7
	0.7	0.3	2.7	7.7	16.0	31.4
	0.9	0.4	4.2	12.6	28.7	68.7
	0.1	0.0	0.1	0.1	0.2	0.4
Generalized entropy	0.3	0.0	0.5	1.1	2.3	4.2
($c = 2$)	0.5	0.1	1.2	3.1	6.1	10.9
	0.7	0.2	2.4	6.5	12.8	24.2
	0.9	0.3	4.1	11.8	25.7	60.0

[a] The two dependent income distributions are the Singh–Maddala distributions $SM(1.6971, 8.3679)$. The pseudo-random numbers are generated using the routine provided in Microsoft FORTRAN90. The sample size of each simulation is 5000 and each simulation is repeated 100 times. All increases are represented in percentage and are calculated as the average of the differences, [(uncorrected standard error − corrected standard error)/corrected standard error] × 100, from the 100 runs.

[b] The two dependent income distributions are lognormal distributions with mean 10 000 and standard deviation 5000. The pseudo-random numbers are generated using the routine provided in Microsoft FORTRAN90. The sample size of each simulation is 5000 and each simulation is repeated 100 times. All increases are represented in percentage and are calculated as the average of the differences, [(uncorrected standard error − corrected standard error)/corrected standard error] × 100, from the 100 runs.

the p-value will be changed. Of course, if only a few significance levels such as 0.01, 0.05 and 0.10 are used, as we usually do in empirical studies, it is possible that the correction for sample dependency may not change the significance level of a given difference.

4.2. Applications to the CPS and PSID data

In this section, we test changes in U.S. family income and personal earnings inequality using the current population survey (CPS) data and the panel study of income dynamics (PSID) data, respectively. The CPS data is an example of partially dependent samples and the PSID data is an example of completely dependent samples. The CPS sample rotates every 2 years and, by design, about half of the households are overlapping (see the CPS website at http://www.bls.census.gov/cps). Considering, however, that some people may move out of their dwellings every year and that some may not respond to the survey, the actual overlap is closer to 40%. The PSID is longitudinal and all individuals are re-interviewed every year and thus the proportion of overlapping in theory is 100%.

Table 2 reports results on annual U.S. family income inequality and its changes from 1990 to 1997. We include in our sample all primary families and subfamilies but exclude all people living in group quarters; we also drop all entries with zero or negative family incomes in order to use the generalized entropy class of inequality. The top half of the table provides estimates of inequality indices (interpolated Gini index and four generalized entropy measures) and the associated standard errors (inside parentheses). The bottom half documents the inequality comparisons of consecutive years. In each cell there are four numbers: the change in the inequality index; the standard error of the inequality change *without* correcting for sample dependency (inside parentheses); the standard error *with* correction for sample dependency (inside brackets); the percentage increase in the standard error if sample dependency is not corrected (inside braces). Consider, for example, the comparison of the Gini coefficient between 1990 and 1991. The first number (0.0032) is the difference between 1991's Gini index and 1990's Gini index, the second number (0.0020) is the standard error of the Gini difference without correcting for sample dependency, the third number [0.0018] is the standard error of the Gini difference with correction for sample dependency, the fourth number {11.2%} reflects the percentage increase in the standard error (from 0.0018 to 0.0020) if sample dependency is not corrected.

The table indicates that failure to correct for sample dependency may increase the standard error by between 3.3% and 17.1%. A difference of this size may affect the significance levels of comparisons even if only three significance levels (i.e., 0.01, 0.05 and 0.10) are used. The effect of correcting sample dependency on statistical inference is indicated by the two superscribed symbols in the top line of each cell: if two symbols are different then the correction changes the significance level. All test statistics are compared with conventional critical values (e.g., 1.96 for the 5% level). The table indicates that not all corrections for sample dependency change the significance levels but in several cases the correction does make a difference. For example, the difference 0.0032 between the 1990 Gini index and the 1991 Gini index is not significant at the 10% level with an uncorrected standard error but becomes significant with a corrected standard error. Interestingly, the magnitude of correction does not have to be large to make a difference and a large magnitude of correction does not necessarily change the significance level. For example, the small 3.3% difference

Table 2
U.S. family income inequality: 1990–1997 (CPS data)[a]

	Int. Gini	GE ($c = 0$)	GE ($c = 0.5$)	GE ($c = 1$)	GE ($c = 2$)
1990	0.3771	0.3261	0.2597	0.2455	0.2739
	(0.0014)	(0.0036)	(0.0018)	(0.0018)	(0.0036)
1991	0.3803	0.3397	0.2657	0.2505	0.2816
	(0.0014	(0.0039)	(0.0019)	(0.0019)	(0.0039)
1992	0.3818	0.3401	0.2676	0.2519	0.2813
	(0.0014)	(0.0038)	(0.0019)	(0.0019)	(0.0038)
1993	0.3884	0.3587	0.2765	0.2604	0.2937
	(0.0015)	(0.0042)	(0.0020)	(0.0020)	(0.0042)
1994	0.3871	0.3627	0.2747	0.2578	0.2867
	(0.0015)	(0.0044)	(0.0019)	(0.0019)	(0.0044)
1995	0.4104	0.3858	0.3164	0.3185	0.4494
	(0.0020)	(0.0049)	(0.0032)	(0.0039)	(0.0049)
1996	0.4146	0.4015	0.3250	0.3282	0.4700
	(0.0021)	(0.0052)	(0.0033)	(0.0039)	(0.0052)
1997	0.4189	0.4159	0.3317	0.3349	0.4809
	(0.0021)	(0.0054)	(0.0033)	(0.0039)	(0.0054)
1990/91	0.0032[b,c]	0.0136[d,e]	0.0060[d,d]	0.0051[c,d]	0.0077[b,b]
	(0.0020)	(0.0053)	(0.0026)	(0.0026)	(0.0053)
	[0.0018]	[0.0051]	[0.0024]	[0.0024]	[0.0050]
	{11.2%}	{3.4%}	{10.5%}	{10.5%}	{4.9%}
1991/92	0.0015[b,b]	0.0004[b,b]	0.0020[b,b]	0.0014[b,b]	− 0.0003[b,b]
	(0.0020)	(0.0054)	(0.0027)	(0.0027)	(0.0054)
	[0.0018]	[0.0052]	[0.0024]	[0.0024]	[0.0049]
	{12.2%}	{3.7%}	{12.0%}	{13.3%}	{9.5%}
1992/93	0.0066[e,e]	0.0186[e,e]	0.0089[e,e]	0.0085[e,e]	0.0124[d,d]
	(0.0021)	(0.0057)	(0.0027)	(0.0028)	(0.0057)
	[0.0018]	[0.0055]	[0.0025]	[0.0025]	[0.0054]
	{11.2%}	{3.3%}	{10.4%}	{10.3%}	{5.2%}
1995/96	0.0042[b,b]	0.0157[d,d]	0.0086[c,d]	0.0097[c,c]	0.0206[e,e]
	(0.0029)	(0.0071)	(0.0046)	(0.0055)	(0.0071)
	[0.0027]	[0.0069]	[0.0042]	[0.0051]	[0.0061]
	{9.0%}	{3.5%}	{8.4%}	{8.0%}	{17.1%}
1996/97	0.0043[b,b]	0.0144[c,d]	0.0068[b,b]	0.0068[b,b]	0.0109[b,b]
	(0.0029)	(0.0075)	(0.0046)	(0.0056)	(0.0075)
	[0.0027]	[0.0073]	[0.0043]	[0.0052]	[0.0067]
	{9.2%}	{3.3%}	{8.3%}	{7.8%}	{12.5%}

[a]The top half of the table shows the inequality index and the associated standard error (in parentheses). The bottom half shows the difference between 2 years' inequality indices; the number in ( ) is the uncorrected standard error of this difference (in parentheses); the corrected standard error (in brackets); the percentage difference of the correction in the standard error [(uncorrected − corrected)/corrected] × 100 (in braces). The first symbol on the shoulder of each difference indicates significance with the uncorrected standard error while the second symbol indicates significance with the corrected standard error.

[b]Indicates insignificance at the 10% level.

[c]Indicates significance at the 10% level.

[d]Indicates significance at the 5% level.

[e]Indicates significance at the 1% level.

Table 3
U.S. personal earnings inequality: 1985–1990 (PSID data)[a]

	Int. Gini	GE ($c = 0$)	GE ($c = 0.5$)	GE ($c = 1$)	GE ($c = 2$)
1985/86	– 0.0106[b,c]	– 0.0293[d,c]	– 0.0202[b,c]	– 0.0177[e,d]	– 0.0139[e,e]
	(0.0063)	(0.0123)	(0.0121)	(0.0181)	(0.0945)
	[0.0038]	[0.0066]	[0.0057]	[0.0083]	[0.0456]
	{87.1%}	{86.2%}	{111.6%}	{117.1%}	{107.2%}
1986/87	– 0.0121[b,c]	– 0.0440[c,c]	– 0.0218[e,c]	– 0.0072[e,e]	0.0963[e,e]
	(0.0069)	(0.0134)	(0.0143)	(0.0236)	(0.1622)
	[0.0036]	[0.0069]	[0.0070]	[0.0126]	[0.1146]
	{92.4%}	{95.1%}	{105%}	{87.6%}	{41.6%}
1987/88	– 0.0103[e,c]	– 0.0433[c,c]	– 0.0223[e,c]	– 0.0133[e,d]	0.0026[e,e]
	(0.0074)	(0.0143)	(0.0161)	(0.0278)	(0.2028)
	[0.0032]	[0.0060]	[0.0047]	[0.0062]	[0.0385]
	{130.8%}	{137.9%}	{243.4%}	{346.5%}	{426.7%}
1988/89	– 0.0117[b,c]	– 0.0399[c,c]	– 0.0258[b,c]	– 0.0255[e,c]	– 0.1128[e,e]
	(0.0071)	(0.0139)	(0.0152)	(0.0252)	(0.1679)
	[0.0033]	[0.0063]	[0.0055]	[0.0088]	[0.0749]
	{114.2%}	{122.2%]	{174.3%}	{184.9%}	{124.3%}
1989/90	– 0.0070[e,b]	– 0.0273[d,c]	– 0.0173[e,d]	– 0.0173[e,e]	– 0.0807[e,e]
	(0.0066)	(0.0127)	(0.0129)	(0.0195)	(0.1003)
	[0.0037]	[0.0072]	[0.0071]	[0.0119]	[0.0760]
	{80.0%}	{77.2%}	{80.4%}	{64.7%}	{31.8%}
1985/87	– 0.0179[d,c]	– 0.0579[c,c]	– 0.0324[d,c]	– 0.0170[e,e]	0.0930[e,e]
	(0.0071)	(0.0138)	(0.0146)	(0.0235)	(0.1530)
	[0.0044]	[0.0087]	[0.0091]	[0.0157]	[0.1249]
	{60.7%}	{58.3%}	{60.7%}	{49.4%}	{22.6%}
1985/88	– 0.0210[c,c]	– 0.0707[c,c]	– 0.0386[d,c]	– 0.0191[e,e]	– 0.1109[e,e]
	(0.0074)	(0.0142)	(0.0153)	(0.0247)	(0.1543)
	[0.0047]	[0.0093]	[0.0095]	[0.0159]	[0.1163]
	{56.7%}	{51.5%}	{61.1%}	{55.7%}	{32.6}
1985/89	– 0.0222[c,c]	– 0.0724[c,c]	– 0.0410[c,c]	– 0.0249[e,b]	0.0317[e,e]
	(0.0072)	(0.0138)	(0.0141)	(0.0214)	(0.1067)
	[0.0047]	[0.0096]	[0.0089]	[0.0136]	[0.0782]
	{63.7%}	{44.8%}	{59.0%}	{57.2%}	{36.4%}
1985/90	– 0.0239[c,c]	– 0.0762[c,c]	– 0.0462[c,c]	– 0.0343[b,d]	– 0.0417[e,e]
	(0.0072)	(0.0136)	(0.0134)	(0.0192)	(0.0813)
	[0.0051]	[0.0102]	[0.0097]	[0.0145]	[0.0706]
	{39.8%}	{33.1%}	{37.8%}	{32.1%}	{15.2%}

[a]Each cell shows the difference between 2 years' inequality indices; the uncorrected standard error of this difference (in parentheses); the corrected standard error (in brackets); the percentage difference of the correction in the standard error [(uncorrected − corrected)/corrected] × 100 (in braces). The first symbol on the shoulder of each difference indicates significance with the uncorrected standard error while the second symbol indicates significance with the corrected standard error. The top half of the table provides comparison of adjacent years between 1985 and 1990 while the bottom half compares 1985 with 1987–1990.

[b]Indicates significance at the 10% level.

[c]Indicates significance at the 1% level.

[d]Indicates significance at the 5% level.

[e]Indicates insignificance at the 10% level.

of correction for the generalized entropy index with $c = 0$ for the comparison between 1996 and 1997 changes the significance level; the large 12.5% for the generalized entropy index with $c = 2$ for the same comparison does not change the significance level. This happens because the first change in inequality (0.0144) is close to be significant at the 5% level and only a small reduction in the standard error is sufficient to make it significant, but the second change in inequality (0.0109) is far from being significant at the 10% level and even a large reduction in the standard error is not enough to make it significant. It follows that the correction for sample dependency is particularly important and necessary if the p-value with uncorrected standard error is already in the neighborhood of being significant.

Table 3 compares U.S. personal earnings inequality from 1985 to 1990. The top half of the table reports five comparisons of consecutive years while the bottom half compares 1985 with other years (1987–1990). All numbers carry the same meanings as those in Table 2 (bottom half). In this illustration, we are interested in changes in earnings inequality of those who were employed, thus we limit our samples to those individuals who worked full time and have positive earnings in both years of each comparison. The effect of the correction on standard errors varies widely, ranging from 15.2% (the generalized entropy index with $c = 2$ for the 1985/1990 comparison) to 426.7% (the generalized entropy index with $c = 2$ for the 1987/1988 comparison). These corrections affect the significance level in 15 of the 25 cases in the top half of the table.

5. Summary and conclusion

Researchers often encounter sample dependency when testing inequality changes across time. Commonly used income samples such as the CPS and PSID are dependent in that they have an overlap in consecutive years. Since conventional inference procedures assume independence among samples, a method of correction for sample dependency is needed.

This paper provided appropriate inference procedures to test the (interpolated) Gini coefficient and the generalized entropy class with dependent samples. We first showed that the estimated changes in inequality indices with completely dependent samples have asymptotic normal distributions and that standard errors can be straightforwardly estimated. We then modified the procedure to adjust the standard errors for inequality comparisons with partially dependent samples. The asymptotic covariance of the partially dependent samples can be easily calculated using a two-step procedure.

The effects of sample dependency on standard errors were documented through a series of simulation studies. We considered two reasonable parametric bivariate distributions and allowed both the degree of correlation and the scope of overlapping to vary. The studies showed that sample dependency may have substantial impacts on standard errors. We also empirically investigated this issue by testing changes in U.S. family income inequality from 1990 to 1997 (using the CPS data) and personal earnings inequality from 1985 to 1990 (using the PSID data). Our empirical results further demonstrated the importance of adjusting standard errors for sample dependency.

While we focused on testing cross-time inequality changes, the method of correction for sample dependency is also applicable to testing marginal changes in inequality. Marginal changes in income inequality refer to the increase or decrease in income inequality of the same population after the population has experienced a change in income. An example of marginal changes is the effect of income transfer programs on income distribution. In the United States, a sizable portion of GNP is spent annually on various welfare programs such as food stamps and temporary assistance to needy families (formerly aid to families with dependent children). This has affected the income distribution. While researchers have generally agreed that welfare programs reduce both income inequality and poverty (see Danziger et al., 1981), they have not arrived at a consensus regarding the magnitude of the reduction. With the social welfare system under scrutiny and in the midst of reforms, accurately measuring the impact of welfare reforms on income inequality is important. Other interesting examples of marginal changes in income inequality include the effect of taxation on income inequality and the effect of wives' participation in the labor force on family income inequality. The two samples of before- and after-event needed to address these issues are completely dependent and the method proposed in this paper should be used to correct for sample dependency. As we have demonstrated here, failure to use such a method may lead to erroneous conclusions.

Acknowledgements

We thank three anonymous referees, an Associate Editor, Professor Arnold Zellner, Dr. Fanhui Kong, Dr. John Bishop and Professor Ram Shanmugam for useful comments and suggestions on earlier versions of the paper. We are also grateful to Xuemei (Alice) Du for her assistance in computer programming. The usual caveats apply.

References

Aghevli, B., Mehran, F., 1981. Optimal grouping of income distribution data. Journal of the American Statistical Association 76, 22–26.

Aitchison, J., Brown, J., 1969. The Lognormal Distribution with Special References to Its Uses in Economics. Cambridge University Press, Cambridge.

Anand, S., 1983. Inequality and Poverty in Malaysia: Measurement and Decomposition. Oxford University Press, London.

Beach, C., Chaykowski, R., Slotsve, G., 1997. Inequality and polarization of male earnings in the United States, 1968–1990. North American Journal of Economics and Finance 8, 135–151.
Bishop, J.A., Formby, J.P., Smith, W.J., 1991a. Lorenz dominance and welfare: changes in the US distribution of income, 1967–1986. Review of Economics and Statistics 71, 134–139.
Bishop, J.A., Chow, K., Formby, J.P., 1991b. A stochastic dominance analysis of growth, recessions and the US income distribution, 1967–1986. Southern Economic Journal 57, 936–946.
Bishop, J.A., Chiou, J., Formby, J.P., 1994. Truncation bias and the ordinal evaluation of income inequality. Journal of Business and Economic Statistics 12, 123–127.
Bishop, J.A., Formby, J.P., Zheng, B., 1997. Statistical inference and the Sen index of poverty. International Economic Review 38, 381–387.
Bishop, J.A., Formby, J.P., Zheng, B., 1998. Inference tests for Gini-based tax progressivity indexes. Journal of Business and Economic Statistics 16, 322–330.
Cowell, F., 1980. On the structure of additive inequality measures. Review of Economic Studies 47, 521–531.
Cowell, F., 1989. Sampling variance and decomposable inequality measures. Journal of Econometrics 42, 27–42.
Danziger, S., Haveman, R., Plotnick, R., 1981. How income transfer programs affect work, savings, and the income distribution: a critical review. Journal of Economic Literature 19, 975–1028.
Dardanoni, V., Forcina, A., 1997. Inference for the Lorenz Curve Ordering. Mimeo, University of California at San Diego.
Davidson, R., Duclos, J., 1997. Statistical inference for the measurement of incidence of taxes and transfers. Econometrica 65, 1453–1465.
Formby, J.P., Smith, W.J., Zheng, B., 1999. The coefficient of variation, stochastic dominance and inequality: a new interpretation. Economics Letters 62, 319–323.
Gail, M., Gastwirth, J.L., 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.
Gastwirth, J.L., 1971. A general definition of the Lorenz curve. Econometrica 39, 1037–1039.
Gastwirth, J.L., 1972. The estimation of the Lorenz curve and Gini index. Review of Economics and Statistics 54, 306–316.
Gastwirth, J.L., 1974. Large sample theory of some measures of income inequality. Econometrica 42, 191–196.
Gastwirth, J.L., Gail, M., 1985. Simple asymptotically distribution-free methods for comparing Lorenz curves and Gini indices obtained from complete data. In: Basmann, R.L., Rhodes Jr., G.F. (Eds.), Advances in Econometrics, Vol. 4. JAI Press, Connecticut.
Goldie, C., 1977. Convergence theorems for empirical Lorenz curves and their inverses. Advances in Applied Probability 9, 765–791.
Hoeffding, W., 1948. A class of statistics with asymptotically normal distribution. Annals of Mathematical Statistics 19, 293–325.
Johnson, M., 1987. Multivariate Statistical Inference. McGraw-Hill, New York.
Kendall, M., Stuart, A., 1977. The Advanced Theory of Statistics. Macmillan, New York.
Lawrance, A., Lewis, P., 1983. Simple dependent pairs of exponential and uniform random variables. Operations Research 31, 1179–1197.
Lewis, P., Orav, E., 1989. Simulation Methodology for Statisticians, Operations Analysts, and Engineers. Wadsworth & Brooks, Pacific Grove, CA.
McDonald, J., 1984. Some generalized functions for the size distribution of income. Econometrica 52, 647–663.
Rao, C.R., 1965. Linear Statistical Inference and Its Applications. Wiley, New York.
Sen, A., 1973. On Economic Inequality. Clarendon Press, Oxford.
Serfling, R.J., 1980. Approximation Theorems of Mathematical Statistics. Wiley, New York.
Shorrocks, A., 1980. The class of additively decomposable inequality measures. Econometrica 48, 613–625.

Singh, S., Maddala, G., 1976. A function for size distribution of incomes. Econometrica 44, 963–970.

Schluter, C., 1996. Income distribution and inequality in Germany: evidence from panel data. STICERD Working Paper, London School of Economics.

Thistle, P., 1990. Large sample properties of two inequality indices. Econometrica 58, 725–728.

Zheng, B., 1996. Statistical inferences for testing marginal changes in Lorenz and generalized Lorenz curves. Mimeo, University of Colorado at Denver.

Zheng, B., 1999. Statistical inferences for testing marginal rank and (generalized) Lorenz dominances. Southern Economic Journal 65, 557–570.