KATHOLIEKE UNIVERSITEIT
LEUVEN

Faculty of Economics and Applied Economics

Department of Economics

Inference for the Measurement of Poverty in the Presence of a Stochastic Weighting Variable

by

Bram THUYSBAERT

Econometrics

Center for Economic Studies
Discussions Paper Series (DPS) 05.02
http://www.econ.kuleuven.be/ces/discussionpapers/default.htm

January 2005

DISCUSSION PAPER

Inference for the measurement of poverty in the presence of a stochastic weighting variable

Bram Thuysbaert*

K.U. Leuven

November 2004

Abstract

Empirical applications of poverty measurement often have to deal with a stochastic weighting variable such as household size. Within the framework of a bivariate distribution function defined over income and weight, I derive the limiting distributions of the decomposable poverty measures and of the ordinates of stochastic dominance curves. The poverty line is allowed to depend on the income distribution. It is shown how the results can be used to test hypotheses concerning changes in poverty. The inference procedures are briefly illustrated using Belgian data.

JEL classification: C40; I32

Keywords: decomposable poverty measure; stochastic dominance; statistical inference

1 Introduction

In recent years, important contributions to the econometric literature on poverty measurement have been made (e.g. Davidson and Duclos, 2000; Zheng, 2001b).

*I am most grateful to Geert Dhaene for helpful comments and many valuable suggestions. Address for correspondence: Bram Thuysbaert, K.U. Leuven, Department of Economics, Naamsestraat 69, B-3000 Leuven, Belgium. Tel. +32 16 32 68 10. Email: bram.thuysbaert@econ.kuleuven.be

By now, the sampling distributions of most commonly used indices have been derived and many testing procedures for stochastic dominance have been proposed. The empirical application of the statistical results, however, lags behind. This may be due to the fact that the application of the theoretical results to the problem at hand is not always straightforward. A frequently encountered problem is the presence of a stochastic weighting variable. This situation occurs naturally when income is weighted by the number of household members. The randomness of the weighting variable affects the sampling distribution of the estimators and has to be accounted for.

Cowell (1989) addresses the problem of stochastic weighting variables. He derives the limiting distributions of the generalized entropy indices and of the Gini coefficient estimated from a sample drawn from a bivariate distribution of income and weight. This paper extends Cowell's analysis to the measurement of poverty. Davidson and Duclos (2000) derive the limiting distribution of the ordinates of stochastic dominance curves and Zheng (2001b) derives the limiting distribution of the decomposable poverty measures with a random poverty line. These authors, however, do not consider stochastic weighting variables. I extend the results of Davidson and Duclos (2000) and Zheng (2001b) so that they can be used in a situation where a stochastic weighting variable is present.

Section 2 presents the indices and dominance results for which inference procedures are developed. Section 3 discusses estimation and hypothesis testing. In section 4, the inference procedures are briefly illustrated with an application to Belgian data.

2 Poverty measurement

Most often, the interest of the empirical researcher is in comparing poverty between regions or over time. Two approaches can be taken. The first is based on the use of *indices.* Here the researcher selects one of the many indices that have been proposed in the literature and compares the values the index takes for the different distributions. An important class of poverty measures is the class of decomposable poverty measures. Let z denote the poverty line. A decomposable, additively separable poverty measure has the form

$$P_F(z) = \int_0^z \pi(x,z)dF(x), \tag{1}$$

where $F(x)$ is an income distribution function with support on the positive half-line and $\pi(x,z)$ is a function evaluating the contribution of the income receiver with income x to total poverty in the population. Throughout, $\pi(x,z)$ is assumed to be differentiable in x and z for $x, z \in (0,\infty)$. It is also assumed that $\partial\pi(x,z)/\partial z > 0$.

Many commonly used indices are decomposable. For the FGT measures (Foster *et al.*, 1984), $\pi(x,z) = (1 - x/z)^a$, where a is a natural number. If $\pi(x,z) = \log(z/x)$, the resulting poverty measure is the Watts measure (Watts, 1986). If $\pi(x,z) = b^{-1}[1 - (x/z)^b]$, a monotonic transformation of the second measure proposed by Clark *et al.* (1981) is obtained.

Since each index establishes a particular ordering over income distributions, the ordering of a given set of distributions may depend on the choice of the index. The use of poverty *dominance* criteria – the second approach – allows the researcher to make poverty comparisons in a more robust way. Dominance ensures that all indices of some well-defined class unanimously prefer one distribution to another. A major drawback, however, is the incompleteness of the ordering associated with each dominance criterion. If dominance curves cross, the criterion is inconclusive.

In the literature, a primal and a dual approach to stochastic dominance have been developed. *Primal* stochastic dominance functions are defined over income. Let $D^1_F(x) = F(x)$ be the first order dominance function of the distribution F and define, for any integer $s \geqslant 2$, the s-th order dominance function as

$$D_F^s(x) = \int_0^x D_F^{s-1}(y)dy.$$

It is well-known in the stochastic dominance literature (e.g. Fishburn, 1976) that $D_F^s(x)$ can be expressed as

$$D_F^s(x) = \frac{1}{(s-1)!}\int_0^x (x-y)^{s-1}dF(y). \tag{2}$$

F is said to dominate G up to x^* at order s if and only if $D_G^s(x) \geqslant D_F^s(x)$ for all $x \in [0, x^*]$.

Poverty dominance criteria relate stochastic dominance of some order to classes of poverty measures. For the sake of brevity, the discussion here is restricted to first and second order dominance. For higher order results, the reader is referred to Davidson and Duclos (2000), Zheng (2000) and Duclos and Makdissi (2004). Let Π^1 denote the class of decomposable poverty measures satisfying $\partial\pi(x,z)/\partial x < 0$ for $x \in (0,z)$. This monotonicity condition requires that a poverty measure increases if a poor person's income decreases. Let Π^2 denote the subclass of Π^1 containing the poverty measures satisfying $\partial^2\pi(x,z)/\partial x^2 > 0$ for $x \in (0,z)$. This convexity condition imposes the transfer principle, requiring an income transfer from a poor person to someone richer to increase poverty. Applying Atkinson's results (1987) to the decomposable poverty measures yields the following proposition.

Proposition 1 *For* $s = 1, 2$: $D_G^s(\lambda z) \geqslant D_F^s(\lambda z)$ *for all* $\lambda \in [0,1]$ $\iff$ $P_G(\lambda z) \geqslant P_F(\lambda z)$ *for all* $P \in \Pi^s$ *and for all* $\lambda \in [0,1]$.

Note that this result requires the maximum poverty line to be the same for both distributions. In practice, however, many circumstances may motivate the use of different maximum poverty lines for F and G. By restricting the set of poverty indices to the indices for which $\pi(x,z)$ is homogeneous of degree zero in x and z, Proposition 1 can be reformulated so that it holds for the case of distribution-specific maximum poverty lines as well. If $\pi(x, z)$ is homogeneous of degree zero, the resulting poverty measures are invariant under rescaling the poverty line and all incomes and may therefore be called relative poverty measures. Let Ψ^1 and Ψ^2 denote the set of relative poverty measures contained in Π^1 and Π^2, respectively, and let z_F and z_G be the maximum poverty lines censoring the distributions F and G. It is not difficult to see that the following proposition holds (e.g. Davidson and Duclos, 2000).

Proposition 2 *For* $s = 1, 2 : \frac{D_G^s(\lambda z_G)}{[z_G]^{s-1}} \geqslant \frac{D_F^s(\lambda z_F)}{[z_F]^{s-1}}$ *for all* $\lambda \in [0, 1] \iff P_G(\lambda z_G) \geqslant P_F(\lambda z_F)$ *for all* $P \in \Psi^s$ *and for all* $\lambda \in [0, 1]$.

Whereas primal dominance functions are defined over income, *dual* dominance functions are defined over population shares. Dominance is checked by comparison of the functions at all population shares $p \in [0, 1]$. The dual approach is therefore also referred to as the p-approach or the quantile approach. In a series of papers, Jenkins and Lambert (1997; 1998a; 1998b) have developed a quantile approach to second order poverty dominance. Following Spencer and Fisher (1992), Jenkins and Lambert suggest the use of the cumulative poverty gap curve which they termed the three I's of poverty or TIP curve. The TIP curve is defined as

$$TIP_F(p, z) = \int_0^{Q_F(p)} (z - x) I(x \leqslant z) dF(x) \qquad \text{with } p \in [0, 1].$$

Jenkins and Lambert noticed that the incidence, intensity and inequality of poverty can be read from the curve. The point on the horizontal axis where the curve becomes flat is the headcount ratio. The TIP ordinate of the headcount ratio is the average poverty gap and the curvature of the TIP curve reflects the inequality of income distribution among the poor.

The TIP curve is interesting not only because it provides a complete graphic poverty profile, but also because it comes equipped with a dominance criterion. Jenkins and Lambert (1998b) show that TIP dominance is equivalent to second order poverty dominance.

Proposition 3 $TIP_G(p,z) \geqslant TIP_F(p,z)$ *for all* $p \in [0,1] \iff P_G(\lambda z) \geqslant P_F(\lambda z)$ *for all* $P \in \Pi^2$ *and for all* $\lambda \in [0,1]$.

The dual statement of Proposition 2 for $s = 2$ makes use of the normalized TIP curve defined as

$$TIP_F^n(p,z) = \frac{1}{z} TIP_F(p,z) \qquad \text{with } p \in [0,1].$$

The result is as follows (Jenkins and Lambert, 1998b).

Proposition 4 $TIP_G^n(p,z_G) \geqslant TIP_F^n(p,z_F)$ *for all* $p \in [0,1] \iff P_G(\lambda z_G) \geqslant P_F(\lambda z_F)$ *for all* $P \in \Psi^2$ *and for all* $\lambda \in [0,1]$.

3 Inference

In recent years, considerable research has gone into the development of procedures for the statistical testing of changes in poverty. Applying the statistical results that can be found in the literature to the problem at hand is not always straightforward. A frequently encountered problem is that the data set contains observations on households, whereas the interest of the researcher is mostly in the distribution of income among individuals. The common practice in empirical work is to make household incomes comparable by the use of some equivalence scale. Equivalent incomes are then weighted by the number of household members or some function thereof. This practice creates a statistical difficulty. Since household size is stochastic, the joint distribution of income and household size has to be considered in deriving asymptotic distributions. A second problem arises if the samples from the distributions that are to be compared are – fully or partially – dependent. This situation is not uncommon when studying the evolution of poverty over time. Sample dependency causes the covariance of the estimates to be non-zero and has to be accounted for in the estimation procedure.

As in Cowell (1989), the randomness of the weighting variable is dealt with by considering a joint distribution function $F(x, w)$ of income x and weight w, where it is assumed that both x and w are positive. Within this framework, I will first derive the limiting distributions of the estimators of the decomposable poverty measures and of the ordinates of stochastic dominance curves. The results are then extended to the dependent sample case. After having discussed estimation of the covariance structure of the estimates, the second subsection focuses on procedures for hypothesis testing.

It is convenient at this stage to introduce some notation. Let

$$\mu_{10} = \int \int w dF(x, w),$$
$$\mu_{11} = \int \int wx dF(x, w),$$

and let m_{10} and m_{11} denote the corresponding sample analogues. Note that $\frac{\mu_{11}}{\mu_{10}}$ is the mean individual income derived from the distribution $F(x, w)$. Also define

$$J_F(y) = \frac{1}{\mu_{10}} \int_0^y \int w dF(x, w).$$

The function J_F may be thought of as the distribution function of the individual incomes. Throughout, it is assumed that J_F is differentiable and has finite population moments up to the required order.

3.1 Estimation

The poverty line

The limiting distributions of the decomposable poverty indices and of the ordinates of primal and dual dominance curves depend on whether the poverty line is deterministic or estimated from the data. I focus on the latter case, which is the more complex one. Suppose that an IID sample of size N is available from $F(x, w)$. Let $\widehat{z}_F$ denote a consistent estimator of the poverty line z_F. I shall suppose that the estimator $\widehat{z}_F$ can be expressed asymptotically as

$$\widehat{z}_F = \frac{1}{m_{10}N}\sum_{i=1}^{N} w_i \zeta_F(x_i) + o_p(N^{-1/2}) \tag{3}$$

with $\zeta_F : \mathbb{R} \to \mathbb{R}$. Frequent choices for the poverty line are fractions of the mean income or (fractions of) quantiles of the distribution. In what follows, I show that the estimators of the mean- and quantile-based poverty lines can be written as in (3).

1. If the poverty line z_F is set to a fraction k of mean income, z_F equals $k\frac{\mu_{11}}{\mu_{10}}$ and is consistently estimated by

$$\widehat{z}_F = k\frac{m_{11}}{m_{10}}.$$

The estimator can be expressed as in (3) by letting

$$\zeta_F(x_i) = kx_i.$$

The limiting distribution of $\widehat{z}_F$ is obtained by a straightforward application of the delta method.

2. When the poverty line z_F is set to a fraction k of the p-quantile of the individual income distribution J_F, things are somewhat less obvious. The

p-quantile of J_F is defined as

$$Q_F(p) = \inf\{s | J_F(s) \geqslant p\}$$

and is estimated by

$$\widehat{Q}_F(p) = \inf\{s \,|\frac{1}{m_{10}N}\sum_{i=1}^{N} w_i I(x_i \leqslant s) \geqslant p\},$$

where $I(\cdot)$ is the indicator function. For univariate distributions, the Bahadur representation (Bahadur, 1966) allows the quantile estimator to be expressed as a sample average of IID variables. Ghosh (1971) provides a relatively simple proof. Along the lines of Ghosh, an asymptotic approximation can be derived for the quantile estimator $\widehat{Q}_F(p)$.

Lemma 1 *Assume that a random sample of size N is available from $F(x, w)$. If $J_F'(Q_F(p)) > 0$ and $\sup_i(w_i/\sum_{i=1}^N w_i) = o_p(N^{-1/2})$, then*

$$\widehat{Q}_F(p) - Q_F(p) = \frac{1}{\mu_{10} J_F'(Q_F(p))N}\sum_{i=1}^{N} w_i\{p - I(x_i \leqslant Q_F(p))\} + o_p(N^{-1/2}).$$

Proof. See Appendix A. ∎

Note that $\widehat{Q}_F(p) - Q_F(p)$ is expressed in Lemma 1 as a sample average of IID variables. Asymptotic normality of $N^{1/2}(\widehat{Q}_F(p) - Q_F(p))$ follows from the Lindeberg-Lévy central limit theorem and the asymptotic variance of $N^{1/2}(\widehat{Q}_F(p) - Q_F(p))$ is

$$\frac{1}{[\mu_{10} J_F'(Q_F(p))]^2}\,\mathrm{var}[w(p - I(x \leqslant Q_F(p)))].$$

To estimate the variance of $\widehat{Q}_F(p)$, an estimate of $J_F'(Q_F(p))$ is needed. Note that

$$J_F'(Q_F(p)) = \frac{E(w|Q_F(p))f_x(Q_F(p))}{\mu_{10}},$$

where f_x denotes the marginal density function of income x derived from $F(x,w)$. A consistent distribution-free estimate of $J'_F(Q_F(p))$ can be obtained using kernel estimation methods.

The quantile-based poverty line $z_F = kQ_F(p)$ is consistently estimated by $\widehat{z}_F = k\widehat{Q}_F(p)$. By Lemma 1 and because $m_{10} = \mu_{10} + o_p(1)$ and $\frac{1}{N}\sum_{i=1}^{N} w_i\{p - I(x_i \leqslant Q_F(p))\} = O_p(N^{-1/2})$,

$$\widehat{z}_F - z_F = \frac{1}{m_{10}N}\sum_{i=1}^{N} w_i \frac{k\{p - I(x_i \leqslant Q_F(p))\}}{J'_F(Q_F(p))} + o_p(N^{-1/2}),$$

which allows us to express $\widehat{z}_F$ as in (3) by letting

$$\zeta_F(x_i) = z_F + \frac{k\{p - I(x_i \leqslant Q_F(p))\}}{J'_F(Q_F(p))}. \tag{4}$$

Decomposable poverty measures

Let us now turn to the decomposable poverty measures. For the bivariate distribution $F(x,w)$, expression (1) becomes

$$P_F(z_F) = \frac{1}{\mu_{10}}\int_0^{z_F}\int w\pi(x, z_F)dF(x,w).$$

Suppose that an IID sample of size N is available from the distribution $F(x,w)$ and suppose that the poverty line is estimated from the sample and admits an asymptotic representation as in (3). The natural estimator of $P_F(z_F)$ is

$$\widehat{P}_F(\widehat{z}_F) = \frac{1}{m_{10}N}\sum_{i=1}^{N} w_i\pi(x_i, \widehat{z}_F)I(x_i \leqslant \widehat{z}_F). \tag{5}$$

Deriving the limiting distribution of $\widehat{P}_F(\widehat{z}_F)$ is nontrivial. The difficulty is that the summation in the numerator of (5) is truncated at a random income level. Zheng (2001b) derives the sampling distribution of the additive and separable poverty measures with a random poverty line. Weighting variables are not considered by Zheng, but his results can be adapted to our needs.

Theorem 1 *Assume that a random sample of size N is available from $F(x, w)$. Suppose that the poverty line is estimated from the sample and can be represented as in (3). Then*

$$\widehat{P}_F(\widehat{z}_F) - P_F(z_F) = \frac{1}{\mu_{10} N} \sum_{i=1}^{N} w_i(\delta_{F,i} - P_F(z_F)) + o_p(N^{-1/2})$$

with

$$\delta_{F,i} = \pi(x_i, z_F) I(x_i \leqslant z_F) + (\zeta_F(x_i) - z_F) \frac{\partial P_F(z_F)}{\partial z_F}.$$

Proof. See Appendix B. ■

Since $\widehat{P}_F(\widehat{z}_F) - P_F(z_F)$ is expressed in Theorem 1 as a sample average of IID variables, it is immediately clear that $N^{1/2}(\widehat{P}_F(\widehat{z}_F) - P_F(z_F))$ is asymptotically normal with mean zero and with asymptotic variance

$$\frac{1}{[\mu_{10}]^2} \operatorname{var}[w(\delta_F - P_F(z_F))]$$

with

$$\delta_F = \pi(x, z_F) I(x \leqslant z_F) + (\zeta_F(x) - z_F) \frac{\partial P_F(z_F)}{\partial z_F}.$$

Note that

$$\frac{\partial P_F(z_F)}{\partial z_F} = \frac{1}{\mu_{10}} \int_0^{z_F} \int w \frac{\partial \pi(x, z)}{\partial z}\bigg|_{z_F} dF(x, w) + \pi(z_F, z_F) J_F'(z_F).$$

All components of the variance can be consistently estimated, either by the sample analogue or by kernel estimation methods. Therefore, by Slutsky's theorem, the variance itself can be consistently estimated. Also note that the joint limiting distribution of k decomposable poverty measures is k-variate normal. The $k \times k$ asymptotic variance matrix is easily derived from the sample average representation in Theorem 1. Finally, if the poverty line is nonrandom, $\zeta_F(x_i) - z_F = 0$. In this case, the formulas simplify considerably.

Ordinates of dominance curves

I now turn to the estimation of the ordinates of stochastic dominance curves and their covariance structure. Primal and dual dominance curves are considered in turn.

Given the results obtained above, deriving the limiting distribution of the ordinates of *primal* dominance curves is straightforward. For the bivariate distribution $F(x,w)$, expression (2) can be used to write

$$D_F^s(\lambda z_F) = \frac{1}{(s-1)!\mu_{10}} \int_0^{\lambda z_F} \int w(\lambda z_F - x)^{s-1} dF(x,w),$$
$$\frac{D_F^s(\lambda z_F)}{[z_F]^{s-1}} = \frac{\lambda^{s-1}}{(s-1)!\mu_{10}} \int_0^{\lambda z_F} \int w \left(\frac{\lambda z_F - x}{\lambda z_F}\right)^{s-1} dF(x,w).$$

Note that the ordinates $D_F^s(\lambda z_F)$ and $\frac{D_F^s(\lambda z_F)}{[z_F]^{s-1}}$ are now expressed as decomposable poverty measures. So, the results obtained for the decomposable poverty measures can be used to estimate the ordinates and their covariance structure.

Besides the primal approach, there is the *dual* approach to stochastic dominance. The TIP curve cumulates poverty gaps. With a size variable, the definition becomes

$$TIP_F(p, z_F) = \frac{1}{\mu_{10}} \int_0^{Q_F(p)} \int w(z_F - x) I(x \leqslant z_F) dF(x,w).$$

Suppose again that we have N drawings from $F(x,w)$. Let the poverty line be estimated from the sample and assume that the estimate allows for a representation as in (3). The natural estimator of $TIP_F(p, z_F)$ is

$$\widehat{TIP}_F(p, \widehat{z}_F) = \frac{1}{m_{10}N} \sum_{i=1}^{N} w_i(\widehat{z}_F - x_i) I(x_i \leqslant \widehat{z}_F) I(x_i \leqslant \widehat{Q}_F(p)). \qquad (6)$$

The difficulty in deriving the limiting distribution of $\widehat{TIP}_F(p, \widehat{z}_F)$ is again the random truncation of the summation in the numerator of (6). Davidson and Duclos (2000) derive the covariance structure of TIP ordinates. They do not consider stochastic weighting variables, but their result can be easily extended. I use $(a)_+$ to denote $\max(0, a)$.

Theorem 2 *Assume that a random sample of size N is available from $F(x, w)$. Suppose that the poverty line is estimated from the sample and admits a representation as in (3). Then*

$$\widehat{TIP}_F(p, \widehat{z}_F) - TIP_F(p, z_F) = \frac{1}{\mu_{10}N} \sum_{i=1}^{N} w_i(\gamma_{F,i} - TIP_F(p, z_F)) + o_p(N^{-1/2})$$

with

$$\begin{aligned} \gamma_{F,i} &= p(z_F - Q_F(p))_+ \\ &\quad + I(x_i \leqslant Q_F(p))[(z_F - x_i)_+ - (z_F - Q_F(p))_+] \\ &\quad + (\zeta_F(x_i) - z_F) \min(J_F(z_F), p). \end{aligned}$$

Proof. See Appendix C. ■

By the same arguments as before, the joint limiting distribution of k TIP ordinates is k-variate normal. The $k \times k$ asymptotic variance matrix is easily derived from the sample average representation in Theorem 2 and can be consistently estimated.

The normalized TIP curve $TIP_F^n(p, z_F)$ is obtained by dividing the TIP curve by the poverty line. A simple adaptation of the proof in Appendix C results in

$$\widehat{TIP^n_F}(p,\widehat{z}_F) - TIP^n_F(p,z_F) = \frac{1}{\mu_{10} z_F N}\sum_{i=1}^{N} w_i(\gamma_{F,i} - \zeta_F(x_i)TIP^n_F(p,z_F)) + o_p(N^{-1/2}).$$

The joint limiting distribution of k normalized TIP ordinates is k-variate normal and the asymptotic variance matrix can be consistently estimated.

S*ample dependency*

Researchers are usually interested in comparing poverty between distributions, say $F(x,w)$ and $G(y,v)$. Suppose that a sample of size M is available from $F(x,w)$ and a sample of size N from $G(y,v)$. Let the statistics s and t be functions of the first and the second sample, respectively, and assume that the statistics can be written asymptotically as sample averages,

$$s = \frac{1}{M}\sum_{i=1}^{M} f(x_i,w_i) + o_p(N^{-1/2}),$$
$$t = \frac{1}{N}\sum_{i=1}^{N} g(y_i,v_i) + o_p(N^{-1/2}),$$

where $f : \mathbb{R}^2 \to \mathbb{R}$ and $g : \mathbb{R}^2 \to \mathbb{R}$ are square integrable with respect to F and G, respectively. For our purposes, this assumption is not restrictive, since the poverty indices and ordinates of the dominance curves all have such asymptotic representations. The joint limiting behavior of s and t depends on whether the samples are independent, fully dependent or partially dependent. Independent samples arise in the case of repeated cross-sectional sampling and imply, trivially, independence of s and t. On the other side of the spectrum, balanced panel data are fully dependent and are conceived as $M = N$ quadruples (x_i, w_i, y_i, v_i) drawn from the joint distribution $H(x,w,y,v)$ of income and household size at different times. It is assumed that $H(x, w, y, v)$ has marginals $F(x, w)$ and $G(y, v)$. In this case, the covariance between s and t follows immediately from the joint asymptotic sample average representations.

In many household panels, however, households rotate so that the samples of different years are partially dependent. Zheng (2001a) proposes a correction method for partial sample dependency. Zheng applies it to inequality measures, but the method is more generally applicable. Suppose that the first Q observations in the samples are matched. That is, (x_1, w_1, y_1, v_1), ..., (x_Q, w_Q, y_Q, v_Q) are independently drawn from $H(x, w, y, v)$. Further suppose that the remaining observations are independently drawn from the marginal distributions $F(x, w)$ and $G(y, v)$. That is, (x_{Q+1}, w_{Q+1}), ..., (x_M, w_M) is independent of (y_i, v_i) and (y_{Q+1}, v_{Q+1}), ..., (y_N, v_N) is independent of (x_i, w_i). It is also assumed that (x_{Q+1}, w_{Q+1}), ..., (x_M, w_M) is independent of (x_1, w_1), ..., (x_Q, w_Q) and that (y_{Q+1}, v_{Q+1}), ..., (y_N, v_N) is independent of (y_1, v_1), ..., (y_Q, v_Q).

Denote $\alpha_x = \frac{1}{M}\sum_{i=1}^{Q} f(x_i, w_i)$, $\beta_x = \frac{1}{M}\sum_{i=Q+1}^{M} f(x_i, w_i)$, $\alpha_y = \frac{1}{N}\sum_{i=1}^{Q} g(y_i, v_i)$ and $\beta_y = \frac{1}{N}\sum_{i=Q+1}^{N} g(y_i, v_i)$. The covariance between s and t can then be written as

$$\begin{aligned} \mathrm{cov}(s, t) &= \mathrm{cov}(\alpha_x + \beta_x, \alpha_y + \beta_y) \\ &= \mathrm{cov}(\alpha_x, \alpha_y) \\ &= \frac{Q}{M \times N} \mathrm{cov}(f(x, w), g(y, v)). \end{aligned}$$

Let $[S, T]' = E([s, t]')$ and let N, $M \to \infty$ along any expansion path. Assume that $Q/(MN)^{1/2} \to q$ with $0 \leqslant q \leqslant 1$. Then, $[M^{1/2}(s - S), N^{1/2}(t - T)]'$ is asymptotically normal with mean zero and asymptotic variance matrix

$$\begin{bmatrix} \mathrm{var}(f(x, w)) & q\,\mathrm{cov}(f(x, w), g(y, v)) \\ q\,\mathrm{cov}(f(x, w), g(y, v)) & \mathrm{var}(g(y, v)) \end{bmatrix}.$$

3.2 Hypothesis testing

The normality of the limiting distribution of the decomposable poverty indices implies that standard inference procedures can be used to test whether a given poverty index is higher in F than in G.

Testing for stochastic dominance is a more complicated matter. Many testing procedures have been proposed (see Davidson and Duclos, 2000 and Barrett and Donald, 2003 for an overview). Some check dominance by comparing the curves at a finite number of points. Others are based on the supremum or infimum of some statistic over the entire domain. The procedures I will discuss make use of a predetermined grid. The advantage of these testing procedures is their flexibility. Primal and dual stochastic dominance can be tested and the procedures can deal with dependent samples and a stochastic size variable.

Suppose the researcher wants to test whether the dominance curve associated with F is above the curve associated with G. Suppose that an IID sample of size M is available from F and an IID sample of size N from G and that N/M converges to a positive constant as $N \to \infty$. Assume that the researcher decides to compare the stochastic dominance curves at k grid points. Let m denote an estimate of μ, the k-dimensional vector of differences between the dominance curves. From the previous subsection, we know that $N^{1/2}(m - \mu)$ tends to a k-variate normal distribution. This is true irrespective of whether the samples are fully dependent, partially dependent or independent. Let S denote a consistent estimate of Σ, the asymptotic variance matrix of m. The vector m and the variance matrix S can be obtained by the estimation methods described in the previous subsection.

Various hypotheses could serve as the null or the alternative in a stochastic dominance test setting. Consider the hypotheses

$$H_0 : \mu = 0$$

$$H_1 : \mu \geqslant 0$$

$$H_2 : \mu \in \mathbb{R}^k$$

$$H_3 : \min(\mu) \leqslant 0$$

$$H_4 : \min(\mu) > 0$$

H_0 corresponds to the hypothesis that both dominance curves are identical, H_1 to (weak) dominance and H_4 to strict dominance.

Beach and Richmond (1985) and Bishop, Formby and Thistle (1992) have proposed a test procedure that has gained some popularity in the stochastic dominance literature. Their procedure constitutes a multiple comparison test. The k subhypotheses $H_{0,i} : \mu_i = 0$ are tested against $H_{A,i} : \mu_i \neq 0$ $(i = 1, 2, ..., k)$. The overall null, H_0, is the intersection of the k hypotheses $H_{0,i}$. If one of the null hypotheses is rejected, the overall null is rejected. The procedure is to compute the t-statistics for each of the k hypotheses $H_{0,i}$. Critical values are taken from the studentized maximum modulus (SMM) distribution. SMM(k, df) is the distribution of the maximum of the absolute value of k independent t-variates with df degrees of freedom. Critical values can be simulated or found in the tables provided by Stoline and Ury (1979). For a test with $k = 10$ and $df = \infty$, for example, the 5% critical value is 2.80.

Actually, the test based on the SMM distribution is a two-sided test of H_0 against H_2. When the null is rejected, however, Bishop, Formby and Thistle (1992) use the test statistics to draw further conclusions. If there is at least one significantly positive t-statistic and no significantly negative t-statistics, the conclusion is that the dominance curve associated with F is above the curve associated with G. If there is at least one significantly negative t-statistic and no significantly positive t-statistics, the conclusion is that the latter curve is above the former. Finally, if there are significantly positive and significantly negative t-statistics, it is concluded that the dominance curves cross.

The Bishop-Formby-Thistle (BFT) test procedure is easy to implement and attractive because of the encompassing decision rule. The procedure, however, does not exploit all the information available in the variance matrix. The elements of the vector m are treated as if they were independent. Yet, in general, the k t-statistics are highly correlated. The BFT test may therefore be expected to be undersized and to have low power.

In the statistical literature (e.g. Perlman, 1969), Wald tests based on distance statistics have been proposed that do make full use of the information in the variance matrix S. Gouriéroux, Holly and Monfort (1982), Kodde and Palm (1986) and Wolak (1989) introduced these test procedures in the econometrics literature. The initial application was to regression coefficients. Recently (Dardanoni and Forcina, 1999; Davidson and Duclos, 2000; Barrett and Donald, 2003), the tests have also been applied to stochastic dominance analysis. The results needed in this paper are contained in Wolak (1989). Wolak provides results that permit testing of H_0 against H_1 and of H_1 against H_2.

The critical step in the procedure is to solve the quadratic programming problem (QP):

$$\min_{\widetilde{m}} (\widetilde{m} - m)' S^{-1} (\widetilde{m} - m)$$

$$\text{subject to } \widetilde{m} \geqslant 0.$$

Let m^* represent the solution to this problem. We can then compute the distance statistics

$$d_{01} = N[m' S^{-1} m - (m - m^*)' S^{-1} (m - m^*)],$$

$$d_{12} = N(m - m^*)' S^{-1} (m - m^*).$$

d_{01} is the Wald statistic for the test of H_0 against H_1 and d_{12} for the test of H_1 against H_2. The asymptotic null distributions of these statistics turn out to be mixtures of chi-squared distributions. The following results (Wolak, 1989) permit calculation of the p-values for the test statistics d_{01} and d_{12}.

Wald test of H_0 against H_1

The asymptotic null distribution of d_{10} is such that

$\lim_{N\to\infty} \Pr(d_{01} \geqslant c) = \Sigma_{j=0}^{k} \Pr(\chi_j^2 \geqslant c) w_j(\Sigma)$.

Wald test of H_1 against H_2

The asymptotic null distribution of d_{12} is such that

$\lim_{N\to\infty} \sup_{\mu\geqslant 0} \Pr(d_{12} \geqslant c) = \Sigma_{j=0}^{k} \Pr(\chi_j^2 \geqslant c) w_{k-j}(\Sigma)$.

The weight $w_j(\Sigma)$ is the probability that the solution to the quadratic programming problem QP for a $N(0, \Sigma)$ variable m has exactly j positive elements. It is convenient to estimate the weights by the use of a Monte Carlo technique. One then draws from $N(0, S)$ and solves, for each of these draws, the quadratic programming problem QP. The weight $w_j(\Sigma)$ is estimated as the proportion of solutions to QP that have exactly j positive elements.

If one is not prepared to conclude that dominance holds unless there is strong evidence in favor of dominance, one may resort to a test developed by Sasabuchi (1980); see also Dardanoni and Forcina (1999). Sasabuchi's procedure tests the hypothesis H_3 that $\min(\mu) \leqslant 0$ against the alternative H_4 that $\min(\mu) > 0$. Denote by m_i the i-th element of m, by s_{ii} the (i, i)-th element of S and by z_α the $(1 - \alpha)$ quantile of the standard normal distribution. Under H_3, $\lim_{N\to\infty} \Pr[\inf(\frac{\sqrt{N} m_i}{\sqrt{s_{ii}}}) > z_\alpha]$ is less than or equal to α, while this limit is one under H_4. That is, the test that rejects H_3 against H_4 if

$$\frac{\sqrt{N} m_i}{\sqrt{s_{ii}}} \geqslant z_\alpha \qquad \text{for all } i = 1, .., k$$

has an upper bound α on its asymptotic size and is consistent. Not unexpectedly, the test has very low power if all the elements of μ are only slightly bigger than zero (Berger, 1989). But if the test rejects H_3, this may be considered as strong evidence in favor of H_4.

4 Empirical illustration

In this section, the inference methods are briefly applied to Belgian data. Some FGT poverty measures are computed and normalized TIP dominance is analyzed. The data set is the Belgian Socio-Economic Panel (SEP). The Centre for Social Policy of the University of Antwerp is responsible for collecting and processing the data. SEP contains data for the years 1985, 1988, 1992 and 1997. An independent data set with observations for 1976 was also made available by the Centre for Social Policy.

All data are monthly household income data. The 1976 data set contains 5098 income observations. For 1985, there are 6471 observations. Of these households, 3779 are also in the 1988 sample and 2900 are in the 1992 sample. In 1992, the sample was extended with the addition of 921 new households. In 1997, 4632 households were interviewed, 2375 of which were new households. The definition of disposable income that was used includes labor income, all social transfers and rental income and is comparable over the different waves. The 1976 data set contains no data on rental income. Official consumer price indices[1] were used to convert incomes to real 1996 euros.

A simple procedure is followed to reconstruct the individual income distribution. First, household income is adjusted by the EU equivalence scale. This scale gives the first adult (a person older than 18) a weight of 1, each subsequent adult 0.5 and each child 0.3. The equivalence scale adjusted household income is then weighted by the (non-weighted) number of household members.

[1] Available at http://mineco.fgov.be/homepull_en.htm.

	'76	'85	'88	'92	'97	'76-'85	'88-'85	'92-'88	'97-'92	'97-'85
P^0	0.0786	0.0505	0.0515	0.0606	0.0729	0.0282	0.0010	0.0091	0.0124	0.0225
s.e.	0.0035	0.0028	0.0042	0.0047	0.0044	0.0045	0.0047	0.0056	0.0064	0.0053
P^1	0.0157	0.0087	0.0076	0.0125	0.0117	0.0070	-0.0011	0.0049	-0.0008	0.0030
s.e.	0.0009	0.0007	0.0008	0.0018	0.0009	0.0011	0.0010	0.0019	0.0020	0.0012
P^2	0.0054	0.0029	0.0019	0.0062	0.0034	0.0026	-0.0009	0.0043	-0.0028	0.0006
s.e.	0.0005	0.0003	0.0003	0.0016	0.0005	0.0006	0.0004	0.0016	0.0017	0.0006

Table 1: FGT measures and standard errors with poverty line set to one-half of contemporary mean income.

The sample of 1976 should be self-weighting. The panels of 1985, 1988, 1992 and 1997 contain sample weights that account for the sampling framework and correct for panel attrition and non-response. Both the sampling weight and household size are random weighting variables. In my computations, the product of these variables was treated as a single stochastic weighting variable. Since the data set did not contain the necessary information on stratification and clustering, complex sample issues were neglected. I corrected for partial sample dependency. Throughout, the significance level of the statistical tests is 5%. The p-values of the Wald statistics were computed by running 1000 replications.

Identifying the poor amounts to the choice of a poverty line. Throughout, the poverty line is set to one-half of contemporary mean income. This choice gives the following chronological sequence of poverty lines for the different years: 494.91, 494.94, 516.88, 562.89 and 589.34. Table 1 gives the estimates and standard errors of the FGT measures for the parameter $a = 0$, 1 and 2. '$A - B$' means that the difference between the poverty index for year A and year B was computed. The headcount ratio P^0 declined from almost 8% in 1976 to about 5% in 1985. From 1985 on, the headcount started rising again to reach more than 7% in 1997. Notice that the difference in the value of the headcount is highly significant for the comparison between 1976 and 1985 and for that between 1985 and 1997. By and large, the poverty gap ratio P^1 displays a similar evolution. Again, there is a significant decline in moving from 1976 to 1985. The 1988 poverty gap ratio is lower than in 1985, but the difference is not significant at the 5% level. From 1988 onward, the poverty gap increases again. The increase from 1985 to 1997 is significant. The distribution-sensitive measure P^2 suggests a somewhat different evolution. From 1976 to 1988, there is a significant decline. The P^2 measure jumps upwards in moving from 1988 to 1992 and declines again afterwards. The difference between 1985 and 1997 is not significantly different from zero.

The P^2 measure suggests that the income distribution among the poor was very unequal in 1992. This is confirmed by Figure 1. Figure 1 plots the normalized TIP curves for the different years. The 1992 TIP curve is indeed very concave. To test whether the TIP curves dominate each other, the test

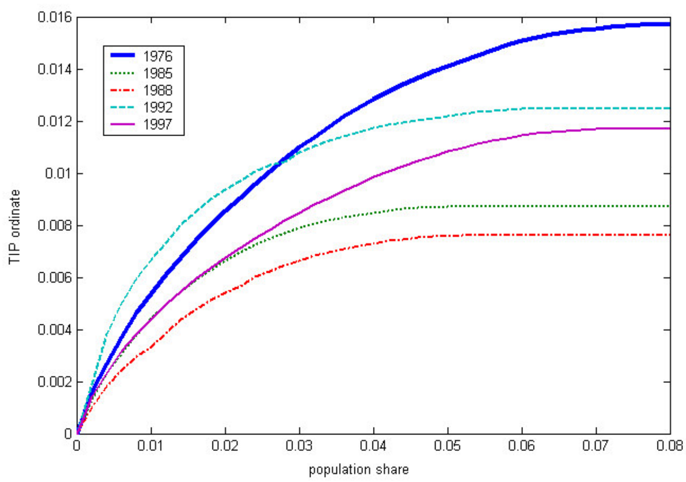

Figure 1: Normalized TIP curves with poverty line set to one-half of contemporary mean income.

procedures described above can be used. I choose to compute the difference between the TIP curves at fractions 0.1, 0.2, ..., 1 of the maximal headcount

	t-statistics				
grid points	'76>'85	'85>'88	'92>'88	'92>'97	'97>'85
0.1	2.58	1.92	3.20	2.27	0.18
0.2	3.03	2.38	3.07	2.04	-0.07
0.3	3.89	2.26	2.85	1.74	0.27
0.4	4.61	2.07	2.72	1.49	0.74
0.5	5.31	1.87	2.65	1.24	1.23
0.6	5.79	1.69	2.60	0.94	1.82
0.7	6.34	1.49	2.54	0.71	2.22
0.8	6.49	1.33	2.55	0.56	2.59
0.9	6.47	1.22	2.63	0.42	2.65
1	6.30	1.13	2.60	0.38	2.60
	p-values in %				
H_0-H_1 (Wald test)	0.00	5.62	0.01	0.19	
H_1-H_2 (Wald test)					61.6
H_3-H_4 (Sasabuchi)	0.49	12.92	0.55	35.19	52.79

Table 2: t-statistics and test results for TIP dominance.

ratio of the two distributions that are to be compared. The test results are in Table 2. '$A > B$' means that the testing procedures check whether A TIP dominates B. The t-statistics are the test statistics for the hypothesis that the difference between the normalized TIP curves at the grid point equals zero. If all t-statistics are found to be positive, the p-value of the test of H_1 against H_2 is a priori known to be 1. In these cases, the p-value of the test of H_0 against H_1 is reported instead.

Sasabuchi's testing procedure rejects non-dominance in favor of strict dominance for the comparison between the 1976 and the 1985 curves and for the comparison between the 1992 and the 1988 curves. The Wald test of H_0 against H_1 does not reject the equality of the 1985 and 1988 curves, but rejects the equality of the 1992 and 1997 curves in favor of TIP dominance of 1992 over 1997. The TIP curves of 1985 and 1997 cross. The crossing, however, does not lead the Wald test to reject the hypothesis that the 1997 TIP curve dominates the 1985 TIP curve.

The general pattern displayed by the poverty indices is thus confirmed by the TIP dominance analysis. Relative poverty decreased from 1976 to the mid-eighties and increased again thereafter. The TIP dominance analysis also indicates that poverty was higher in 1992 than in 1997 according to all relative poverty measures satisfying the transfer principle.

5 Conclusion

Inference for poverty measurement often involves a stochastic weighting variable. This situation occurs naturally when the observations on income are at the household level and when the researcher's interest is in the distribution of income among individuals. In empirical work, household incomes are typically rescaled using an equivalence scale and are then weighted by the number of household members. The consideration of a bivariate distribution defined over income and weight allowed us to take the randomness of the weighting variable into account. The limiting distributions of the decomposable poverty measures and of the ordinates of poverty dominance curves were derived within the bivariate framework. The poverty line was allowed to depend on the income distribution, with special attention to mean- and quantile-based poverty lines. It was then shown how the results on the limiting distributions of the estimators can be used to test changes in the value of a poverty index and to test poverty dominance. Belgian data were used to illustrate the inference procedures.

The results in this paper can be extended along several lines. First, results were derived only for poverty dominance curves. It is not difficult to derive similar results for welfare and inequality dominance curves such as (generalized) Lorenz curves. The same procedures can then be applied to test welfare and inequality dominance. Second, this paper does not touch upon complex sample issues like stratification and clustering. Extending the results to the complex sample case first requires that an appropriate asymptotic framework be specified. The limiting distributions of the estimators then have to be derived within the chosen framework. Under the assumption that the number of clusters tends to infinity, the usual complex sample variance formulas can be used for the estimators considered in this paper (see e.g. Zheng, 2001b; Zheng, 2002). Third, I have focused on test procedures for stochastic dominance that make use of a predetermined grid. The arbitrariness involved in the grid specification is not desirable from an inference perspective and constitutes an important drawback of these procedures. Recently, Barrett and Donald (2003) have proposed a Kolmogorov-Smirnov type test for stochastic dominance which effectively considers the entire domain of the dominance curves. The test requires the samples to be randomly drawn from univariate income distributions. It would be of relevance to extend the testing framework so that stochastic weights and complex sample design can be dealt with. This is an issue for future research.

Appendices

Appendix A: Proof of Lemma 1

The proof parallels that of Ghosh (1971, p. 1958-1959). Define

$$J_F(y) = \frac{1}{\mu_{10}} \int_0^y \int w dF(x, w)$$

and

$$\widehat{J}_F(y) = \frac{1}{m_{10} N} \sum_{i=1}^{N} w_i I(x_i \leqslant y).$$

Let

$$t_N = N^{1/2}(\widehat{Q}_F(p) - Q_F(p)).$$

We then have for any c

$$t_N \leqslant c \Longleftrightarrow u_N \leqslant c_N$$

where

$$\begin{aligned} u_N &= N^{1/2}\{J_F(Q_F(p) + cN^{-1/2}) - \widehat{J}_F(Q_F(p) + cN^{-1/2})\}(J'_F(Q_F(p)))^{-1}, \\ c_N &= N^{1/2}\{J_F(Q_F(p) + cN^{-1/2}) - \widehat{J}_F(\widehat{Q}_F(p))\}(J'_F(Q_F(p)))^{-1}. \end{aligned}$$

Now, a Taylor expansion gives

$$J_F(Q_F(p) + cN^{-1/2}) = p + cN^{-1/2} J'_F(Q_F(p)) + o(N^{-1/2}). \tag{7}$$

Under the condition that $\sup_i(w_i / \sum_{i=1}^N w_i) = o_p(N^{-1/2})$, we also have

$$\widehat{J}_F(\widehat{Q}_F(p)) = p + o_p(N^{-1/2}). \tag{8}$$

From (7) and (8),

$$c_N \rightarrow_p c.$$

Let

$$v_N = N^{1/2}\{J_F(Q_F(p)) - \widehat{J}_F(Q_F(p))\}(J'_F(Q_F(p)))^{-1}.$$

Then

$$\begin{aligned} u_N - v_N = N^{1/2}\{&J_F(Q_F(p) + cN^{-1/2}) - J_F(Q_F(p)) \\ &-\widehat{J}_F(Q_F(p) + cN^{-1/2}) + \widehat{J}_F(Q_F(p))\}(J'_F(Q_F(p)))^{-1}. \end{aligned}$$

Because $cN^{-1/2} = o(1)$ and $J_F(y) - \widehat{J}_F(y) = O_p(N^{-1/2})$ for any y,

$$J_F(Q_F(p)+cN^{-1/2}) - J_F(Q_F(p)) - \widehat{J}_F(Q_F(p)+cN^{-1/2}) + \widehat{J}_F(Q_F(p)) = o_p(N^{-1/2})$$

so that

$$u_N - v_N \to_p 0.$$

Because $c_N \to_p c$ and $u_N - v_N \to_p 0$, we have for every $\varepsilon > 0$

$$\begin{aligned} \lim_{N\to\infty} \Pr(t_N \leqslant c, v_N \geqslant c + \varepsilon) &= \lim_{N\to\infty} \Pr(u_N \leqslant c_N, v_N \geqslant c + \varepsilon) = 0, \\ \lim_{N\to\infty} \Pr(t_N \geqslant c + \varepsilon, v_N \leqslant c) &= \lim_{N\to\infty} \Pr(u_N \geqslant c_N + \varepsilon, v_N \leqslant c) = 0. \end{aligned}$$

This establishes the second condition of Lemma 1 of Ghosh (1971). The first condition, that v_N is bounded in probability, is also satisfied. By Lemma 1 of Ghosh, $t_N - v_N = o_p(1)$. So,

$$\begin{aligned} \widehat{Q}_F(p) - Q_F(p) &= \frac{p - \frac{1}{m_{10}N}\sum_{i=1}^{N} w_i I(x_i \leqslant Q_F(p))}{J'_F(Q_F(p))} + o_p(N^{-1/2}) \\ &= \frac{\frac{1}{N}\sum_{i=1}^{N} w_i\{p - I(x_i \leqslant Q_F(p))\}}{\mu_{10} J'_F(Q_F(p))} + o_p(N^{-1/2}), \end{aligned}$$

where μ_{10} could be substituted for m_{10} because $m_{10} = \mu_{10} + o_p(1)$ and

$\frac{1}{N}\sum_{i=1}^{N} w_i\{p - I(x_i \leqslant Q_F(p))\} = O_p(N^{-1/2})$.

Appendix B: Proof of Theorem 1

The proof parallels that of Zheng (2001b, p. 341-342). Write $\widehat{P}_F(\widehat{z}_F)$ as $\frac{p_F}{m_{10}}$ with

$$\begin{aligned}
p_F &= \int_0^{\widehat{z}_F}\int w\pi(x,\widehat{z}_F)d\widehat{F}(x,w) \\
&= \int_0^{z_F}\int w\pi(x,\widehat{z}_F)d\widehat{F}(x,w) + \int_{z_F}^{\widehat{z}_F}\int w\pi(x,\widehat{z}_F)d\widehat{F}(x,w) \\
&= (i) + (ii),
\end{aligned} \tag{9}$$

say, where $\widehat{F}(x,w)$ is the empirical distribution function. After a Taylor expansion of $\pi(x,\widehat{z}_F)$, the first integral becomes

$$\begin{aligned}
(i) \quad = \quad & \int_0^{z_F}\int w\pi(x,z_F)d\widehat{F}(x,w) \\
& + (\widehat{z}_F - z_F)\int_0^{z_F}\int w\left.\frac{\partial\pi(x,z)}{\partial z}\right|_{z_F} d\widehat{F}(x,w) + o_p(N^{-1/2}) \\
= \quad & \int_0^{z_F}\int w\pi(x,z_F)d\widehat{F}(x,w) \\
& + (\widehat{z}_F - z_F)\int_0^{z_F}\int w\left.\frac{\partial\pi(x,z)}{\partial z}\right|_{z_F} dF(x,w) + o_p(N^{-1/2}),
\end{aligned}$$

because $\widehat{z}_F - z_F$ and $\widehat{F} - F$ are $O_p(N^{-1/2})$. For the same reason,

$$\begin{aligned}
(ii) &= \int_{z_F}^{\widehat{z}_F}\int w\pi(x,\widehat{z}_F)dF(x,w) + O_p(N^{-1}) \\
&= (\widehat{z}_F - z_F)\pi(z_F,z_F)f_x(z_F)E(w|z_F) \\
&\quad +(\widehat{z}_F - z_F)\int_{z_F}^{\widehat{z}_F}\int w\left.\frac{\partial\pi(x,z_F)}{\partial z}\right|_{z_F} dF(x,w) + o_p(N^{-1/2}) \\
&= (\widehat{z}_F - z_F)\pi(z_F,z_F)f_x(z_F)E(w|z_F) + o_p(N^{-1/2}),
\end{aligned}$$

after a Taylor expansion and neglecting terms that are $o_p(N^{-1/2})$. Combining these results, we get

$$\begin{aligned} p_F &= \frac{1}{N}\sum_{i=1}^{N} w_i \pi(x_i, z_F) I(x_i \leqslant z_F) \\ &+ (\widehat{z}_F - z_F)\left\{ \int_0^{z_F}\int w \frac{\partial \pi(x,z)}{\partial z}\bigg|_{z_F} dF(x,w) + \pi(z_F, z_F) f_x(z_F) E(w|z_F) \right\} \\ &+ o_p(N^{-1/2}). \end{aligned} \tag{10}$$

From (3), we have that

$$\widehat{z}_F - z_F = \frac{1}{m_{10}N}\sum_{i=1}^{N} w_i(\zeta_F(x_i) - z_F) + o_p(N^{-1/2}).$$

Since $m_{10} \to_p \mu_{10}$ and $\frac{1}{N}\sum_{i=1}^{N} w_i(\zeta_F(x_i) - z_F) = O_p(N^{-1/2})$,

$$\widehat{z}_F - z_F = \frac{1}{\mu_{10}N}\sum_{i=1}^{N} w_i(\zeta_F(x_i) - z_F) + o_p(N^{-1/2}). \tag{11}$$

Substituting this result into (10), it follows that

$$\begin{aligned} \widehat{P}_F(\widehat{z}_F) - P_F(z_F) = \frac{1}{m_{10}N}\sum_{i=1}^{N} & w_i\{\pi(x_i, z_F) I(x_i \leqslant z_F) \\ & + (\zeta_F(x_i) - z_F)\frac{\partial P_F(z_F)}{\partial z_F} - P_F(z_F)\} + o_p(N^{-1/2}). \end{aligned}$$

By a similar argument as above, μ_{10} may now be substituted for m_{10}.

Appendix C: Proof of Theorem 2

The proof parallels that of Davidson and Duclos (2000, p. 1461). Write

$\widehat{TIP}_F(p, \widehat{z}_F)$ as $\frac{t_F}{m_{10}}$ with

$$\begin{aligned} t_F &= \int\int w(z_F - x)I(x \leqslant \widehat{Q}_F(p))I(x \leqslant \widehat{z}_F)d\widehat{F}(x,w) \\ &\quad + (\widehat{z}_F - z_F)\int\int wI(x \leqslant \widehat{Q}_F(p))I(x \leqslant \widehat{z}_F)d\widehat{F}(x,w) \\ &= (i) + (ii), \end{aligned} \tag{12}$$

say. Using the same techniques as above, it can be shown that

$$\begin{aligned} (i) &= \int\int w(z_F - x)I(x \leqslant Q_F(p))I(x \leqslant z_F)d\widehat{F}(x,w) \\ &\quad + (\widehat{Q}_F(p) - Q_F(p))(z_F - Q_F(p))I(Q_F(p) \leqslant z_F)E(w|Q_F(p))f_x(Q_F(p)) \\ &\quad + o_p(N^{-1/2}) \\ &= \frac{1}{N}\sum_{i=1}^{N} w_i(z_F - x_i)I(x_i \leqslant Q_F(p))I(x_i \leqslant z_F) \\ &\quad + \frac{1}{N}\sum_{i=1}^{N} w_i(p - I(x_i \leqslant Q_F(p)))(z_F - Q_F(p))I(Q_F(p) \leqslant z_F) \\ &\quad + o_p(N^{-1/2}), \end{aligned}$$

after substitution of the Bahadur representation into the second term.
The second term in (12) is

$$\begin{aligned} (ii) &= (\widehat{z}_F - z_F)\min\left(\int_0^{z_F}\int wdF(x,w), \int_0^{Q_F(p)}\int wdF(x,w)\right) + O_p(N^{-1}) \\ &= \mu_{10}(\widehat{z}_F - z_F)\min(J_F(z_F), p) + O_p(N^{-1}). \end{aligned}$$

Substituting these results and (11) into (12), it follows upon rearranging that

$$\widehat{TIP}_F(p,\widehat{z}_F) - TIP_F(p,z_F) = \frac{1}{m_{10}N}\sum w_i\{p(z_F - Q_F(p))_+ + I(x_i \leqslant Q_F(p)) \times [(z_F - x_i)_+ - (z_F - Q_F(p))_+] + (\zeta_F(x_i) - z_F)\min(J_F(z_F),p) - TIP_F(p,z_F)\} + o_p(N^{-1/2}).$$

where $(a)_+$ denotes $\max(0,a)$. Again, m_{10} can be replaced by μ_{10} .

References

Atkinson, A.B. (1987), On the measurement of poverty, *Econometrica*, 55, 749-764.

Bahadur (1966), A note on quantiles in large samples, *Annals of Mathematical Statistics*, 37, 577-580.

Barrett, G. and S. Donald (2003), Consistent tests for stochastic dominance, *Econometrica*, 71, 71-104 .

Beach, C.M., & J. Richmond (1985), Joint confidence intervals for income shares and Lorenz curves, *International Economic Review*, 26, 439-450.

Berger, R. (1989), Uniformly more powerful tests for hypotheses concerning linear inequalities and normal means, *Journal of the American Statistical Association*, 84, 192-199.

Bishop, J.A., J.P. Formby and P. Thistle (1992), Convergence of South and Non-South income distributions, 1969-1979, *The American Economic Review*, 82, 262-272.

Clark, S., R. Hemming and D. Ulph (1981), On indices for the measurement of poverty, *Economic Journal*, 30, 348-361.

Cowell, F. (1989), Sampling variance and decomposable inequality measures, *Journal of Econometrics*, 42, 27-41.

Dardanoni, V. and A. Forcina (1999), Inference for Lorenz orderings, *Econometrics Journal*, 2, 49-75.

Davidson, R. and J.-Y. Duclos (2000), Statistical inference for stochastic dominance and for the measurement of poverty and inequality, *Econometrica*, 68, 1435-1464.

Duclos, J.-Y. and P. Makdissi (2004), Restricted and unrestricted dominance for welfare, inequality and poverty orderings, *Journal of Public Economic Theory*, 6, 145-164.

Fishburn, P.C. (1976), Continua of stochastic dominance relations for bounded probability distributions, *Journal of Mathematical Economics*, 3, 295-311.

Foster, J.E., J. Greer and E. Thorbecke (1984), A class of decomposable poverty measures, *Econometrica*, 52, 761-776.

Ghosh, J.K. (1971), A new proof of the Bahadur representation of quantiles and an application, *Annals of Mathematical Statistics*, 42, 1957-1961.

Gouriéroux, C. and A. Holly and A. Monfort (1982), Likelihood ratio test, Wald test, and Kuhn-Tucker test in linear models with inequality constraints on the regression, *Econometrica*, 50, 63-80.

Jenkins, S. and P. Lambert (1997), Three I's of poverty curves, with an analysis of UK poverty trends, *Oxford Economic Papers*, 49, 317-327.

Jenkins, S. and P. Lambert (1998a), Ranking poverty gap distributions: further TIPs for poverty analysis, *Research on Economic Inequality*, 8, 31-38.

Jenkins, S. and P. Lambert (1998b), Three I's of poverty curves and poverty dominance: TIPs for poverty analysis, *Research on Economic Inequality*, 8, 39-56.

Kodde, D. and F. Palm (1986), Wald criteria for jointly testing equality and inequality restrictions, *Econometrica*, 54, 1243-1248.

Perlman, M.D. (1969), One-sided testing problems in multivariate analysis, *Annals of Statistics*, 40, 549-567.

Sasabuchi, S. (1980), A test of multivariate normal mean with composite hypotheses determined by linear inequalities, *Biometrika*, 67, 429-439.

Spencer, B. and S. Fisher (1992), On comparing distributions of poverty gaps, *Sankhya: The Indian Journal of Statistics. Series B*, 54, 114-126.

Stoline, M.R., & Ury, H.K. (1979), Tables of the studentized maximum modulus distribution, *Technometrics*, 21, 87-93.

Watts, H. (1968), An economic definition of poverty, in D.P. Moynihan (ed.), *On Understanding Poverty*, New York, Basic Books, 316-325.

Wolak, F.A. (1989), Testing inequality constraints in linear econometric models, *Journal of Econometrics*, 41, 205-235.

Zheng, B. (2000), Poverty orderings, *Journal of Economic Surveys*, 14, 427-466.

Zheng, B. (2001a), Statistical inference for testing inequality indices with dependent samples, *Journal of Econometrics*, 101, 315-335.

Zheng, B. (2001b), Statistical inference for poverty measures with relative poverty lines, *Journal of Econometrics*, 101, 337-356.