Statistical Inference for Interdistributional Lorenz Curves*

John A. Bishop
Department of Economics
East Carolina University

K. Victor Chow
Department of Finance
West Virginia University

Lester A. Zeager
Department of Economics
East Carolina University

Abstract

We provide methods for making stochastic dominance comparisons of interdistributional Lorenz curves (ILCs) for population subgroups. We offer alternative tests for significant differences between a single ILC and the 45-degree line in a given year and between two ILCs across years. The latter comparisons show whether the ILCs are converging to (or diverging from) equality over time. Our tests apply to any moment of the distribution and impose no prior restrictions on the functional form of the underlying distribution. We illustrate the methods with data on white and non-white incomes drawn from the Current Population Survey.

November 5, 2004

* The authors are grateful for valuable comments from John Formby and other session participants at the Eastern Economic Association meetings (Washington, DC) and at the Department of Economics seminar, East Carolina University. We take sole responsibility for any remaining shortcomings of the paper.

Statistical Inference for Interdistributional Lorenz Curves

Interdistributional Lorenz curves (ILC) were devised by Butler and McDonald (1987) as a method for measuring inequality *across* two populations – in contrast to the dispersion of incomes *within* a single population (as in the traditional Lorenz curve), but the essential idea has a longer history. Deutsch and Silber (1999) trace interdistributional inequality to Gini (1916, 1959), who introduced the notions of "Transvariazione" (extent of overlapping) and "Ipertransvariazione" (difference between the concentrations of each distribution below a reference point in the other distribution). Subsequent extensions and refinements have been offered by Dagum (1980, 1987), Shorrocks (1982), Vinod (1985), Gastwirth (1985), and Deutsch and Silber (1987). Deutsch and Silber (1999) provide a useful survey of this literature and they claim that the ILCs of Butler and McDonald (1987) capture the notion of "economic advantage" in the prior literature.

What distinguishes ILCs from the earlier literature is a dominance approach to interdistributional inequality. Yet, in their empirical application, Butler and McDonald (1987) collapse the information provided by the ILC into a single (Pietra) index number before making the comparisons. Regardless of the index selected (Pietra, Gini, etc.), the findings will be less general than dominance tests that compare the ILCs directly. In this paper we explain and illustrate methods for statistical inference that use the ILCs directly. We illustrate the methods with data on incomes by race. In Butler and McDonald (1987) as well as the prior literature, comparisons by race have been the standard application of interdistributional inequality.

The ILC and its index number predecessors rely on a different notion of inequality than the Lorenz curve employs. The latter is purely relative (doubling the incomes in a population has no effect), while the former is sensitive to levels of income. ILCs are constructed using normalized incomplete moments for population subgroups at target incomes in the pooled distribution (i.e., the proportion of a subgroup with incomes less than the target income, the proportion of subgroup income accounted for by members below the target income, etc.). These proportions clearly depend on the income levels for subgroups. Our approach to statistical inference for ILCs covers any incomplete moment and places no prior restrictions on the functional form of the underlying distribution. We can test whether an ILC for a given year differs significantly from the 45-degree line and whether ILCs for two years differ significantly. The latter comparisons reveal whether ILCs are converging toward or diverging from the 45-degree line. We use data from various years of the Current Population Surveys (CPS) to illustrate our approach.

In previous research, Bishop, *et. al.* (2003) generated ILCs for one moment (income shares), but did not offer statistical inference procedures. Bishop, *et. al.* (2004) proposed statistical procedures for comparisons of income shares across groups in a given year, but did not provide a rigorous test for comparisons across years. This paper offers a more comprehensive presentation by permitting comparisons for *any moment* (population shares, income shares, etc.) and by making rigorous comparisons across groups and over time (which allows tests for converging or diverging ILCs).

Section 2 of the paper explains ILCs and statistical inference procedures for comparing them. Section 3 briefly describes our data sample. Section 4 illustrates the statistical inference procedures with data from the CPS. Section 5 provides a summary of our findings.

2. INTERDISTRIBUTIONAL LORENZ CURVES

We first define the essential concepts that provide the foundation for interdistributional Lorenz curves. Let x be a continuous income variable with a probability density $f(x)$ and let $F(x)$ represent the cumulative distribution function (CDF) of x. Let the inverse CDF of x be written $0 \leq F^{-1}(p) \leq \infty$ and, without loss of generality, let $\tau = F^{-1}(p)$ define target incomes. When $p = 0.1, 0.2, \ldots, 1.0$, the target incomes become the decile order statistics. Let I_τ^x be an indicator variable such that $I_\tau^x = 1$ if $x \leq \tau$ and $I_\tau^x = 0$ otherwise.

Given a target income τ, we can define the *h*-th partial moment for $x < \tau$ of the density function $f(x)$ as

$$(1) \qquad M(\tau; h, x) = \int_0^\tau x^h f(x)dx = \int_0^\infty (xI_\tau^x)^h \, dF(x) = E\left[(xI_\tau^x)^h\right],$$

where *E* is the expectation operator. For $h = 0$ the partial moment reduces to $F(\tau)$, which involves accumulating population shares. Following Butler and McDonald (1987), we define the normalized incomplete moment of x for $x \leq \tau$ as

$$(2) \qquad \phi(\tau, h, x) = M(\tau, h, x) / E(x^h),$$

For $h=1$ the normalized incomplete moment gives Lorenz ordinates, $\phi(\tau,1,x)=L(\tau;x)$, which becomes clear if we write the Lorenz curve in the form proposed by Bishop, Chow, and Formby (1994),

$$(3) \qquad L(\tau;x)=\mu_x^{-1}\int_0^{\tau} xf(x)dx=\mu_x^{-1}\int_0^{\infty} xI_\tau^x dF(x)=E\left[xI_\tau^x\right]/E(x),$$

where $E(x)$ is the mean of *x*. We can interpret $L(\tau;x)$ as the proportion of income of *x* received by individuals with incomes *x* less than or equal to a target income τ.

To represent population subgroups, let incomes be classified by *K* mutually exclusive groups $\{\Phi_k, k=1,2,..,K\}$ and define an indicator variable G_k^x such that $G_k^x=1$ if $x\in\Phi_k$ and $G_k^x=0$ otherwise. This indicator variable allows us to rewrite (3), because $E(xI_\tau^x|G_k^x=1)=E(xG_k^xI_\tau^x)/E(G_k^x)$ and $E(x|G_k^x=1)=E(xG_k^x)/E(G_k^x)$. Bishop, Chow, and Zeager (2002) use this approach to show that:

THEOREM 1. *$L(\tau,x)$ can be decomposed by $L(\tau,x^{(k)})$ for k = 1, 2, …, K in that*

$$(4) \qquad L(\tau,x)=\sum_{k=1}^{K}P^{(k)}\cdot L(\tau,x^{(k)}),$$

where $P^{(k)}=E\left[x\cdot G_k^x\right]/E(x)$. We can interpret $P^{(k)}$ as the income share of subgroup k with respect to the income variable *x*.

We show here that similar reasoning can be applied to cases in which $h\neq 1$. That is, expression (2) can also be decomposed by population group *k* (*k* = 1, 2, …, *K*):

$$(5) \qquad \phi(\tau,h,x^{(k)})=M(\tau,h,x^{(k)})/E\left[(x\cdot G_k^x)^h\right]=E\left[(x\cdot G_k^x\cdot I_\tau^x)^h\right]/E\left[(x\cdot G_k^x)^h\right].$$

Since $\phi(\tau, h, x^{(k)})$ is expressed as the expected value of a function of the random income times the indicator variables for targets and population groups, we can easily determine the property of its decomposition from the overall $\phi(\tau, h, x)$ as follows:

THEOREM 2. *$\phi(\tau, h, x)$ can be decomposed by $\phi(\tau, h, x^{(k)})$ for k = 1, 2, …, K in that*

(6) $$\phi(\tau, h, x) = \sum_{k=1}^{K} w^{(k)} \cdot \phi(\tau, h, x^{(k)}),$$

where $w^{(k)} = E\left[(x \cdot G_k^x)^h\right] \big/ E(x^h)$. We can interpret $w^{(k)}$ as the *h*-th moment of the income share of subgroup *k* with respect with the income variable *x*.

Butler and McDonald (1987) use $\phi(\tau, h, x)$ to define two "natural" ILCs for population subgroups. For ease of presentation, let K = 2. Then the first natural ILC is obtained by plotting $\phi(\tau, 0, x^{(1)})$ against $\phi(\tau, 0, x^{(2)})$. If $\phi(\tau, 0, x^{(1)})$ = $\phi(\tau, 0, x^{(2)})$ at each target income (τ), then the ILC corresponds to the 45-degree line. But, if one group (k = 1) is "disadvantaged" [$\phi(\tau, 0, x^{(1)})$ > $\phi(\tau, 0, x^{(2)})$ at each τ], then the ILC lies below the 45-degree line, like an ordinary Lorenz curve. The second natural ILC is obtained by plotting $\phi(\tau, 1, x^{(1)})$ against $\phi(\tau, 1, x^{(2)})$ and is interpreted in similar fashion.

To develop an inference test for the interdistributional Lorenz curves, we first select a set of m income classes or target income levels, denoted by $\{\tau_i | i = 1, 2, ..., m\}$, to which there correspond *K* sets of ILC ordinates { $\phi(\tau_i, h, x^{(k)})$ | *i = 1, 2, …, m, and k = 1,2,…,K*}. This approach allows us to relax the assumption of a continuous CDF, because the Lorenz and concentration ordinates correspond to a set of target incomes instead of a set of quantile functions. Empirically, the targets are selected as a set of sample quantiles ($\hat{\xi}_p$) of the income variable *x*, i.e., p_1 = 0.1, p_2 = 0.2, …, p_9 = 0.9, which in our application are sample deciles. Let a random sample of size N be given from the population. If the CDF of *x* is strictly monotonic, then $\hat{\xi}_p$ has the property of strong or almost sure consistency (Rao 1965, 335).

Let ($x_1, x_2, \ldots, x_N$) be a set of identical and independently distributed (i.i.d.) random sample incomes drawn from the population density *f(x)*. According to equation (5), the decomposed interdistributional Lorenz ordinates can be estimated as

(7) $$\hat{\phi}^h_{i,(k)} = \hat{\phi}(\tau_i, h, x^{(k)}) = \left[N^{-1} \sum_{j=1}^{N} (x_j G_k^{x_j} I_{\tau_i}^{x_j})^h \right] \Bigg/ \left[N^{-1} \sum_{j=1}^{N} (x_j G_k^{x_j})^h \right]$$

Let $\underset{1x(mK)}{\Phi} = (\Phi_1, \Phi_2, ..., \Phi_{mK})' = \left((\phi^h_{1,(1)}, \ldots, \phi^h_{m,(1)}), (\phi^h_{1,(2)}, \ldots, \phi^h_{m,(2)}), \ldots, (\phi^h_{1,(K)}, \ldots, \phi^h_{m,(K)}) \right)'$ be a vector of *mK* decomposed ILC ordinates. The estimates of the vector Φ can be written as $\underset{1x(mK)}{\hat{\Phi}} = (\hat{\Phi}_1, \hat{\Phi}_2, ..., \hat{\Phi}_{mK})' = \left((\hat{\phi}^h_{1,(1)}, \ldots, \hat{\phi}^h_{m,(1)}), (\hat{\phi}^h_{1,(2)}, \ldots, \hat{\phi}^h_{m,(2)}), \ldots, \right.$

$\left. (\hat{\phi}^h_{1,(K)}, \ldots, \hat{\phi}^h_{m,(K)}) \right)'$. From equations (5) through (7), the decomposed ILC ordinates are functions of $E\left[(xG_k^x I_{\tau_i}^x)^h \right]$ and $E\left[(xG_k^x)^h \right]$ for *i* = 1,2, …, *m* and *k* = 1,2,…, *K*. To derive the asymptotic sampling distribution of $\overline{\Phi}$, it is necessary to determine the sampling distributions of these estimates, $\overline{(xG_k^x I_{\tau_i}^x)^h}$,and $\overline{(xG_k^x)^h}$ for *i* = 1,2, …, *m* and *k* = 1,2,…, *K*.

We define the vector of $K(m+1)$ parameter estimators as

$\underset{1x[K(m+1)]}{\overline{\Psi}} = (\overline{\psi}_1, \overline{\psi}_2, ..., \overline{\psi}_{K(m+1)})' = \left(\left(\overline{(xG_1^x I_{\tau_1}^x)^h}, \ldots, \overline{(xG_1^x I_{\tau_m}^x)^h}, \overline{(xG_1^x)^h} \right), \ldots, \left(\overline{(xG_K^x I_{\tau_1}^x)^h}, \ldots, \right. \right.$
$\left. \left. \overline{(xG_K^x I_{\tau_m}^x)^h}, \overline{(xG_K^x)^h} \right) \right)'$.

THEOREM 3. *Suppose* $(x_1, x_2, ..., x_N)$ *are i.i.d. random samples of a size of N drawn from the population density function f(x). Given a set of predefined target incomes* $\{\tau_i | i = 1,2,3,..., m\}$ *such that* $0 < \tau_1 < \ldots < \tau_m < \infty$, *and a population decomposed into K mutually exclusive groups, the vector* $\sqrt{N}(\overline{\Psi} - \Psi)$ *converges in probability to a K(m+1) variate normal distribution with mean zero and a variance-covariance* $\Omega = (\sigma_{i,j})$, *where*

$$\underset{(K(m+1))x(K(m+1))}{\Omega} = \begin{bmatrix} \left[\alpha_{ij}^{11}\right] & \cdots & \left[\alpha_{ij}^{1K}\right] \\ \vdots & \ddots & \vdots \\ \left[\alpha_{ij}^{K1}\right] & \cdots & \left[\alpha_{ij}^{KK}\right] \end{bmatrix}, \text{ and}$$

$$\left[\alpha_{ij}^{kl}\right]_{(m+1)x(m+1)} = \begin{cases} Cov[(xG_k^x I_{\tau_i}^x)^h, (xG_l^x I_{\tau_j}^x)^h] & \text{for } i, j \le m \\ Cov[(xG_k^x I_{\tau_i}^x)^h, (xG_l^x)^h] & \text{for } i \le m, \quad j = (m+1) \\ Cov[(xG_k^x)^h, (xG_l^x I_{\tau_j}^x)^h] & \text{for } i = (m+1), \ j \le m \\ Cov[(xG_k^x)^h, (xG_l^x)^h] & \text{for } i = j = (m+1) \end{cases},$$

where *Cov* denotes the covariance measure.

PROOF. Given that the income samples *x* and the indicator variables *G* and *I* are *i.i.d.*, the *h*-th power function *i.i.d.* random variable is also *i.i.d.* From direct calculations, it can be shown that $E(\overline{\psi}_i) = \psi_i$, $i = 1,2,\ldots, K(m+1)$. Then, for large samples, the Kolmogorov Strong Law of Large Numbers implies that $\overline{\psi}_i$ converges in probability to ψ_i. From the Lindeberg-Levy Central Limit Theorem, we obtain the result that $\sqrt{N}(\overline{\psi_i} - \psi_i)$ converges in distribution to $N(0, \sigma_i^2)$. Finally, from the Cramer-Wald Theorem, it can be shown that $\sqrt{N}(\overline{\Psi} - \Psi)$ converges to a multivariate normal distribution, $N(\underline{0}, \Omega)$. Q.E.D.

Theorem 3 allows us to analyze the sampling distribution of the estimated decomposed ILC ordinates. Applying Rao's (1965) theorem on the limiting distribution of differentiable functions of random variables, the limiting distribution of $\hat{\Phi}$ is also multivariate normal. We summarize this result in the following theorem.

THEOREM 4. *Under the conditions of Theorem 3, the vector of estimated decomposed ILC ordinates* $\underset{1x(mK)}{\hat{\Phi}} = (\hat{\Phi}_1, \hat{\Phi}_2, ..., \hat{\Phi}_{mK})' = \left((\hat{\phi}^h_{1,(1)}, \ldots, \hat{\phi}^h_{m,(1)}), (\hat{\phi}^h_{1,(2)}, \ldots, \hat{\phi}^h_{m,(2)}), \ldots, (\hat{\phi}^h_{1,(K)}, \ldots, \hat{\phi}^h_{m,(K)}) \right)'$ *is asymptotically normal in that* $\sqrt{N}(\hat{\Phi} - \Phi)$ *has a limiting Km-variate normal distribution with mean zero and covariance matrix* $V = J\Omega J' = (v_{ij})$, *where* Ω *is defined in Theorem 3 and J is defined as* $\underset{(2Km)x(2Km+k)}{J} = \left[\delta\overline{\Phi}_j / \delta\overline{\Psi}_j \right]_{\overline{\Psi}=\Psi}$.

Then, the covariance estimate of the *k*-th and *l*-th estimated decomposed ILC ordinates, $\hat{\phi}^h_{j,(k)}$ and $\hat{\phi}^h_{j,(l)}$, can then be determined as follows:

$$\begin{aligned}\overline{Cov}(\hat{\phi}^h_{j,(k)}, \hat{\phi}^h_{j,(l)}) = {} & a_k a_l \left[\overline{(xG^x_k I^x_{\tau_j})^h (xG^x_l I^x_{\tau_j})^h} - \left(\overline{(xG^x_k I^x_{\tau_j})^h}\right)\left(\overline{(xG^x_l I^x_{\tau_j})^h}\right) \right] + \\ & c_{jk} c_{jl} \left[\overline{(xG^x_k)^h (xG^x_l)^h} - \left(\overline{(xG^x_k)^h}\right)\left(\overline{(xG^x_l)^h}\right) \right] + \\ & a_k c_{jk} \left[\overline{(xG^x_k I^x_{\tau_j})^h (xG^x_k)^h} - \left(\overline{(xG^x_k I^x_{\tau_j})^h}\right)\left(\overline{(xG^x_k)^h}\right) \right] + \\ & a_l c_{jl} \left[\overline{(xG^x_l I^x_{\tau_j})^h (xG^x_l)^h} - \left(\overline{(xG^x_l I^x_{\tau_j})^h}\right)\left(\overline{(xG^x_l)^h}\right) \right]\end{aligned}$$

where $a_k = \left(\overline{(xG^x_k)^h}\right)^{-1}$, $a_l = \left(\overline{(xG^x_l)^h}\right)^{-1}$, $c_{jk} = -\left(\overline{(xG^x_k I^x_{\tau_j})^h}\right)\left(\overline{(xG^x_k)^h}\right)^{-2}$, and $c_{jl} = -\left(\overline{(xG^x_l I^x_{\tau_j})^h}\right)\left(\overline{(xG^x_l)^h}\right)^{-2}$ for $j = 1, 2, ..., m$.

One can perform a goodness-of-fit test for a *marginal* change in the subgroup moments. That is, one can test for differences in the decomposed moments for subgroup 1 (e.g., $\underset{mx1}{\hat{\psi}} = (\hat{L}^{(1)}_1 \text{-} \hat{C}^{(1)}_1, ..., \hat{L}^{(1)}_m \text{-} \hat{C}^{(1)}_m)'$). Under the null hypothesis that $H_0: \psi = \underline{0}$, an appropriate test statistic is

(8) $$c = N\hat{\psi}'\hat{\Theta}^{-1}\hat{\psi},$$

where $\hat{\Theta} = \rho \hat{V} \rho'$,

and
$$\rho = \begin{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix} & \begin{bmatrix} -1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & \cdots & 0 \end{bmatrix} \end{bmatrix}_{m \text{ x } 2Km} .$$

From Theorem 3, the c-statistic is also asymptotically distributed as a (central) chi-squared variate with *m*-degrees of freedom.

Alternatively, one could test a joint hypothesis such that $\{ H_{0j} : \psi_j = L_j^{(1)}$ - $C_j^{(1)} = 0 | j = 1,..., m\}$. Then the appropriate test statistics are

(9) $$Z_j = \sqrt{N} \frac{\hat{\psi}_j}{\sqrt{\hat{\Theta}_{jj} / N}} \text{, for } j = 1, 2, ..., m .$$

Let $Z^* = \max_{1 \le j \le m} |Z_j|$ be the largest absolute value of the test statistics. We then apply the Sidak (1967) probability inequality and the results in Hochberg (1974) and Richmond (1982) to control the size of the multiple sub-hypothesis tests.

THEOREM 5. *Let* $Z = (Z_1, ..., Z_m)'$ *be a vector of m test statistics corresponding to* (9). *From Theorem 4, the distribution of vector Z converges asymptotically to an m-variate normal distribution. Under the null hypothesis, the confidence interval of at least* $100(1-\alpha)$ *percent for the extreme statistic,* $Z^*(\tau)$*, can be defined as:*

(10) $$Z^*(\tau) \pm SMM(\alpha; m; \infty) ,$$

where $SMM(\alpha; m; \infty)$ *is the asymptotic critical value of the* α*-point of the Studentized Maximum Modulus (SMM) distribution (Stoline and Ury 1979) with parameter m and* ∞ *degrees of freedom.*

Further, let $Z^{*+} = \max_{1 \le j \le m} Z_j$, and $Z^{*-} = \min_{1 \le j \le m} Z_j$. The asymptotic joint confidence interval of at least $100(1-\alpha)$ percent is:

(11) $$-SMM(\alpha; m; \infty) \le Z^{*-} ... \le Z^{*+} \le SMM(\alpha; m; \infty) .$$

We emphasize that test statistic c in (8) illustrates only one possibility, and that the results of Theorem 3 and the SMM approach for controlling the joint test size can be applied to a wide range of hypothesis tests for subgroup income distribution comparisons. We also note that the proposed methodology employs a finite-target testing approach, so it may have low power in detecting tail inequality for fat-tail distributions. Thus, the power of the tests is an issue for further research.

3. Data

The data for our empirical analysis are drawn from the Current Population Surveys (CPS) for 1977, 1987, 1992, 1997, and 2002. We avoid comparisons over the years 1993-95, when the CPS made substantial changes in the top-coding of incomes in the public-use sample (Burkhauser, et. al., 2004). We also restrict the samples to primary families (i.e., excluding single-person families and unrelated individuals). We correct for inflation to allow pooling of incomes across time, but make no adjustments for the size and composition of the family.

To simplify the exposition we will refer to nonwhites as “blacks” in the application. Table 1 presents the mean incomes of whites and blacks, adjusted for inflation, over the years we consider. Before the change in top-coding (1976-91), the mean incomes for both population subgroups increased slowly, 5.45 percent for whites and 3.86 percent for blacks. After the change in top-coding (1996-2001), mean incomes increased much more rapidly, 13.54 percent for whites and 11.37 for blacks. As the next section will show, ILCs are sensitive to changes in both the level and dispersion of the distributions by population subgroup.

4. Application

This section estimates interdistributional Lorenz curves (ILCs) for whites and blacks. We test whether ILCs differ from the 45-degree line in a particular year and from each other across years. The former tests show whether one population subgroup has an "economic advantage" over another (Deutsch and Silber 1999), while the latter tests reveal whether the advantage is narrowing or widening over time.

Figure 1 illustrates how we implement statistical inference procedures for ILCs. We select the order statistics (upper income cutoffs) for the deciles in the distribution of pooled incomes (across subgroups and time) as target incomes and estimate each ILC at ten points. For each target income, we plot the corresponding black and white shares on the horizontal and vertical axes, respectively. By construction, the black-white difference will be the vertical distance between the 45-degree line and the ILC. We can then test the differences for statistical significance using methods described in section 2. Given ILCs for two years, we can test whether differences in differences are statistically significant, i.e., whether an ILC for two subgroups is converging or diverging over time.

Table 2a presents statistical tests for white-black ILCs, based upon population shares (h=0) in 1976 and 1991 (before the change in top-coding for the CPS public-use samples). Column (1) shows the target incomes, which are decile order statistics for the distribution of incomes pooled across subgroup and time. Columns (2) and (3) report the estimated population shares in 1976 for whites and blacks at or below each target income, with standard errors in parentheses. Column (4) gives differences in population shares by race. Columns (5)–(7) give the corresponding information for 1991. Column (8) reports the "difference in differences" over time, while column (9) provides the test statistics Z_j, j=1,2,…,10, from equation (9) in section 2. We also give the chi-squared test statistic c from equation (8) in section 2 in the bottom row of Table 2a.

An inspection of columns (4) and (7) of Table 1a reveals that the estimated differences between the 45-degree line and the 1976 and 1991 ILCs are large relative to their standard errors. The chi-squared statistics in both columns are highly significant as well. Both results indicate that whites have an economic advantage over blacks for these years. Column (8) alerts us to a possible crossing of ILCs for 1976 and 1991, because the difference in differences yields both negative and positive signs. However, the SMM test statistics in column (9) do not support a crossing, as no positive test statistic is significant at the ten-percent level. Hence, we find that the white-black ILC (for population shares) shifted *away from* the 45-degree line during 1976-91, creating a widening advantage for whites over blacks during the period.

To explore why the advantage of whites over blacks (in population shares) changed over time, consider the rows of Table 2a in more detail. Recall (Table 1) that mean incomes rose ─ albeit slowly ─ during 1976-91 and grew slightly faster for whites than blacks. Nevertheless Table 2a shows that population shares at or below fixed target incomes *rose* in the bottom four deciles for whites and blacks, which mean that incomes were falling at the bottom of the distribution. Population shares declined in the upper six deciles for whites and blacks, indicating rising incomes (as reflected in the rising means). These patterns indicate a widening dispersion of incomes in both subgroups. Changes in both the level and dispersion of incomes in each subgroup influence the comparisons of ILCs over time, because they affect the population shares at or below fixed incomes.

Table 2b presents the statistical tests for white-black ILCs, based on income shares (h=1), rather than population shares, in 1976 and 1991. Once again, the black-white differences in columns (4) and (7) are large relative to their standard errors and the chi-squared statistics in these columns are highly significant, which implies an advantage for whites over blacks using income shares. In column (8), a few positive "difference in differences" are statistically significant, but the negative one (in the top decile) is not, so the ILC for income shares *converged* toward the 45-degree line, unlike the population-share ILC in Table 2a. This finding is reinforced by the chi-squared statistic in column (8), which is statistically significant. Figure 2 illustrates the movement of the income-share ILC toward the 45-degree line over time.

Table 3 provides a summary of our ILC comparisons. The first row shows the results for 1976-91 that have been described above. The second row identifies a period (1986-91) for which we find no significant differences between ILCs, using population shares or income shares. The last row identifies a later period (1996-2001) in which a comparison of ILCs yields a statistically significant crossing. In all cases except the crossing, the chi-squared statistic and the dominance comparisons generate identical conclusions, but the chi-squared statistic is misleading in the case of a crossing.

5. Conclusions

We present distribution-free stochastic dominance tests for differences in the interdistributional Lorenz curves (ILCs) proposed by Butler and McDonald (1987). We illustrate the procedures with ILCs constructed from population and income shares. The data on white and nonwhite family incomes are drawn from the CPS public-use files for various years between 1976 and 2001. We avoid years in which the income top-codes changed. Our comparisons illustrate periods of converging, diverging, statistically equivalent, and crossing ILCs.

REFERENCES

Bishop, J. A. , K. V. Chow, and J. P. Formby, (1994) "Testing for Marginal Changes in Income Distributions with Lorenz and Concentration Curves," *International Economic Review*, 35, 479-488.

Bishop, J. A., K. V. Chow, F. Luo, and L. A. Zeager, (2003) "Changes in Economic Advantage by National Origin after German Unification," *Estadistica*, 55: 197-212 (available at http://www.ecu.edu/econ/wp/02/ecu0206.pdf).

Bishop, J. A., K. V. Chow, and L. A. Zeager, (2002) "Decomposing Lorenz and Concentration Curves," *International Economic Review*, 44, 965-978.

Bishop, J. A., K. V. Chow, and L. A. Zeager, (2004) "Lorenz Decomposition and Interdistributional Lorenz Comparisons," *Research on Income Inequality*, 12, forthcoming (available at http://www.ecu.edu/econ/wp/03/RIIvol12BCZ.pdf).

Burkhauser, R. V., J. S. Butler, S. Feng, and A. J. Houtenville, (2004) "Long Term Trends in Earnings Inequality: What the CPS Can Tell Us," *Economics Letters*, 82, 295-299.

Butler, R. J., and J. B. McDonald, (1987) "Interdistributional Income Inequality," *Journal of Business and Economic Statistics*, 5, 13-18.

Dagum, C. (1980) "Inequality Measures Between Income Distributions with Applications," *Econometrica*, 48, 1791-1803.

Dagum, C. (1987) "Measuring the Economic Affluence Between Populations of Income Receivers," *Journal of Business and Economic Statistics*, 5, 5-12.

Deutsch, J., and J. Silber, (1997) "Gini's 'Transvariazione' and the Measurement of Inequality: Within and Between Distributions," *Empirical Economics*, 22, 547-554.

Deutsch, J., and J. Silber, (1999) "Inequality Decomposition by Population Subgroup and the Analysis of Interdistributional Inequality", in J. Silber (Ed.), *Handbook of Income Inequality Measurement* (Dordrecht: Kluwer) (pp. 363-397).

Gastwirth, J. L. (1985) "Comment on 'Measurement of Economic Distance Between Blacks and Whites' by H.D. Vinod," *Journal of Business and Economic Statistics*, 3, 405-409.

Gini, C. (1916) Il Concetto di "Transvariazione" e le sue prime applicazioni," Studi di Economia, Finanza e Statistica, editi del Giornali degli Economisti e Revista de Statistica, Reprinted in Gini, 1959.

Gini, C. (1959) *Memorie de Metadologia Statistica: Volume Secondo – Transvariazione.* Rome: Libreria Goliardica.

Hochberg, Y. (1974) "Some Generalizations of T-Method in Simultaneous Inference, *Journal of Multivariate Analysis*, 4, 224-234.

Rao, C. R. (1965) *Linear Statistical Inference and its Application*, (New York: Wiley).

Richmond, J. (1982) "A General Method for Constructing Simultaneous Confidence Intervals," *Journal of the American Statistical Association*, 77, 455-460.

Shorrocks, A. F. (1982) "On the Distance Between Income Distributions," *Econometrica*, 50, 1337-1339.

Sidak, Z. (1967) "Rectangular Confidence Regions for the Mean of Multivariate Normal Distributions," *Journal of the American Statistical Association*, 62, 626-633.

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

Vinod, H. D. (1985) "Measurement of Economic Distance Between Blacks and Whites," *Journal of Business and Economic Statistics*, 3, 78-88.

Figure 1
Comparing an ILC with the 45-Degree Line

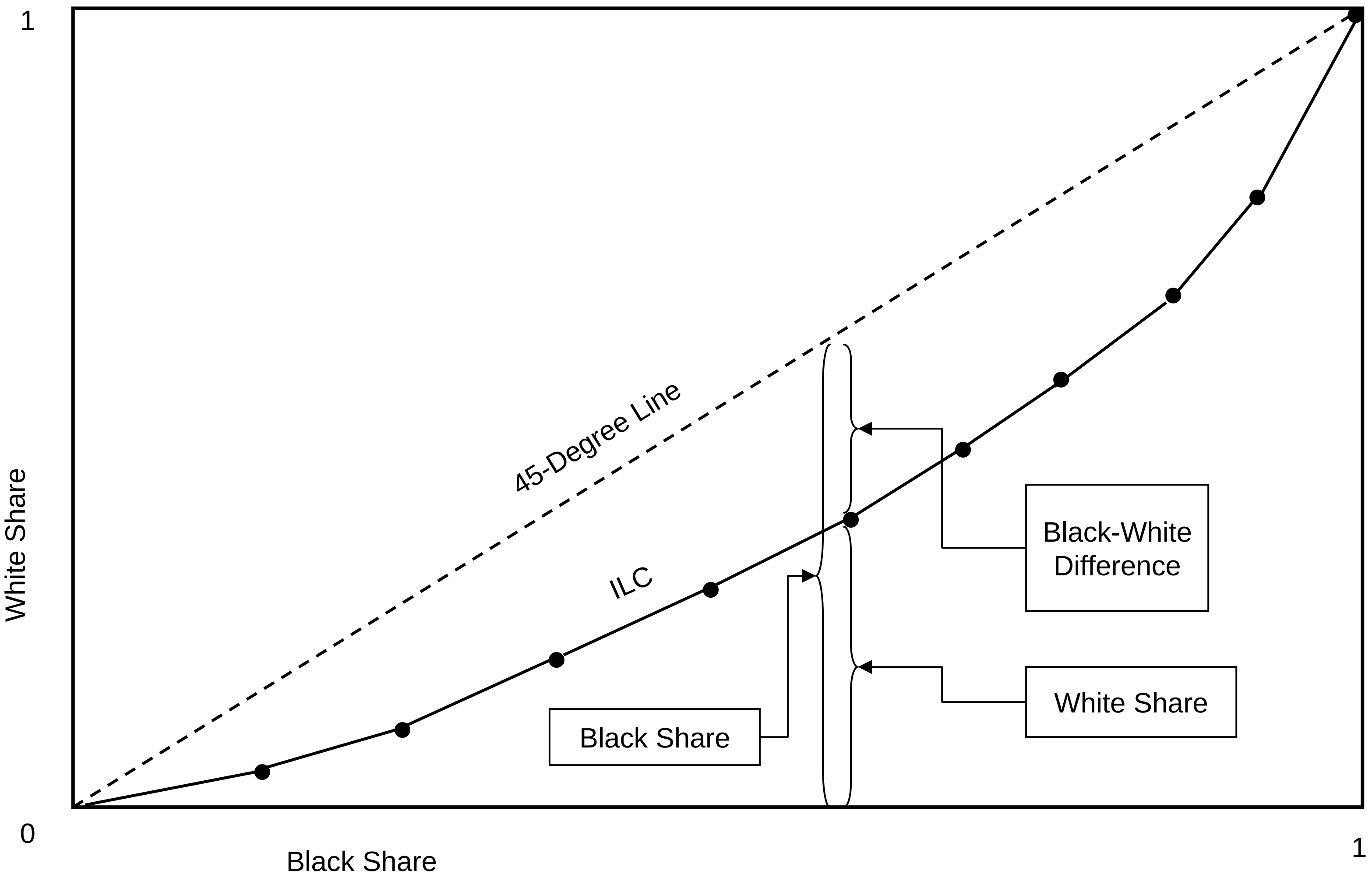

Table 1
Mean Incomes for Whites and Blacks

Year	White	Black
1976	53,393 (183.0)	35,483 (426.0)
1986	57,193 (217.0)	37,736 (497.0)
1991	56,303 (215.0)	36,853 (502.0)
1996	61,646 (335.0)	40,886 (776.0)
2001	69,995 (301.0)	45,535 (541.0)

Note: All incomes are expressed in 2001 dollars. The numbers in parentheses are standard errors.

Table 2a
Black-White Interdistributional Lorenz Ordinates (Population Shares) for White vs Black: 1976 and 1991

	1976			1991			Converging ILCs?	
Target Income (1)	**White Share (2)**	**Black Share (3)**	**Black-White Difference (4)=(3)-(2)**	**White Share (5)**	**Black Share (6)**	**Black-White Difference (7)=(6)-(5)**	**Difference in Differences 8=(4)-(7)**	**Test Statistic (9)**
14,728	0.0768 (0.0014)	0.2352 (0.0070)	0.1584 (0.0071)	0.0945 (0.0015)	0.2811 (0.0072)	0.1865 (0.0073)	-0.0281 (0.0102)	-2.76
23,058	0.1767 (0.0020)	0.4083 (0.0081)	0.2317 (0.0083)	0.1901 (0.0021)	0.4145 (0.0078)	0.2244 (0.0081)	-0.0073 (0.0112)	0.63
31,025	0.2767 (0.0023)	0.5277 (0.0082)	0.2510 (0.0085)	0.2926 (0.0024)	0.5290 (0.0079)	0.2364 (0.0083)	0.0146 (0.0112)	1.23
39,042	0.3878 (0.0025)	0.6290 (0.0079)	0.2412 (0.0083)	0.4000 (0.0026)	0.6313 (0.0077)	0.2314 (0.0081)	0.0099 (0.0112)	0.85
47,808	0.5038 (0.0026)	0.7355 (0.0072)	0.2317 (0.0077)	0.4995 (0.0026)	0.7139 (0.0072)	0.2143 (0.0077)	0.0173 (0.0108)	1.59
56,798	0.6245 (0.0025)	0.8191 (0.0063)	0.1946 (0.0068)	0.6010 (0.0026)	0.7931 (0.0065)	0.1921 (0.0069)	0.0025 (0.0097)	0.26
67,814	0.7390 (0.0023)	0.8901 (0.0051)	0.1510 0.0000	0.7006 (0.0024)	0.8592 (0.0055)	0.1586 (0.0060)	-0.0076 (0.0082)	-0.92
82,377	0.8455 (0.0019)	0.9438 (0.0038)	0.0983 (0.0042)	0.8005 (0.0021)	0.9137 (0.0045)	0.1131 0.0049	-0.0148 (0.0065)	-2.28
108,398	0.9353 (0.0013)	0.9847 (0.0020)	0.0494 (0.0024)	0.9031 (0.0016)	0.9648 (0.0029)	0.0617 (0.0033)	-0.0123 (0.0041)	-3.02
Chi-Square Statistic			1456.7			1267.2	40.0	

Note: The target incomes are the decile order statistics for the distribution pooled across years and population subgroups expressed in 2001 dollars. Numbers in parenthesis are standard errors. The test statistic for converging ILCs is a student maximum modulus (SMM). The critical values of the SMM for the 5-percent and 10-percent levels are 2.76 and 2.52, respectively.

Table 2b
Black-White Interdistributional Lorenz Ordinates (Income Shares) for White vs Black: 1976 and 1991

Target Income (1)	1976 White Share (2)	1976 Black Share (3)	1976 Black-White Difference (4)=(3)-(2)	1991 White Share (5)	1991 Black Share (6)	1991 Black-White Difference (7)=(6)-(5)	Converging ILCs? Difference in Differences 8=(4)-(7)	Converging ILCs? Test Statistic (9)
14,728	0.0139 (0.0003)	0.0634 (0.0025)	0.0495 (0.0025)	0.0150 (0.0003)	0.0608 (0.0023)	0.0458 (0.0023)	0.0037 (0.0034)	1.09
23,058	0.0489 (0.0007)	0.1538 (0.0047)	0.1049 (0.0048)	0.0470 (0.0006)	0.1273 (0.0041)	0.0802 (0.0041)	0.0246 (0.0063)	3.91
31,025	0.0992 (0.0011)	0.2446 (0.0066)	0.1454 (0.0067)	0.0955 (0.0011)	0.2100 (0.0060)	0.1146 (0.0061)	0.0309 (0.0090)	3.42
39,042	0.1717 (0.0016)	0.3416 (0.0083)	0.1744 (0.0085)	0.1600 (0.0016)	0.3029 (0.0028)	0.1429 (0.0029)	0.0315 (0.0116)	2.71
47,808	0.2613 (0.0022)	0.4728 (0.0099)	0.2115 (0.0101)	0.2353 (0.0021)	0.4013 (0.0093)	0.1660 (0.0095)	0.0456 (0.0139)	3.27
56,798	0.3762 (0.0027)	0.5917 (0.0108)	0.2156 (0.0111)	0.3268 (0.0026)	0.5105 (0.0106)	0.1837 (0.0109)	0.0319 (0.0156)	2.04
67,814	0.5058 (0.0031)	0.7135 (0.0110)	0.2077 (0.0114)	0.4335 (0.0030)	0.6231 (0.0114)	0.1897 (0.0118)	0.0180 (0.0164)	1.10
82,377	0.6549 (0.0034)	0.8232 (0.0103)	0.1683 (0.0108)	0.3649 (0.0034)	0.7308 (0.0115)	0.1659 (0.0120)	0.0024 (0.0162)	0.15
108,398	0.8140 (0.0032)	0.8365 (0.0076)	0.1225 (0.0083)	0.7341 (0.0036)	0.8628 (0.0103)	0.1288 (0.0109)	0.0063 (0.0127)	-0.45
Chi-Square Statistic			746.8			564.2	27.0	

Note: The target incomes are the decile order statistics for the distribution pooled across years and population subgroups expressed in 2001 dollars. Numbers in parenthesis are standard errors. The test statistic for converging ILCs is a student maximum modulus (SMM). The critical values of the smm for the 5-percent and 10-percent levels are 2.76 and 2.52, respectively.

Table 3
Summary of ILC Comparisons

Period	Population Shares (h=0)	Income Shares (h=1)
1976-1991	Diverging (40.0)	Converging (27.0)
1986-1991	No Difference (5.9)	No Difference (3.4)
1996-2001	Crossing (48.4)	Converging (45.1)

Note: The numbers in parentheses are chi-squared statistics.