Consistent Tests for Poverty Dominance Relations

Garry F. Barrett*

Stephen G. Donald**

Yu-Chin Hsu†

This version: June 15, 2015

* E-mail: garry.barrett@sydney.edu.au. School of Economics, University of Sydney, Sydney, NSW, 2006 Australia.

** Corresponding author. E-mail:donald@eco.utexas.edu. Department of Economics, University of Texas at Austin, Austin TX, 78712 U.S.A.

† E-mail: ychsu@econ.sinica.edu.tw. Institute of Economics, Academia Sinica, Taipei, 115 Taiwan.

‡ Acknowledgement: We thank Esfandiar Massoumi and two referees for helpful comments. All errors and omissions are our own responsibility.

Abstract

This paper considers methods for comparing poverty in two income distributions. We first discuss the concept and usefulness of the Poverty Gap Profile (PGP) for comparing poverty in two populations. Dominance of one PGP over another suggests poverty dominance for a wide class of indices which may be expressed as functionals of the PGP. We then discuss hypotheses that can be used to test poverty dominance in terms of the PGP and introduce and justify a test statistic based on empirical PGP's where we allow for the poverty line to be estimated. A method for obtaining critical values by simulation is proposed that takes account of estimation of the poverty line. The finite sample properties of the methods are examined in the context of a Monte Carlo simulation study and the methods are illustrated in an assessment of relative consumption poverty in Australian over the period 1988/89-2009/10.

JEL classification: C01, C12, C21

Keywords: Poverty gap profile, poverty gap profile dominance, hypothesis testing, poverty line.

1 Introduction

Since the pioneering work of Sen (1976), research in poverty measurement has sought to develop measures which take into account the incidence, depth and distributional aspects of poverty. More recent research has focused on developing methods for making poverty comparisons which are robust to the normative properties of a specific poverty index. This line of research has culminated in the contributions of Atkinson (1987), Shorrocks (1995) and Jenkins and Lambert (1997) on poverty quasi-orderings based on the distribution of poverty gaps or 'poverty shortfalls'.

Poverty Gap Profiles (PGP) have key properties which enable them to play a central role in poverty analysis, analogous to the role of Lorenz curves in inequality measurement. First, the PGP is an intuitive graphical device for illustrating the three fundamental aspects of poverty evident in an income distribution. Second, the normative criteria incorporated in the PGP dominance quasi-ordering (focus, monotonicity, S-concavity) are widely accepted as a minimal set of properties desired of normative poverty measures. The normative properties of the PGP are transparent and directly related to the stochastic dominance (SD) criteria used in the welfare and inequality measurement literature (Ravallion, 1994). Further, poverty comparisons based on poverty gap dominance relations are robust to the additional normative properties embodied in a specific poverty index, and to the common scaling of poverty lines. Although the poverty gap dominance criteria provide a partial ordering of distributions, many popular poverty indices can be expressed as functionals of PGPs and can be used to generate a complete, cardinal ordering of distributions.

In this paper we develop consistent tests for poverty gap dominance relations that compare two estimated PGP's for two independent samples of individual or household income.[1] The proposed test is Cramer-von Mises type test statistic based on an integral of the positive difference between two empirical PGPs. An obvious alternative would have been to use Kolmogorov-Smirnov type tests as has been used for tests of SD proposed in McFadden (1989) and elaborated and extended by Barrett and Donald (2003) and Linton, Maasoumi and Whang (2005). Although it appears that such a test seems to

[1] The methods can also be applied to the distribution of earnings, consumption or wealth. We use the term income as synonymous for any measure of economic wellbeing.

work in practice, based on simulations, it is more difficult to properly justify its theoretical properties since there is a non-differentiability in the PGP at the poverty line. When the poverty line is estimated, as in our case, this makes it difficult to take into account the effect of this on the limiting distribution.

The advantage of testing PGP dominance over SD is that PGP dominance directly tests the hypothesis - poverty dominance - of interest. As an inferential problem, the proposed PGP test, unlike SD testing, permits the poverty lines to differ across distributions and, importantly, to be sample dependent. Further, the results could also be used to derive statistical properties of poverty indices that are defined as functionals of the PGP. The results provide the foundation for empirical poverty comparisons based on PGPs within a framework of formal statistical inference.

In the next section of the paper the PGP is defined. Section 3 states the hypotheses of interest which relate to dominance relations between two PGPs defined on two populations. In Section 4 we define the empirical version of the PGP and Cramer-von Mises type test and derive the properties of tests based on the test statistic with simulated critical values. Section 5 shows how one can adapt the results to allow inference based on the normalized PGP. Section 6 provides a small scale Monte Carlo study which examines how well the asymptotic arguments work in small samples. The inference procedures for PGP dominance relations are illustrated in Section 7 with an application to Australian consumption data and Section 8 concludes.

2 Poverty Gap Profiles

In this section we consider the definition and properties of the PGP. To derive the PGP let income X be distributed with distribution function $F(x) = \Pr(X \leq x)$. Let z denote a poverty line which may be known or possibly unknown. In the case of an unknown z we will assume that it is estimated as a function of an estimated sample quantile (e.g. median) or moment (e.g. mean). Given a poverty line z we can define the population version of the "head-count ratio" as $F(z) = \Pr(X \leq z)$ which is simply the proportion of the population with incomes below the poverty line. Although the empirical version of the head-count ratio is popular, this summary measure of poverty has been widely criticized since it captures only the incidence of poverty and ignores the depth of poverty and the inequality in incomes among the poor (Sen, 1976). Shorrocks (1995, 1998) and Jenkins and Lambert (1997, 1998a, 1998b) suggested the use of the PGP, and indices based on this curve, as a general approach to obtaining poverty dominance orderings which are sensitive to these three aspects of poverty measurement. The PGP is also known by a variety of names, such as the "deprivation profile" of Shorrocks (1998), the "Three I's of Poverty (TIP)" curve by Jenkins and Lambert (1997, 1998a, 1998b),[2] and is dual to the poverty deficit curve introduced by Atkinson (1987). To define the PGP let

$$D(z, X) = \max\{z - X, 0\} = (z - X) \cdot 1(X \leq z) \tag{1}$$

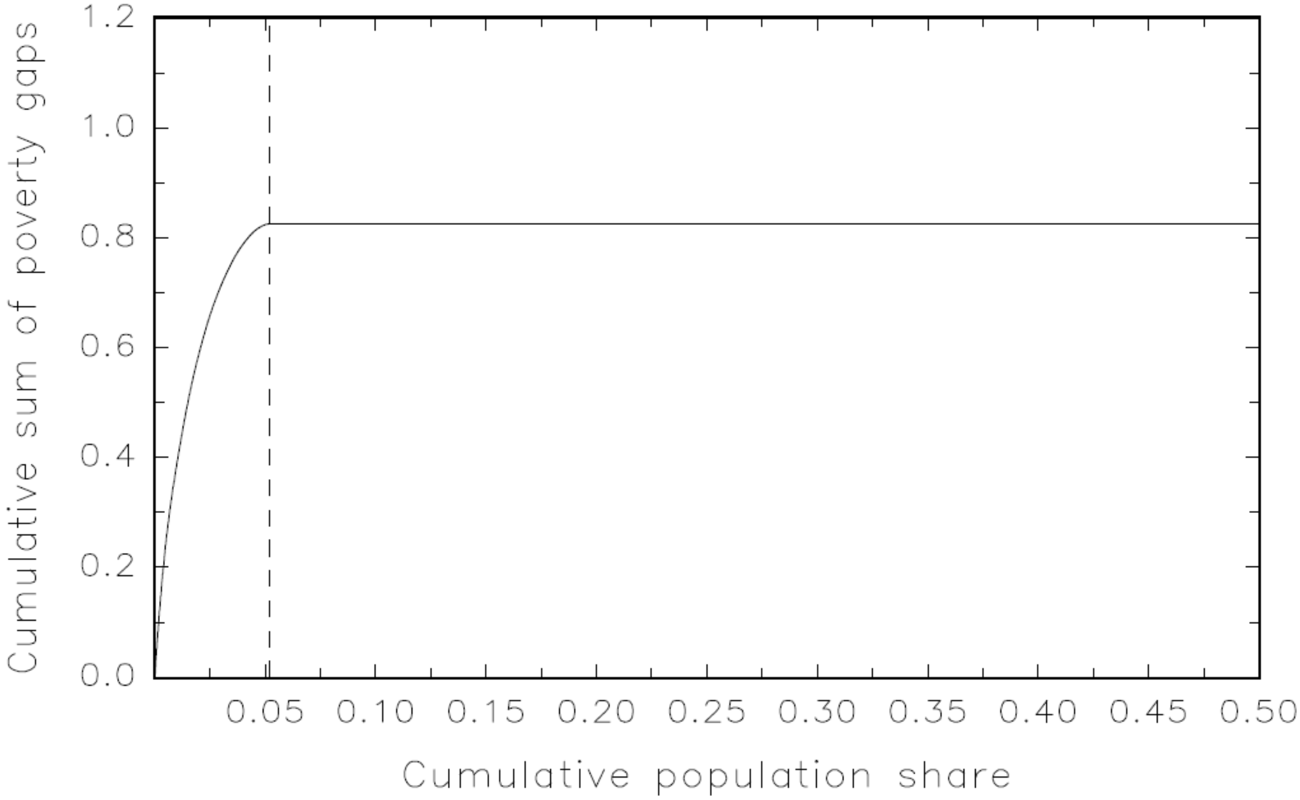
be the poverty gap (or 'income deficit') for the randomly drawn income X where $1(\cdot)$ denotes indicator function. This gives the difference between the income of an individual and the poverty line and is zero whenever an individual has an income greater than the poverty line. Further, let $Q(p)$ be the pth quantile of income so that by construction $F(Q(p)) = p$. The PGP P is then simply represented as

$$\begin{aligned} P(p; z) &= E(D(z, X) \cdot 1(X \leq Q(p))) \\ &= 1(Q(p) \leq z) \cdot (pz - G(p)) + 1(Q(p) > z) \cdot (F(z)z - G(F(z))) \end{aligned} \tag{2}$$

where $G(p) = E(X \cdot 1(X \leq Q(p)))$ is the Generalized Lorenz curve, or the mean income for the poorest $100p\%$ of the population. The curve P gives the average poverty gap for the poorest $100p\%$ of the population whenever p is a value that is below the head-count ratio. For values of p above the head-count ratio, the poverty gap profile gives the average poverty gap for the population. The expression in (2) shows that the PGP can be expressed as the difference between the poverty line z scaled by the cumulative population share p (zp maps the line of maximal poverty) and the Generalized Lorenz ordinate $G(p)$ (which is cumulative mean income scaled by cumulative population share p) over the poor segment of the population, and is equal to the population mean poverty gap $(F(z)z - G(F(z)))$ for all p at and above the poverty line. This expression demonstrates the duality between the PGP and the Generalized Lorenz curve defined over the poor segment of the population. Equivalently, the PGP is dual to the Generalized Lorenz curve

[2] In addition Yitzhaki (1999) refereed to the curve as the 'absolute rotated Lorenz curve' and, in earlier work on wage discrimination, Jenkins (1991, 1994) labeled the curve the 'inverse generalized Lorenz curve'.

Figure 1: Poverty Gap Profile



for the censored income distribution $\min\{X, z\}$. The relationship between the PGP and the Generalized Lorenz curve highlights that the PGP is a useful graphical device for depicting key dimensions of poverty for a given income distribution.

As has been shown by Shorrocks (1995) and Jenkins and Lambert (1997) the PGP curve captures the three fundamental elements of poverty: the point at which the curve levels out is the head-count ratio, the height of the profile at the head-count ratio gives the average poverty gap (or mean income deficit for the full population) and the degree of concavity of the curve indicates the degree of income inequality among the poor segment of the population. Figure 1 illustrates a typical PGP.

The curve is of interest in its own right, and many popular poverty indices can be expressed as functionals of the PGP. Further, it has been shown that if the PGP for one distribution dominates that for another, then all poverty indices which satisfy a set of basic properties will indicate that there is less poverty in the dominated distribution for all values of the poverty line up to z. Our results concerning the empirical PGP could be used to derive statistical properties of empirical versions of such poverty indices.

The development of the PGP in (2) is based on absolute poverty gaps. A related approach is to consider relative poverty gaps whereby the poverty deficit is normalized by the poverty line $D^R(z, X) = \max\{\frac{z-X}{z}, 0\} = (\frac{z-X}{z}) \cdot 1(X \le z)$. Consequently, the "normalized PGP" is simply equal to the absolute PGP of (2) scaled by $\frac{1}{z}$:

$$P^R(p; z) = \frac{P(p; z)}{z} = 1(Q(p) \le z) \cdot \left(p - \frac{G(p)}{z}\right) + 1(Q(p) > z) \cdot \left(F(z) - \frac{G(F(z))}{z}\right) \quad (3)$$

Below we first discuss tests of PGP dominance using the PGP based on absolute poverty gaps with an estimated z and then discuss how this approach can be extended to tests based on P^R.[3]

3 Tests of PGP Dominance

We are interested in testing whether there is a dominance relationship between two PGPs based on two distributions. We use subscripts 1 and 2 on the various curves and poverty lines to distinguish them apart. Thus for instance the two income distributions are F_1 and F_2, the two poverty levels are z_1 and z_2 and so on. With this notation in hand we can state the hypothesis of interest as follows,

$$H_0^1 : P_2(p; z_2) \le P_1(p; z_1) \text{ for all } p \in [0, 1].$$
$$H_1^1 : P_2(p; z_2) > P_1(p; z_1) \text{ for some } p \in [0, 1]. \quad (4)$$

The null hypothesis is that the P_1 is everywhere at least as large as P_2. This will be referred to as Weak PGP Dominance of P_1 over P_2. As shown by Jenkins and Lambert (1997), an implication of this is that poverty will be ranked as more severe in F_1 than in F_2 for a wide class of poverty indices. The way that we have set up these hypotheses is consistent with much of the recent literature on testing stochastic dominance (see McFadden (1989), Davidson and Duclos (2000), Barrett and Donald (2003), Linton, Maasoumi

[3] This approach to poverty measurement is based on a "relative" perspective where the poverty threshold is not defined as an absolute standard applicable across distributions but is relative to properties of the given distribution.

and Whang (2005)) and Lorenz dominance relations (Dardanoni and Forcina (1999), Barrett, Donald and Bhattacharya (2014)). Here the P_1 lies above (strictly, nowhere below) P_2 and in that sense is (weakly) dominant. From a social welfare perspective, the P_1 is closer to the line of maximal poverty than is P_2, and thereby implying greater poverty in F_1 compared to poverty in F_2. Note that the null hypothesis also includes the case where the PGP curves coincide. This can occur when the poverty lines are identical and when the Generalized Lorenz curves up to the poverty line are identical.[4,5] The alternative is true whenever P_2 is above P_1 for some point.

We could just as well reverse the roles of P_1 and P_2 and test similar hypotheses. This would allow one to determine whether a PGP curve dominated another in a stronger sense. In particular if one considered the hypotheses,

$$\begin{aligned} &H_0^2 : P_1(p; z_1) \leq P_2(p; z_2) \text{ for all } p \in [0,1]. \\ &H_1^2 : P_1(p; z_1) > P_2(p; z_2) \text{ for some } p \in [0,1]. \end{aligned} \tag{5}$$

then the hypotheses H_0^1 and H_1^2 together imply the strong dominance of P_1 over P_2 so that in principle one could use the tests to determine whether or not there is strong dominance. Note also that the hypotheses H_0^1 and H_0^2 together imply that the PGP curves are identical. For completeness, it may be of interest to formally test the null hypothesis of PGP equality,

$$\begin{aligned} H_0^{eq} &: \quad P_1(p; z_1) = P_2(p; z_2) \text{ for all } p \in [0,1]. \\ H_1^{eq} &: \quad P_1(p; z_1) \neq P_2(p; z_2) \text{ for some } p \in [0,1]. \end{aligned}$$

4 Properties of the Test Statistics

4.1 Estimator for PGP Curve

In this subsection we discuss how to estimate $P(p; z)$ based on a random sample of n observations drawn from F – these will be denoted by $\mathcal{X} = \{X_1, ..., X_n\}$. Let $\hat{z}$ be an estimator for z. Let $\hat{F}(x)$ and $\hat{Q}(p)$ denote the empirical distribution and quantile

[4] A further implication is that the headcount ratios are identical.

[5] Another possibility is that $F_1(x + a) = F_2(x)$ for all $x \in R$ and $z_1 = z_2 + a$. That is, F_1 is a location shift of F_2 and the poverty line z_1 is shifted accordingly.

function of income such that

$$\hat{F}(x) = \frac{1}{n}\sum_{i=1}^{n}(X_i \leq x), \qquad \hat{Q}(p) = \inf\{x : \hat{F}(x) \geq p\}.$$

The empirical poverty gap profile can be obtained simply by taking the empirical counterparts of the objects that define $P(p;z)$ to get

$$\hat{P}(p;\hat{z}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{z} - X_i) \cdot 1(X_i \leq \hat{z}) \cdot 1(X_i \leq \hat{Q}(p)).$$

It is straightforward to see that

$$\hat{P}(p;\hat{z}) = 1(\hat{Q}(p) \leq \hat{z}) \cdot (\hat{F}(\hat{Q}(p))\hat{z} - \hat{G}(p)) + 1(\hat{Q}(p) > \hat{z}) \cdot (\hat{F}(\hat{z})\hat{z} - \hat{G}(\hat{F}(\hat{z})))$$

with

$$\hat{G}(p) = \frac{1}{n}\sum_{i=1}^{n} X_i \cdot 1(X_i \leq \hat{Q}(p))$$

being the empirical Generalized Lorenz Curve at p.

4.2 Test Statistics

Our aim is to make inference regarding PGP dominance based on independent random samples from two populations. We make the following assumptions.

Assumption 4.1 *Assume that:*

1. $\{X_i^j\}_{i=1}^{n_j}$ *is a random sample from* F_j *and the sample for* $j = 1$ *is independent from the sample for* $j = 2$.

2. *the sampling scheme is such that as* $n_1 \to \infty$

$$\lim_{n_1\to\infty} \frac{n_1 n_2}{n_1 + n_2} \to \infty, \text{ and } \lim_{n_1\to\infty} \frac{n_1}{n_1 + n_2} \to \lambda \in (0,1)$$

The first part is the standard independent random samples assumption that would be appropriate in situations where we have two separate random samples from non-overlapping populations, such as countries or regions, or two random samples drawn at two different points in time for the same population. Note we allow for differing sample sizes. The requirement in (ii) is that, for the asymptotic analysis, the number of observations in each sample is not fixed as the other grows and it requires that the sample sizes are growing to infinity at the same rate. Note that the simple random sampling assumption can be relaxed in ways that are discussed in Bhattacharya (2005). It is also possible to allow for dependent sampling, such as with matched-pair sampling with multiple values for each observational unit, such as with panel data, as considered by Barrett, Donald and Bhattacharya (2014) for Lorenz dominance testing. For notational simplicity, we simply write n when taking limits and use n as the subscript when there is no confusion.

With the two independent samples, denote the respective empirical PGPs as $\hat{P}_1(p;\hat{z}_1)$ and $\hat{P}_2(p;\hat{z}_2)$. The proposed test of PGP dominance is the Cramer-von Mises type test with test statistic defined as

$$\hat{T}_1 = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \int_0^1 \max\left\{\hat{P}_2(p;\hat{z}_2) - \hat{P}_1(p;\hat{z}_1), 0\right\} dp,$$

the integral of the positive part of the difference between empirical PGPs with scaling factor $\sqrt{n_1 n_2/(n_1+n_2)}$. To derive the limiting null distribution of $\hat{T}_1$, we make the following assumptions.

Assumption 4.2 *Assume that for j=1 and 2*

1. *z_j is an interior point of $[x_l, x_u]$ where $0 \leq x_l < x_u < \infty$.*
2. *F_j is continuous on $[x_l, x_u]$ with probability density function $f_j(x)$ that is bounded away from zero on $[x_l, x_u]$.*

The first part of the assumption is simply that we know that the poverty line is finite, but we allow z_j to be known or unknown. For example, z_j can be a function of a sample quantile (e.g. median) or moment (e.g. mean). The second part of the assumption is that the distribution of income is continuous in a region that is slightly larger than the interval that contains all incomes below the poverty line. The support of income need not be finite since the PGP basically ignores the values of incomes that are above the poverty line. The requirement that $f_j(x)$ is bounded away from zero on $[x_l, x_u]$ is needed to allow one to obtain desirable asymptotic properties of the estimated quantile function estimator on this range.

Assumption 4.3 *Assume that for* $j = 1$ *and* 2*, the estimator* $\hat{z}_j$ *satisfies that*

$$\sqrt{n_j}(\hat{z}_j - z_j) = \frac{1}{\sqrt{n_j}} \sum_{i=1}^{n_j} \psi_{z_j}(X_{ji}; z) + o_p(1)$$

where $\psi_{z_j}(X_j; z_j)$ *is measurable with* $E[\psi_{z_j}(X_j; z_j)] = 0$ *and* $E[|\psi_{z_j}(X_j; z_j)|^{2+\delta}] < \infty$ *for some* $\delta > 0$*.*

Assumption 4.3 requires that $\hat{z}_j$ is asymptotically normally distributed with variance equal to $Var(\psi_{z_j}(X_j; z_j))$. This is not a restrictive assumption. For example, if $z_j = E[X_j]$ and let $\hat{z}_j$ be the mean estimator, then Assumption 4.3 would be satisfied with $\psi_{z_j}(X_j; z_j) = X_j - z_j$. If z_j is the median and let $\hat{z}_j$ be the sample median, then Assumption 4.3 would be satisfied with $\psi_{z_j}(X_j; z_j) = -f_j(z_j)^{-1}(1(X_j \leq z_j) - 1/2)$. If z_j is half of the median and let $\hat{z}_j$ be half of the sample median as is the case in the simulations and empirical studies, then Assumption 4.3 would be satisfied with $\psi_{z_j}(X_j; z_j) = -0.5 f_j(2z_j)^{-1}(1(X_j \leq 2z_j) - 1/2)$.

Let $\mathcal{P}_j(p)$ for $p \in [0, 1]$ denote a Gaussian process with covariance kernel generated by $\psi_{P_j}(X_j; p)$ such that

1. when $x_\ell \leq Q(p) \leq z_j$,

$$\begin{aligned} \psi_{P_j}(X_j; p) = &(z_j - X_j) \cdot 1(X_j \leq Q_j(p)) - P_j(p; z_j) \\ &+ p \cdot \psi_{z_j}(X_j; z_j) - (z_j - Q_j(p)) \cdot (1(X_j \leq Q_j(p)) - p), \end{aligned} \tag{6}$$

2. when $z_j \leq Q(p)$,

$$\psi_{P_j}(X_j; p) = (z_j - X_j) \cdot 1(X_j \leq z_j) - P_j(p; z_j) + F_j(z_j) \cdot \psi_{z_j}(X_j; z_j). \tag{7}$$

Define $P_o = \{p \in [0, 1] : \ P_1(p; z_1) = P_2(p; z_2)\}$ to be the contact set as in Linton, Song and Whang (2010).

Proposition 4.4 *Given Assumptions 4.1, 4.2 and 4.3, under* H_0^1*, we have*

$$\hat{T}_1 \xrightarrow{d} \int_{P_o} \max\left\{\sqrt{\lambda}\mathcal{P}_2(p) - \sqrt{1-\lambda}\mathcal{P}_1(p), 0\right\} dp,$$

where $\mathcal{P}_1$ *and* $\mathcal{P}_2$ *are two mutually independent Gaussian processes.*

The result that the limiting null distribution only depends on those p's in the contact set P_o is standard in the literature. In our case, however, the proof is more difficult than usual. This is due to the fact that we allow for the possibility of estimating the poverty line and also have to deal with a non-differentiability of the PGP curves at the point z_j. One should note that the result could also be used to derive the statistical properties of poverty indices that are functionals of the PGP and could be used to develop and justify inference methods along the lines of Barrett and Donald (2009).

4.3 Critical Value

To describe how we approximate the critical value we first describe our simulation method that is used to approximate the limiting process and also introduce the recentering method that is commonly employed in the literature to improve the power of a test for null hypotheses involving inequality constraints. Let $\{U_{1i}\}_{i=1}^{n_1}$ and $\{U_{2i}\}_{i=1}^{n_2}$ denote two sequences of i.i.d. random variables with mean 0, variance 1 and $E[|U_{ji}|^{2+\delta}] < \infty$ for some $\delta > 0$ that are independent of the samples – these could be standard normal random variables. Define the simulated processes $\hat{\mathcal{P}}_j^u(p)$ as

$$\hat{P}_j^u(p) = \frac{1}{n_j} \sum_{i=1}^{n_j} U_i \cdot \hat{\psi}_{P_j}(X_{ji}; p)$$

where $\hat{\psi}_{P_j}(X_{ji}; p)$ is the estimated influence function such that

1. when $\hat{Q}(p) \leq \hat{z}_j$,

$$\begin{aligned} \hat{\psi}_{P_j}(X_{ji}; p) = (\hat{z}_j - X_j) \cdot 1(X_j \leq \hat{Q}(p)) - \hat{P}_j(p; \hat{z}_j) + p \cdot \hat{\psi}_{z_j}(X_j; \hat{z}_j) \\ - (\hat{z}_j - \hat{Q}_j(p)) \cdot (1(X_j \leq \hat{Q}_j(p)) - p), \end{aligned}$$

2. when $\hat{Q}(p) > \hat{z}_j$,

$$\hat{\psi}_{P_j}(X_{ji}; p) = (\hat{z}_j - X_j) \cdot 1(X_j \leq \hat{z}_j) - \hat{P}_j(p; \hat{z}_j) + \hat{F}_j(\hat{z}_j) \cdot \hat{\psi}_{z_j}(X_j; \hat{z}_j).$$

Note that $\hat{\psi}_{z_j}(X_j; \hat{z}_j)$ is the estimated influence function for the estimator of z_j. If $z_j = E[X_j]$, then we have $\hat{\psi}_{z_j}(X_j; \hat{z}_j) = X_{ji} - \hat{z}_j$. If z_j is the median of the distribution of X_j, then we have $\hat{\psi}_{z_j}(X_j; \hat{z}_j) = -\hat{f}_j^{-1}(\hat{z}_j) \cdot (1(X_{ij} - \hat{z}_j) - 1/2)$ where $\hat{f}_j(\hat{z}_j)$ is a consistent kernel estimator for $f_j(z_j)$. If z_j is half of the median of the distribution of X_j, then we have $\hat{\psi}_{z_j}(X_j;\hat{z}_j) = -0.5\hat{f}_j^{-1}(2\hat{z}_j)\cdot(1(X_{ij}-2\hat{z}_j)-1/2)$ where $\hat{f}_j(2\hat{z}_j)$ is a consistent kernel estimator for $f_j(2z_j)$.

We use the recentering method described in Donald and Hsu (2013). The recentering method is similar to the generalized moment selection method of Andrews and Shi (2013) and the contact set method of Linton, Song and Whang (2010). These methods are proposed to improve the power of tests that involve inequality constraints by avoiding use of the least favorable configuration. For a sequence of positive numbers a_n, define

$$\hat{\mu}(p) = \big(\hat{P}_2(p;\hat{z}_2) - \hat{P}_1(p;\hat{z}_1)\big)\cdot 1\Big(\sqrt{\frac{n_1 n_2}{n_1+n_2}}\big(\hat{P}_2(p;\hat{z}_2) - \hat{P}_1(p;\hat{z}_1)\big) < -a_n\Big). \tag{8}$$

Define the simulated test statistic as

$$\hat{T}_1^u = \sqrt{\frac{n_1 n_2}{n_1+n_2}}\int_0^1 \max\big\{\hat{P}_2^u(p) - \hat{P}_1^u(p) + \hat{\mu}(p), 0\big\}dp.$$

Let α denote the significance level. The simulated critical value is defined as

$$\hat{c}^1_{\eta,n} = \max\{\tilde{c}^1_n, \eta\},$$

where η is an arbitrarily small positive number, say 10^{-6} and

$$\tilde{c}^1_n = \inf\Big\{c:\ P\Big(\hat{T}_1^u \geq c\Big) \leq \alpha\Big\},$$

i.e., $\tilde{c}^1_n$ is the $(1-\alpha)$-th quantile of $\hat{T}_1^u$. With the critical value in hand the decision rule for the test is,

Reject H_0^1, if $\hat{T}_1 > \hat{c}^1_{\eta,n}$.

This method can also be used to generate p-values by finding the proportion of simulated $\max\{\hat{T}_1^u,\eta\}$ that exceed the test statistic value $\hat{T}_1$. A decision rule based on the p-value would be equivalent to one based on comparing the test statistic to the critical value.

4.4 Size and Power Properties of the Proposed Test

The following result describes the behavior of our test procedure under the null and alternative hypotheses. To derive this result we impose the following conditions on a_n.

Assumption 4.5 *Let a_n be a sequence of negative numbers such that $\lim_{n\to\infty} a_n = -\infty$ and $\lim_{n\to\infty} n_1^{-1/2} a_n = 0$.*

Theorem 4.6 *Suppose that Assumptions 4.1, 4.2, 4.3 and 4.5 hold and we reject* H_0^1 *if* $\hat{T}_1 > \hat{c}^1_{\eta,n}$. *Then,*

1. *suppose that* H_0^1 *is true and the Lebesgue measure of* P_o *is zero, then* $\lim_{n\to\infty} P(\hat{T}_1 > \hat{c}^1_{\eta,n}) = 0$,

2. *suppose that* H_0^1 *is true and the Lebesgue measure of* P_o *is strictly greater than zero, then* $\lim_{\eta\to 0}\lim_{n\to\infty} P(\hat{T}_1 > \hat{c}^1_{\eta,n}) = \alpha$, *and*

3. *suppose that* H_1^1 *is true, then* $\lim_{n\to\infty} P(\hat{T}_1 > \hat{c}^1_{\eta,n}) = 1$.

The first two results describe the size of our test and show that size is no larger than the nominal size α and that this nominal size is achieved as long as the contact set is non-empty when η is chosen to be small. The third result shows that the test is consistent against a fixed alternative.

The test of H_0^2 against H_1^2 is exactly analogous to this procedure. The test statistic for testing H_0^{eq} against H_1^{eq} is defined as

$$\hat{T}_{eq} = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \int_0^1 \left|\hat{P}_2(p;\hat{z}_2) - \hat{P}_1(p;\hat{z}_1)\right| dp$$

and it is straightforward to show that under H_0^{eq}

$$\hat{T}_{eq} \xrightarrow{d} \int_0^1 \left|\sqrt{\lambda}\mathcal{P}_2(p) - \sqrt{1-\lambda}\mathcal{P}_1(p)\right| dp.$$

For the significance level α, the simulated critical value is defined as

$$\hat{c}^{eq}_n = \inf\left\{c: \ P\left(\hat{T}^u_{eq} \geq c\right) \leq \alpha\right\}$$

where

$$\hat{T}^u_{eq} = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \int_0^1 \left|\hat{P}^u_2(p) - \hat{P}^u_1(p)\right| dp.$$

The decision rule is to "reject H_0^{eq} when $\hat{T}^u_{eq} > \hat{c}^{eq}_n$." Critical values can be obtained in a manner similar to that for the test of weak dominance.

4.5 Uniform Size Control

As discussed in Andrews and Shi (2013), pointwise asymptotics, as considered above, may not provide a good approximation to the finite-sample properties of test statistics for null hypotheses involving inequality constraints. Hence, in this subsection, we extend our pointwise results to a uniform result by adopting the methods of Andrews and Shi (2013).

We first modify our recentering function. Let b_n be a sequence of positive numbers. Define $\hat{\phi}_n(z)$ as

$$\hat{\phi}_n(p) = -b_n \cdot 1\Big(\sqrt{\frac{n_1 n_2}{n_1+n_2}}\big(\hat{P}_2(p;\hat{z}_2) - \hat{P}_1(p;\hat{z}_1)\big) < -a_n\Big). \tag{9}$$

where a_n is defined as above. Note that $\hat{\phi}_n(p)$ is a modification of $\hat{\mu}(p)$, where the recentering parameter $\big(\hat{P}_2(p;\hat{z}_2) - \hat{P}_1(p;\hat{z}_1)\big)$ is replaced by $-b_n$ when $\sqrt{n_1 n_2/(n_1+n_2)}\big(\hat{P}_2(p;\hat{z}_2) - \hat{P}_1(p;\hat{z}_1)\big)$ is less than $-a_n$. Following Andrews and Shi (2013), b_n will be picked in a way such that $\hat{\phi}_n(z) \leq P_2(p;\hat{z}_2) - P_1(p;\hat{z}_1)$ for all $p \in [0,1]$ with probability approaching one. We define the critical value $\hat{c}^u_{\eta,n}$ as

$$\hat{c}^{1,u}_{\eta,n} = \tilde{c}^{1,u}_{\eta,n} + \eta,$$
$$\tilde{c}^{1,u}_{\eta,n} = \sup\left\{ q \Big| P^u\Big(\sqrt{\frac{n_1 n_2}{n_1+n_2}} \int_0^1 \max\big\{\hat{P}^u_2(p) - \hat{P}^u_1(p) + \hat{\phi}_n(p), 0\big\} dp \leq q\Big) \leq 1-\alpha+\eta \right\},$$

where the subscript u denotes the critical value we use to derive the uniformity property of our test.

We imposes conditions on a_n and b_n that are similar to those in Assumptions GMS1 and GMS2 in Andrews and Shi (2013).

Assumption 4.7 *Assume that:*

1. *a_n satisfies Assumption 4.5.*

2. *b_n is a sequence of positive numbers such that (i) $\sqrt{n_1} b_n$ is non-decreasing and (ii)* $\lim_{n\to\infty} \sqrt{n_1} b_n / a_n = 0$.

We characterize the set of data generating processes (DGPs) such that our test will have uniform size. As in the proof of Andrews and Shi (2013), the key is to characterize a subset of DGPs that is “compact" in some sense. Note that a PGP curve is fully characterized by the CDF and the poverty line, and the asymptotics of a PGP estimator are fully characterized by the CDF, the poverty line and the influence function of the poverty line estimator. Therefore, let $\theta = (F, z, \psi_z)$ be an index of a DGP. Let Θ denote a collection of DGPs such that the following conditions are satisfied.

Assumption 4.8 *Assume that for all $\theta \in \Theta$,*

1. *F is continuous on $[x_l, x_u]$ where $0 \leq x_l < x_u \leq M$ with probability density function $f(x)$ such that $\delta \leq f(x) \leq M$ for some $0 < \delta < 1/2$ and $M > 0$,*
2. *$z \in [x_l + \delta, x_u - \delta]$ and $\delta \leq F(z) \leq 1 - \delta$, and*
3. *the estimator $\hat{z}$ is such that uniformly over $\theta \in \Theta$,*

$$\left|\sqrt{n}(\hat{z} - z) - \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi_z(X_i; z)\right| = o_p(1)$$

where $\psi_z(X; z)$ is measurable with $E[\psi_z(X; z)] = 0$ and $E[|\psi_z(X; z)|^{2+\delta}] < M$.

Let $\mathcal{H}_2$ denote the set of all covariance kernel functions defined on $([0, 1 - \delta] \cup \{1\})^2$. For each θ, let $h_{2,\theta}$ denote a covariance kernel function on $([0, 1 - \delta] \cup \{1\})^2$ generated by $\psi_P(X; p)$ such that

1. when $0 \leq p \leq 1 - \delta$,

$$\begin{aligned} \psi_P(X; p) = {} & (z - X) \cdot 1(X \leq Q(p)) - P(p; z) \\ & + p \cdot \psi_z(X; z) - (z - Q(p)) \cdot (1(X \leq Q(p)) - p), \end{aligned} \tag{10}$$

2. when $p = 1$

$$\psi_P(X; 1) = (z - X) \cdot 1(X \leq z) - P(1; z) + F(z) \cdot \psi_z(X; z). \tag{11}$$

Obviously, $h_{2,\theta} \in \mathcal{H}_2$. Let the supremum norm on $\mathcal{H}_2$ be $d_h(h_2^1, h_2^2) = \sup_{p', p'' \in [0, 1-\delta] \cup \{1\}} |h_2^1(p', p'') - h_2^2(p', p'')|$ for any $h_2^1, h_2^2 \in \mathcal{H}_2$. Note that (10) and (11) are used to make it easier to characterize the set for which our test has uniform size.[6] Define $\Theta^2 = \Theta \times \Theta$ where Θ satisfies Assumption 4.8 and $\Theta_0^2 = \{(\theta_1, \theta_2) \in \Theta^2 | P_1(p; z_1) \leq P_2(p; z_2) \text{ for } p \in [0, 1]\}$. That is, Θ_0^2 is the subset of Θ^2 such that the null hypothesis holds.

[6] Please see Appendix for details.

Theorem 4.9 *Suppose Assumption 4.1, 4.7 and 4.8 hold. Then for every compact subset* $\mathcal{H}_{2,cpt}$ *of* $\mathcal{H}_2$,

1. $\limsup_{n\to\infty} \sup_{\{(\theta_1,\theta_2)\in\Theta_0^2 | h_{2,\theta_1}, h_{2,\theta_2}\in\mathcal{H}_{2,cpt}\}} P(\hat{T}_1 > \hat{c}^{1,u}_{\eta,n}) \leq \alpha$, *and*

2. $\lim_{\eta\to 0}\limsup_{n\to\infty} \sup_{\{(\theta_1,\theta_2)\in\Theta_0^2 | h_{2,\theta_1}, h_{2,\theta_2}\in\mathcal{H}_{2,cpt}\}} P(\hat{T}_1 > \hat{c}^{1,u}_{\eta,n}) = \alpha$.

Theorem 4.9 is similar to Theorem 2 and Theorem B1 of Andrews and Shi (2013), and Theorem 6.1 of Donald and Hsu (2013). The first part shows that our test has correct size uniformly over a set of DGPs and the second part shows that our test is at most infinitesimally conservative asymptotically. Similar to the pointwise case, the method with uniform size control can also be used to generate p-values by finding the proportion of simulated $\hat{T}_1^u + \eta$ that exceed the test statistic value $\hat{T}_1$.

5 Tests for Normalized PGP Dominance

In this section, we briefly summarize how to test for normalized PGP dominance. Recall that the normalized PGP curve is defined as the corresponding PGP curve divided by the associated poverty line:

$$P_j^R(p, z_j) = \frac{P_j(p, z_j)}{z_j} \quad \text{for } p \in [0,1].$$

which can be estimated by

$$\hat{P}_j^R(p, \hat{z}_j) = \frac{\hat{P}_j(p, \hat{z}_j)}{\hat{z}_j} \quad \text{for } p \in [0,1].$$

The hypotheses that the normalized PGP curve for the population F_1 is everywhere at least as large as that for the population F_2 are defined as

$$\begin{aligned} &H_0^{1R} : P_2^R(p; z_2) \leq P_1^R(p; z_1) \text{ for all } p \in [0,1] \\ &H_1^{1R} : P_2^R(p; z_2) > P_1^R(p; z_1) \text{ for some } p \in [0,1] \end{aligned} \tag{12}$$

The proposed test statistic for H_0^{1R} against H_1^{1R} for normalized PGP dominance is

$$\hat{T}_{1,R} = \sqrt{\frac{n_1 n_2}{n_1+n_2}} \int_0^1 \max\left\{\hat{P}_2^R(p; \hat{z}_2) - \hat{P}_1^R(p; \hat{z}_1), 0\right\} dp.$$

Define $P_o^R = \{p \in [0,1] : \; P_1^R(p; z_1) = P_2^R(p; z_2)\}$. Under the same conditions and under H_0^{1R}, we can show that

$$\hat{T}_{1,R} \xrightarrow{d} \int_{P_o^R} \max\left\{\sqrt{\lambda}\mathcal{P}_2^R(p) - \sqrt{1-\lambda}\mathcal{P}_1^R(p), 0\right\} dp,$$

where $\mathcal{P}_1^R(p)$ and $\mathcal{P}_2^R(p)$ are two mutually independent Gaussian processes with covariance kernels generated by $\psi_{P_j}^R(X_j; p)$ such that

$$\psi_{P_j}^R(X_j; p) = \frac{1}{z_j}\Big(\psi_{P_j}(X_j; p) - P_j^R(p; z_j) \cdot \psi_{z_j}(X_j; z_j)\Big). \tag{13}$$

This is shown in the Appendix. Therefore a critical value can be constructed in a similar fashion to that described above for the PGP curve. A test for normalized PGP dominance with uniform size control can also be constructed using methods similar to those described above.

6 Monte Carlo Results

In this section we consider a small scale Monte Carlo experiment in which we gauge the extent to which the preceding asymptotic arguments hold in small samples. We consider a few cases that illustrate the properties of the tests in a variety of situations and consider both the size and power properties of the tests. We use distributions in the log-normal family because they are easy to simulate and have been widely used in empirical work on income distributions. We generate two sets of samples from two (possibly different) distributions. In each case we generate X^1 and X^2 as (independent) log-normal random variables using the equations,

$$\begin{aligned} X_i^1 &= \exp(\sigma_1 Y_{1i} + \mu_1) \\ X_j^2 &= \exp(\sigma_2 Y_{2j} + \mu_2) \end{aligned}$$

where the Y_{1i} and Y_{2j} are independent $N(0,1)$.

The first series of experiments consider tests of absolute PGP dominance with the poverty line estimated using half the sample median. In Case 1, $\mu_1 = \mu_2 = 0.85$ and $\sigma_1 = \sigma_2 = 0.6$. These parameters generate distributions with means equal to 2.8 and standard deviations equal to 1.8 – the ratio of which is similar to US CPS income data. In Case 1 the PGPs for the two populations are identical and we are interested in the size properties of the testing procedure. The second case, Case 2, $\mu_1 = 0.85$, and $\sigma_1 = 0.6$ while $\mu_2 = 0.75$ and $\sigma_2 = 0.6$. In this case, when using half the sample median as the poverty line, one can show that the PGP for X^2 is below that for X^1 (the PGP curve for 2 lies below that for 1 everywhere except at the origin). In this case we should expect that we do not reject the hypothesis H_0^1 but we should reject H_0^2. We consider tests of both of these hypotheses as well as H_0^{eq}. Note also that in this case we should expect that the test will reject H_0^1 less often than the nominal size of the test because of the result in Proposition 4.6. In Case 3 $\mu_1 = 0.85$, and $\sigma_1 = 0.6$ while $\mu_2 = 0.85$ and $\sigma_2 = 0.62$, resulting in X^2 have greater inequality and poverty depth and severity (though lower incidence) and distribution X^2 strongly PGP dominating X^1. Therefore we expect to not reject H_0^2 and reject H_0^1 and H_0^{eq}. For Case 4, $\mu_1 = 0.85$, and $\sigma_1 = 0.6$ while $\mu_2 = 0.85$ and $\sigma_2 = 0.7$. This is similar to Case 3 and is used to examine power as the violation of the nulls H_0^1 and H_0^{eq} is larger in this case. Each of these specifications results in poverty incidence or head-count ratios of between 0.12 and 0.16.

In performing the test of poverty dominance we use the decision rule described above based on the appropriate simulated critical values. For all of the experiments we considered sample sizes of $n_j = 200, 500, 1000$. The number of simulations used to estimate the critical values is 1000. To account for estimation of the poverty line by half the sample median we use,

$$\psi_{z_j}(X_{ji}; \hat{z}_j) = -0.5\hat{f}_j(\hat{z}_j)^{-1}\big[1\big(X_{ji} - 2\hat{z}_j\big) - 1/2\big]$$

where $\hat{f}_j(2\hat{z}_j)$ is a nonparametric estimator for $f_j(2z_j)$ such that

$$\hat{f}_j(2\hat{z}_j) = \frac{1}{nh}\sum_{i=1}^{n_j} K\Big(\frac{X_{ji} - 2\hat{z}_j}{h_j}\Big)$$

with $K(u) = 3/4(1-u^2)$ for $|u| \leq 1$. The bandwidth is set at $h_j = 1.06\hat{\sigma}_j n_j^{-1/5}$ where $\hat{\sigma}_j$ is the sample deviation of the sample j.

We implement the test with uniform size control. For the recentering parameter, we set

$a_n = -0.1 \cdot \sqrt{0.3(\log(n_1+n_2))} \cdot \hat{\sigma}_j$ and $b_n = 0.1\sqrt{\frac{n_1+n_2}{n_1 n_2}}\sqrt{0.4\frac{\log(n_1+n_2)}{\log(\log(n_1+n_2))}} \cdot \hat{\sigma}_j$ for the PGP dominance case where $0.1\sqrt{0.3(\log(n_1+n_2))}$ and $0.1\sqrt{\frac{n_1+n_2}{n_1 n_2}}\sqrt{0.4\frac{\log(n_1+n_2)}{\log(\log(n_1+n_2))}}$ are similar to what is suggested in Andrews and Shi (2013) and $\hat{\sigma}_j$ is the standard deviation of the variable in question. In our simulations, the $\hat{\sigma}_j$ is roughly equal to 2 for each case. For the normalized PGP dominance case, we could use $a_n = -0.1\sqrt{0.3(\log(n_1+n_2))} \cdot \frac{\hat{\sigma}_j}{\hat{z}_j}$ and $b_n = \sqrt{\frac{n_1+n_2}{n_1 n_2}}\sqrt{0.4\frac{\log(n_1+n_2)}{\log(\log(n_1+n_2))}} \cdot \frac{\hat{\sigma}_j}{\hat{z}_j}$. Since in all of our simulations the $\hat{z}_j$'s are close to 1 it becomes convenient to use $a_n = -2 \cdot \sqrt{0.3(\log(n_1+n_2))}$ and $b_n = 0.1\sqrt{\frac{n_1+n_2}{n_1 n_2}}\sqrt{0.4\frac{\log(n_1+n_2)}{\log(\log(n_1+n_2))}}$ for all simulations. We set $\eta = 10^{-6}$. The number of points that we use to approximate the integral is 200. For each experiment the total number of Monte Carlo replications was set at 1000. The table reports the proportion of times that the respective null hypothesis was rejected for three different nominal significance levels α.

The Monte Carlo results based on the regular PGP curve are contained in Table 1. Results based on the normalized PGP curve are found in Table 2. In Table 1, Case 1 shows that the tests have actual size close to nominal for all the tests, even with the smallest sample size considered. In Case 2, the test of PGP dominance is able to detect the violation of the null H_0^2 (and H_0^{eq}), with rejection rates that exceed the nominal size for all sample sizes and a rejection rate that increases with the sample size. The true null of H_0^1 is rejected less often than the nominal size. In Case 3, the false null H_0^1 is rejected more often, and the true null H_0^2 is rejected less often, than the nominal size of the test. In Case 4, where there is a stronger PGP dominance of X^2 over X^1 we see that the rejection of H_0^1 occurs with higher frequency and shows the power of the test. Overall, these small scale experiments suggest even in small samples the absolute PGP dominance tests have size and power properties that are consistent with our theoretical results. The recentering has little impact on the properties of the tests in these specifications. The results based on the normalized PGP's in Table 2 display similar properties although it appears in Case 2 that the two normalized PGP's are very similar since all the tests have rejection rates close to nominal size.

Table 1: Monte Carlo Results: Rejection Rates - Absolute PGP

		H_0^1			H_0^2			H_0^{eq}		
	n_j	**0.10**	**0.05**	**0.01**	**0.10**	**0.05**	**0.01**	**0.10**	**0.05**	**0.01**
Case 1	200	0.110	0.060	0.013	0.094	0.045	0.007	0.106	0.053	0.012
	500	0.112	0.057	0.014	0.111	0.055	0.008	0.110	0.052	0.011
	1000	0.105	0.050	0.014	0.114	0.067	0.014	0.117	0.054	0.015
Case 2	200	0.069	0.040	0.011	0.169	0.093	0.022	0.131	0.070	0.021
	500	0.033	0.013	0.005	0.215	0.133	0.046	0.143	0.092	0.032
	1000	0.024	0.011	0.002	0.275	0.166	0.049	0.176	0.100	0.032
Case 3	200	0.173	0.094	0.016	0.074	0.027	0,003	0.123	0.056	0.014
	500	0.203	0.108	0.030	0.039	0.022	0.003	0.129	0.075	0.021
	1000	0.255	0.163	0.042	0.041	0.020	0.002	0.178	0.100	0.024
Case 4	200	0.484	0.337	0.144	0.010	0.003	0	0.332	0.239	0.087
	500	0.692	0.563	0.313	0.001	0	0	0.567	0.456	0.227
	1000	0.896	0.814	0.607	0	0	0	0.817	0.738	0.506

Table 2. Monte Carlo Results: Rejection Rates - Normalized PGP

		H_0^1			H_0^2			H_0^{eq}		
	n_j	**0.10**	**0.05**	**0.01**	**0.10**	**0.05**	**0.01**	**0.10**	**0.05**	**0.01**
Case 1	200	0.099	0.051	0.014	0.116	0.061	0.018	0.116	0.069	0.020
	500	0.106	0.055	0.009	0.099	0.053	0.016	0.107	0.053	0.013
	1000	0.103	0.053	0.005	0.106	0.060	0.013	0.108	0.048	0.012
Case 2	200	0.095	0.050	0.013	0.119	0.066	0.014	0.115	0.056	0.017
	500	0.112	0.047	0.005	0.109	0.064	0.011	0.112	0.057	0.008
	1000	0.099	0.048	0.015	0.104	0.064	0.015	0.110	0.056	0.014
Case 3	200	0.205	0.105	0.031	0.063	0.038	0.008	0.144	0.086	0.023
	500	0.233	0.129	0.046	0.050	0.024	0.005	0.151	0.080	0.030
	1000	0.263	0.162	0.056	0.020	0.010	0.001	0.169	0.106	0.030
Case 4	200	0.540	0.400	0.182	0.006	0.004	0.001	0.404	0.307	0.140
	500	0.816	0.704	0.441	0	0	0	0.705	0.577	0.365
	1000	0.954	0.920	0.769	0	0	0	0.918	0.867	0.691

7 Empirical Example: Australian Consumption Poverty 1988-2009

In this section we illustrate the methods of testing for poverty dominance relations by assessing contemporary trends in relative consumption poverty in Australia over the period 1988/89-2009/10. The data are from the Australia Bureau of Statistics Household Expenditure Survey (HES) conducted in 1988/89, 1993/94, 1998/99, 2003/04 and 2009/10 (hereafter referenced by the first year of the survey period). The welfare concept examined is consumption as measured by expenditure on the set of non-durables consisting of food, alcohol and tobacco, fuel, clothing, personal care, medical care, transport, recreation, utilities and current housing services. Current housing services for renters is equal to rent paid while for home-owners it is imputed from a regression of rent payments on a series of indicator variables for number of bedrooms and location of residence by survey year for the subsample of renters. The sample is restricted to families where the household reference person is 25-60 years of age.

Family consumption is divided by the adult equivalent scale (AES) equal to the square-root of family size. To minimize reporting errors multiple-family households are excluded. The HES is a stratified random sample and for each observation there is an associated weight representing the inverse probability of selection into the survey. The observational weights are multiplied by the number of family members to make the sample representative of individuals; the adjusted weights were used throughout the analysis.

Summary statistics are reported in Table 3. Nominal prices are inflated to 2010 real values using the CPI. The mean budget share of the non-durable commodity bundle was 68 percent in 1988. Over the sample period non-durable consumption grew at an average annual rate of 2.36 percent. The poverty line in each year is set equal to half the median consumption level; the growth in median consumption translated into an increasing absolute value of the poverty threshold over time. Point estimates for the headcount ratio and mean poverty gap suggest that the incidence and depth of consumption poverty increased over the 21 year observation period. Plots of the empirical absolute and normalized PGPs and differences for adjacent surveys, are presented in Figures 3-12.

Table 3: HES 1988-2009 Summary Statistics

Year	n	Consumption Distribution						
		Mean μ	*Median*	$\frac{z}{\mu}$	$F(z)$	$E(P)$	$E(P\|P>0)$	*Modified Sen*
1988	4654	173.02	154.68	0.447	0.052	0.825	15.885	0.011
1993	5396	186.50	165.67	0.444	0.038	0.585	15.407	0.007
1998	4645	201.40	179.41	0.445	0.049	0.900	18.545	0.010
2003	4583	215.14	194.85	0.453	0.061	1.204	19.812	0.012
2009	5009	251.42	224.96	0.447	0.067	1.378	20.712	0.012

Table 4 presents the test results based on comparisons of absolute PGPs. The null hypothesis is that distribution 1 weakly PGP dominated distribution 2, against the alternative that the null is false. In this case we report p-values for each test which gives the proportion of simulated draws that exceed the calculated test statistic value. To do this 5000 simulations were used. We account for estimation of the poverty line using the same method as used in the simulations and use re-centering based on $a_n = -10\sqrt{0.3(\log(n_1+n_2))}$ and $b_n = 10\sqrt{\frac{n_1+n_2}{n_1 n_2}}\sqrt{0.4\frac{\log(n_1+n_2)}{\log(\log(n_1+n_2))}}$ since the standard deviations in of the data sets are close to 100. For the normalized PGP dominance case, we set $a_n = -0.1\sqrt{0.3(\log(n_1+n_2))}$ and $b_n = 0.1\sqrt{\frac{n_1+n_2}{n_1 n_2}}\sqrt{0.4\frac{\log(n_1+n_2)}{\log(\log(n_1+n_2))}}$. We set $\eta = 10^{-6}$. The number of points that we use to approximate the integral is 1,000. The results were not sensitive to these choices.

The first row of Table 4 is for the test with distribution 1 corresponding to 1988 and distribution 2 corresponding to 1993. The results show that the null of H_0^1 cannot be rejected at any conventional level of significance, while H_0^2 is rejected at the 5% level. The p-value for the null of PGP equality $H_0^{eq}$of 0.027 also implies rejection at the 5% levels of significance. These results indicate that the absolute PGPs shifted down, and consumption poverty unambiguously decreased between 1988 and 1993. The results concerning the other years suggest that the 1993 consumption distribution PGP dominated the 1998 distribution. One can reject H_0^1 (and H_0^{eq}) in this instance but one cannot reject H_0^2. This strong PGP dominance of the 1998 distribution over the 1993 distribution implies that poverty rose from 1993 to 1998. A comparison of the 1998 and 2003 distributions show that the 2003 weakly poverty dominated 1998 (though the null of poverty equality cannot be rejected at the 5% level). Comparing the 2003 and

2009 consumption distributions show that the none of the null hypotheses considered can be rejected, which implying that the respective PGPs coincide. Across the full observation period, the 2003 consumption distribution strongly poverty dominates the 1988 distribution implying an increase in poverty over the two decades.

Table 4: PGP Dominance Test Results

F_1	F_2	Test	*Test Statistic*	*P-value*	Inference
1988	1993	H_0^1	0.001	0.650	Pov $\downarrow$
		H_0^2	11.719	0.014	
		H_0^{eq}	11.720	0.029	
1993	1998	H_0^1	15.397	0.004	Pov $\uparrow$
		H_0^2	0.001	0.665	
		H_0^{eq}	15.298	0.009	
1998	2003	H_0^1	14.096	0.028	Pov $\uparrow$
		H_0^2	0.004	0.626	
		H_0^{eq}	14.100	0.057	
2003	2009	H_0^1	8.485	0.168	Pov $\sim$
		H_0^2	0.000	1.000	
		H_0^{eq}	8.485	0.330	
1988	2009	H_0^1	26.553	$<$0.001	Pov $\uparrow$
		H_0^2	0	1.000	
		H_0^{eq}	26.553	$<$0.001	

Table 5: Normalized PGP Dominance Test Results

F_1	F_2	Test	*Test Statistic*	*P-value*	Inference
1988	1993	H_0^1	<0.001	0.650	Pov $\downarrow$
		H_0^2	0.176	0.003	
		H_0^{eq}	0.176	0.006	
1993	1998	H_0^1	0.145	0.014	Pov $\uparrow$
		H_0^2	<0.001	0.657	
		H_0^{eq}	0.145	0.027	
1998	2003	H_0^1	0.107	0.080	Pov $\sim$
		H_0^2	<0.001	0.542	
		H_0^{eq}	0.108	0.154	
2003	2009	H_0^1	<0.001	0.555	Pov $\sim$
		H_0^2	0.004	0.475	
		H_0^{eq}	0.005	0.959	
1988	2009	H_0^1	0.075	0.152	Pov $\sim$
		H_0^2	<0.001	0.585	
		H_0^{eq}	0.075	0.327	

Table 5 contains the test results based on the normalized PGPs with the poverty lines estimated again using half the sample median. As evident from the figures, this rescaling changes the shape of the PGPs and potentially the poverty orderings. As shown by the test results in Table 5, the normalized PGP comparison also indicate a decline in consumption poverty from 1988 to 1993, then a reversal to 1998. The relative PGPs were generally stable from 1998 to 2009, and over the full sample period there was no discernible change in relative consumption poverty.

8 Conclusion

In this paper we propose methods for testing for poverty dominance relations based on the empirical Poverty Gap Profile. The tests are non–parametric and consistent being based on global comparisons of the complete PGP at every empirical ordinate. The proposed test statistics have non-standard, case specific limiting distributions and we demonstrate that asymptotically valid inferences could be drawn using simulations. The tests of poverty dominance are shown to have a good performance in small samples and were illustrated in the context of an analysis of consumption poverty in Australia over the period 1988/89-2009/10.

Appendix A: Implementation Procedure

This appendix describes in detail the algorithm for finding the appropriate critical value for testing PGP dominance. We present the critical value for uniform size case and the critical value for pointwise size control can be obtained similarly.

1. Calculate $\hat{z}_j$, $\hat{F}_j(x)$ and $\hat{Q}_j(p)$.

2. Calculate the estimated PGP curve $\hat{P}_j(p;\hat{z}_j)$ as

$$\hat{P}_j(p;\hat{z}_j) = \frac{1}{n_j}\sum_{i=1}^{n_j}(\hat{z}_j - X_j)\cdot 1(X_j \le \hat{z}_j)\cdot 1(X_j \le \hat{Q}_j(p))$$

   for an evenly spaced grid on $[0,1]$ with say 201 points, i.e., $[0, 0.005, \ldots, 1]$.

3. Calculate the test statistic $\hat{T}_1$ by a Riemann sum

$$\hat{T}_1 = \sqrt{\frac{n_1 n_2}{n_1+n_2}}\sum_{k=1}^{200}\frac{1}{200}\max\left\{\hat{P}_2(k/200;\hat{z}_2) - \hat{P}_1(k/200;\hat{z}_1), 0\right\}.$$

4. Calculate the recentering function $\hat{\phi}_N(p)$ according to (9) for $p \in [0, 0.005, \ldots, 1]$.

5. Calculate estimated influence function $\hat{\psi}_{P_j}(p;\hat{z}_j)$ according to (6) and (7) for $p \in [0, 0.005, \ldots, 1]$.

6. Generate pseudo random variables $\{U_{ji}^b\}_{i=1}^{n_j}$ from the standard normal distribution for $b = 1, \ldots, B$, say $B = 1000$.

7. Calculate $\hat{P}_j^{ub}(p;\hat{z}_j)$ by

$$\hat{P}_j^{ub}(p;\hat{z}_j) = \frac{1}{n_j}\sum_{i=1}^{n_j} U_{ji}^b \cdot \hat{\psi}_{P_j}(p;\hat{z}_j)$$

   for $p \in [0, 0.005, \ldots, 1]$ and for $b = 1, \ldots, B$.

8. Calculate simulated test statistic $\hat{T}_1^{ub}$ by a Riemann sum

$$\hat{T}_1^{ub} = \sqrt{\frac{n_1 n_2}{n_1+n_2}}\sum_{k=1}^{200}\frac{1}{200}\max\Big\{\hat{P}_2^{ub}(k/200;\hat{z}_2) - \hat{P}_1^{ub}(k/200;\hat{z}_1) + \hat{\phi}_N(k/200), 0\Big\} \quad (14)$$

   for $b = 1, \ldots, B$.

9. Rank $\hat{T}_1^{ub}$ in ascending order, i.e., $\hat{T}_1^{u(1)} \leq \hat{T}_1^{u(2)} \leq \ldots \leq \hat{T}_1^{u(B)}$.

10. Let $\eta = 10^{-6}$ and calculate $\hat{c}_{\eta,n}^{1,u} = \hat{T}_1^{u((1-\alpha)B)+1} + \eta$. That is, if $\alpha = 5\%$ and $B = 1000$, then $\hat{c}_{\eta,n}^{1} = \hat{T}_1^{u(951)} + \eta$ and if $\alpha = 10\%$ and $B = 1000$, then $\hat{c}_{\eta,n}^{1} = \hat{T}_1^{u(901)} + \eta$.

11. Reject H_0^1 if $\hat{T}_1^{ub} > \hat{c}_{\eta,n}^{1,u}$.

Appendix B: Proofs of Results

Proof of Proposition 4.4: We first discuss the asymptotics of the $P(p, z)$ estimators where we drop the subscripts first for notational simplicity. Define $\bar{p} = F(z)$ and $p_u = F(x_u)$.

Let $w(X, s, s') = (s - X) \cdot 1(X \leq s) \cdot 1(X \leq s')$ where $s, s' \in [x_l, x_u]$. It is straightforward to see that $\{w(X, s, s') : \ s, s' \in [x_l, x_u]\}$ is a Donsker class with envelope function $2x_u$ that is of finite second moment. Then by functional central limit theorem as in van der Vaart and Wellner (1996), we have

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \big(w(X_i, s, s') - E[w(X, s, s')] \big) \Rightarrow \mathcal{W}(s, s')$$

where $\mathcal{W}(s, s')$ is a Gaussian process with covariance kernel generated by $w(X, s, s')$. This implies that

$$\sup_{s,s' \in [x_l, x_u]} \Big| \frac{1}{n} \sum_{i=1}^{n} w(X_i, s, s') - E[w(X, s, s')] \Big| \xrightarrow{p} 0.$$

Note that as in (2), we have

$$\begin{aligned} E[w(X, s, s')] &= 1(s' \leq s) \cdot (F(s')s - G(F(s'))) + 1(s' > s) \cdot (F(s)s - G(F(s)) \\ &= F(\min\{s', s\})s - G(F(\min\{s', s\})). \end{aligned}$$

Then for $p \in [0, p_u]$ with $p_u = F(x_u)$, $P(p; z) = E[w(X, z, Q_F(p))]$ and

$$\hat{P}(p; \hat{z}) = \frac{1}{n} \sum_{i=1}^{n} w(X_i, \hat{z}, \hat{Q}(p)).$$

Given that $E[w(X, s, s')]$ is uniformly continuous on $[x_l, x_u]^2$, $\hat{z} \xrightarrow{p} z$ and $\sup_{p \in [0,p_u]} |\hat{Q}(p) - Q(p)| \xrightarrow{p} 0$, it follows that $\sup_{p \in [0,p_u]} |\hat{P}(p; \hat{z}) - P(p; z)| \xrightarrow{p} 0$. Let δ_n be a sequence of positive numbers with $\delta_n \to 0$ and $\sqrt{n}\delta_n \to \infty$. We claim that

$$\lim_{n \to \infty} P\big(\sup_{p \leq \bar{p} - \delta_n} \hat{Q}(p) < \hat{z} \big) = 1, \quad \lim_{n \to \infty} P\big(\inf_{p \geq \bar{p} + \delta_n} \hat{Q}(p) > \hat{z} \big) = 1. \tag{15}$$

We show the first result with the argument for the second one being similar. Let $\epsilon_n = z - Q(p - \delta_n)$ and note that it is straightforward to see that $\epsilon_n \to 0$ and $\sqrt{n}\epsilon_n \to \infty$. It is also true that $\sup_{p \leq \bar{p} - \delta_n} \sqrt{n}|\hat{Q}(p) - Q(p)| = O_p(1)$ and $\sqrt{n}(\hat{z} - z) = O_p(1)$. Then the first result in (15) follows since,

$$
\begin{aligned}
&\sup_{p\le\bar p-\delta_n}\sqrt{n}(\hat Q(p)-\hat z)\\
\le&\sup_{p\le\bar p-\delta_n}\sqrt{n}(\hat Q(p)-Q(p)-\hat z+z)+\sup_{p\le\bar p-\delta_n}\sqrt{n}(Q(p)-z)\\
\le&O_p(1)-\sqrt{n}\epsilon_n\to-\infty.
\end{aligned}
$$

Next, we derive the asymptotic properties of $\sqrt{n}(\hat P(p;\hat z)-P(p;z))$. Note that

$$
\begin{aligned}
&\sqrt{n}(\hat P(p;\hat z)-P(p;z))\\
=&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(w(X_i,\hat z,\hat Q(p))-E[w(X,\hat z,\hat Q(p))]\big)\\
&\quad+\sqrt{n}\big(E[w(X,\hat z,\hat Q(p))]-P(p,z)\big)\\
=&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(w(X_i,\hat z,\hat Q(p))-E[w(X,\hat z,\hat Q(p))]\big)\\
&\quad+\sqrt{n}\Big\{F(\min\{\hat Q(p),\hat z\})\hat z-G_F(F(\min\{\hat Q(p),\hat z\}))\\
&\qquad-\big[F(\min\{Q(p),z\})z-G(F(\min\{Q(p),z\}))\big]\Big\}\equiv I_1+I_2.
\end{aligned}
$$

Note that

$$
\begin{aligned}
I_1=&\Big(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(w(X_i,\hat z,\hat Q(p))-E[w(X,\hat z,\hat Q(p))]\big)\\
&\qquad-\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(w(X_i,z,Q(p))-E[w(X,z,Q(p))]\big)\Big)\\
&\qquad+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(w(X_i,z,Q(p))-E[w(X,z,Q(p))]\big)\\
=&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(w(X_i,z,Q(p))-E[w(X,z,Q(p))]\big)+o_p(1),
\end{aligned}\tag{16}
$$

where the $o_p(1)$ result holds because of the stochastic equicontinuity of the empirical process. This results holds uniformly over $p\in[0,p_u]$. For the second term, we need to consider three cases: (a) $p\le\bar p-\delta_n$, (b) $p>\bar p+\delta_n$ and (c) $|p-\bar p|<\delta_n$ because $F(\min\{s',s\})s-G(F(\min\{s',s\}))$ is not differentiable at $s=s'$. For case (a),

$\min\{Q(p), z\} = Q(p)$ and by (15), we have

$$
\begin{aligned}
&\sqrt{n}\big(F(\min\{\hat{Q}(p), \hat{z}\})\hat{z} - G(F(\min\{\hat{Q}(p), \hat{z}\})) \\
&\qquad - \big[F(\min\{Q(p), z\})z - G(F(\min\{Q(p), z\}))\big]\big) \\
=&\sqrt{n}\big(F(\hat{Q}(p))\hat{z} - G(F(\hat{Q}(p))) - \big[F(Q(p))z - G(F(Q(p)))\big]\big) + o_p(1) \\
=&\sqrt{n}\big(F(\hat{Q}(p))\hat{z} - F(Q(p))z\big) + \sqrt{n}\big(G(F(\hat{Q}(p))) - G(F(Q(p)))\big) + o_p(1) \\
=&\sqrt{n}\big(p \cdot (\hat{z} - z) + z \cdot f(Q(p))(\hat{Q}(p) - Q(p))\big) + o_p(1) \\
&\qquad - \sqrt{n}\big(Q(p) \cdot f(Q(p)) \cdot (\hat{Q}(p) - Q(p))\big) + o_p(1) \\
=&p \cdot \sqrt{n}(\hat{z} - z) + (z - Q(p)) \cdot f(Q(p))\sqrt{n}(\hat{Q}(p) - Q(p)) + o_p(1) \\
=&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(p \cdot \psi_z(X_i; z) - (z - Q(p)) \cdot (1(X_i \leq Q(p)) - p)\big) + o_p(1).
\end{aligned} \tag{17}
$$

Therefore, (16) and (17) together imply

$$
\begin{aligned}
\sup_{p \leq \bar{p} - \delta_n}\Big|&\sqrt{n}(\hat{P}(p; \hat{z}) - P(p; z)) - \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big(\big(w(X_i, z, Q(p)) - E[w(X, z, Q(p))]\big) \\
&\qquad + p \cdot \psi_z(X_i; z) - (z - Q(p)) \cdot (1(X_i \leq Q(p)) - p)\Big)\Big| = o_p(1).
\end{aligned} \tag{18}
$$

Similarly, for case (b), we have

$$
\begin{aligned}
&\sqrt{n}\big(F(\min\{\hat{Q}(p), \hat{z}\})\hat{z} - G(F(\min\{\hat{Q}(p), \hat{z}\})) \\
&\qquad - \big[F(\min\{Q(p), z\})z - G(F(\min\{Q(p), z\}))\big]\big) \\
=&\sqrt{n}\big(F(\hat{z})\hat{z} - G(F(\hat{z})) - \big[F(z)z - G(F(z))\big]\big) + o_p(1) \\
=&\sqrt{n}\big(F(\hat{z})\hat{z} - F(z)z\big) + \sqrt{n}\big(G(F(\hat{z})) - G(F(z))\big) + o_p(1) \\
=&\sqrt{n}\big((F(z) + zf(z)) \cdot (\hat{z} - z)\big) + o_p(1) - \sqrt{n}\big(zf(z) \cdot (\hat{z} - z)\big) + o_p(1) \\
=&F(z)\sqrt{n}(\hat{z} - z) + o_p(1).
\end{aligned} \tag{19}
$$

We also have

$$
\begin{aligned}
\sup_{p \geq \bar{p} + \delta_n}\Big|&\sqrt{n}(\hat{P}(p; \hat{z}) - P(p; z)) \\
&\qquad - \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big(w(X_i, z, Q(p)) - E[w(X, z, Q(p))] + F(z) \cdot \psi_z(X_i; z)\Big)\Big| = o_p(1).
\end{aligned} \tag{20}
$$

For case (c), it is straightforward to see that

$$\sup_{|p-\bar{p}|\leq\delta_n}\left|\sqrt{n}(\hat{P}(p;\hat{z})-P(p;z))\right| = O_p(1). \tag{21}$$

Note that for $p_u \leq p \leq 1$,

$$\sup_{p\geq p_u}|\sqrt{n}(\hat{P}(p;\hat{z})-P(p;z))-\sqrt{n}(\hat{P}(p_u;\hat{z})-P(p_u;z))| = o_p(1.)$$

To obtain the result in Proposition 4.4 let $p_1 = F_1(z_1)$ and $p_2 = F_2(z_2)$ and without loss of generality, we derive the result for the case where $p_1 = p_2 = p_m$, $z_1 = z_2$ and $P_1(p,z_1) = P_2(p,z_2)$ for all $p \in [0,1]$. Note that in this case, $P_o = [0,1]$. Proofs for cases where there exist points such that $P_2(p;z_2) < P_1(p;z_1)$ is similar to that of Lemma 2.1 of Donald and Hsu (2015). Note that

$$\begin{aligned}
\hat{T}_1 &= \sqrt{\frac{n_1n_2}{n_1+n_2}}\int_0^1 \max\left\{\hat{P}_2(p;\hat{z}_2)-\hat{P}_1(p;\hat{z}_1),0\right\}dp \\
&= \sqrt{\frac{n_1n_2}{n_1+n_2}}\int_{p_m-\delta_n}^{p_m+\delta_n} \max\left\{\hat{P}_2(p;\hat{z}_2)-\hat{P}_1(p;\hat{z}_1),0\right\}dp \\
&\quad+\sqrt{\frac{n_1n_2}{n_1+n_2}}\int_0^{p_m-\delta_n} \max\left\{\hat{P}_2(p;\hat{z}_2)-\hat{P}_1(p;\hat{z}_1),0\right\}dp\Big) \\
&\quad+\sqrt{\frac{n_1n_2}{n_1+n_2}}\int_{p_m+\delta_n}^1 \max\left\{\hat{P}_2(p;\hat{z}_2)-\hat{P}_1(p;\hat{z}_1),0\right\}dp \\
&= \hat{T}_{11}+\hat{T}_{12}+\hat{T}_{13}
\end{aligned} \tag{22}$$

Note that given that $\delta_n \to 0$ and

$$\sqrt{\frac{n_1n_2}{n_1+n_2}}\sup_{p\in[0,1]}\left|\hat{P}_2(p;\hat{z}_2)-\hat{P}_1(p;\hat{z}_1)\right| = O_p(1),$$

we have $\hat{T}_{11} = o_p(1)$. Also,

$$\begin{aligned}
\hat{T}_{12} &= \sqrt{\frac{n_1n_2}{n_1+n_2}}\int_0^{p_m-\delta_n} \max\left\{\frac{1}{n_2}\sum_{i=1}^{n_2}\psi_{P_2}(X_{2i};p)-\frac{1}{n_1}\sum_{i=1}^{n_1}\psi_{P_1}(X_{1i};p),0\right\}+o_p(1) \\
&= \sqrt{\frac{n_1n_2}{n_1+n_2}}\int_0^{p_m} \max\left\{\frac{1}{n_2}\sum_{i=1}^{n_2}\psi_{P_2}(X_{2i};p)-\frac{1}{n_1}\sum_{i=1}^{n_1}\psi_{P_1}(X_{1i};p),0\right\}+o_p(1) \\
&= \int_0^{p_m} \max\left\{\sqrt{\frac{n_1}{n_1+n_2}}\frac{1}{\sqrt{n_2}}\sum_{i=1}^{n_2}\psi_{P_2}(X_{2i};p)-\sqrt{\frac{n_2}{n_1+n_2}}\frac{1}{\sqrt{n_1}}\sum_{i=1}^{n_1}\psi_{P_1}(X_{1i};p),0\right\}+o_p(1) \\
&\xrightarrow{D} \int_0^{p_m} \max\left\{\sqrt{\lambda}\mathcal{P}_2(p)-\sqrt{1-\lambda}\mathcal{P}_1(p)\right\}dp
\end{aligned} \tag{23}$$

where the first equality follows from (18) and (20), the second equality holds for reasons similar to the result that $\hat{T}_{11}$ is $o_p(1)$. The last line follows from continuous mapping theorem and the fact that

$$\sqrt{\frac{n_1}{n_1+n_2}}\frac{1}{\sqrt{n_2}}\sum_{i=1}^{n_2}\psi_{P_2}(X_{2i};p)-\sqrt{\frac{n_2}{n_1+n_2}}\frac{1}{\sqrt{n_1}}\sum_{i=1}^{n_1}\psi_{P_1}(X_{1i};p)$$
$$\Rightarrow\sqrt{\lambda}\mathcal{P}_2(p)-\sqrt{1-\lambda}\mathcal{P}_1(p)$$

and $\int_{P_o}\{\cdot,0\}dp$ is a continuous functional. By the same argument, we can show that

$$\begin{aligned}\hat{T}_{13}&=\sqrt{\frac{n_1n_2}{n_1+n_2}}\int_{p_m+\delta_n}^{1}\max\Big\{\frac{1}{n_2}\sum_{i=1}^{n_2}\psi_{P_2}(X_{2i};p)-\frac{1}{n_1}\sum_{i=1}^{n_1}\psi_{P_1}(X_{1i};p),0\Big\}+o_p(1)\\&=\sqrt{\frac{n_1n_2}{n_1+n_2}}\int_{p_m+\delta_n}^{1}\max\Big\{\frac{1}{n_2}\sum_{i=1}^{n_2}\psi_{P_2}(X_{2i};p)-\frac{1}{n_1}\sum_{i=1}^{n_1}\psi_{P_1}(X_{1i};p),0\Big\}+o_p(1)\\&\xrightarrow{D}\int_{p_m}^{1}\max\big\{\sqrt{\lambda}\mathcal{P}_2(p)-\sqrt{1-\lambda}\mathcal{P}_1(p)\big\}dp.\end{aligned}$$

Therefore, we have

$$\hat{T}_1\xrightarrow{D}\int_0^1\max\big\{\sqrt{\lambda}\mathcal{P}_2(p)-\sqrt{1-\lambda}\mathcal{P}_1(p)\big\}dp.$$

Proof of Theorem 4.6: The proof of Theorem 4.6 (i) and (ii) is similar to that for Theorem 4.1 in Donald and Hsu (2013) except that we need to allow for the non-differentiability around $F_1(z_1)$ and $F_2(z_2)$. This can be handled with the same argument as in Proposition 4.4. The proof of Theorem 4.6 (i) is similar to that for Theorem 4.2 in Donald and Hsu (2013).

Proof of Theorem 4.9: Let $\mathcal{H}_1$ denote the set of all functions from $[0,1]$ to $[-\infty,0]$. Let $h=(h_1,h_2)$, where $h_1\in\mathcal{H}_1$ and $h_2\in\mathcal{H}_2$, and define

$$T(h)=\int_0^1\max\{(\Psi_{h_2}(z)+h_1(z)),0\}dp.$$

Define $c_0(h_1,h_2,1-\alpha)$ as the $(1-\alpha)$-th quantile of $T(h)$. The key is to show that for any sequence of $(\theta_{1,\ell_n},\theta_{2,\ell_n})\in\{\Theta_0^2|h_{2,\theta_1},h_{2,\theta_2}\in\mathcal{H}_{2,cpt}\}$, there is a further subsequence k_n of ℓ_n such that (a) $\big((F_{1,k_n},z_{1,k_n}),F_{2,k_n},z_{2,k_n})\big)\to\big((F_1^*,z_1^*),F_2^*,z_2^*)\big)$ such that the null hypothesis holds, and (b) $h_{2,\theta_{1,k_n}}\to h_{2,1}^*\in\mathcal{H}_{2,cpt}$ and $h_{2,\theta_{2,k_n}}\to h_{2,2}^*\in\mathcal{H}_{2,cpt}$. Note that (a) is implied by Assumption 4.8 since for all n, $(F_{1,\ell_n},F_{2,\ell_n})$ belongs to a compact subset by the Arzelà-Ascoli Theorem, e.g., Theorem 6.2.61 of Corbae, Stinchcombe and Zeman (2009). So does $(z_{1,\ell_n},z_{2,\ell_n})$. Also (b) holds because we impose $\mathcal{H}_{2,cpt}$.

Define $\tilde{h}_2^{\theta_1}(p', p'') = h_2^{\theta_1}(p', p'')$ if $p', p'' \in [0, F_1(z_1)]$, $\tilde{h}_2^{\theta_1}(p', p'') = h_2^{\theta_1}(p', 1)$ if $p' \in [0, F_1(z_1)]$ and $p'' \in (F_1(z_1), 1]$, and $\tilde{h}_2^{\theta_1}(p', p'') = h_2^{\theta_1}(p', 1)$ if $p', p'' \in (F_1(z_1), 1]$. Also, $\tilde{h}_2^{\theta_1}(p', p'') = \tilde{h}_2^{\theta_1}(p'', p')$. It is true that if $h_{2,\theta_{1,k_N}} \to h^*_{2,\theta_1^*} \in \mathcal{H}_{2,cpt}$, then $\tilde{h}_{2,\theta_{1,k_n}} \to \tilde{h}^*_{2,\theta_1^*}$. Define $h_{1,n}^{\theta_1,\theta_2} = \sqrt{n_1 n_2/(n_1+n_2)}(P_1(p) - P_2(p))$ and $h_{1,n}^{\theta_1,\theta_2}$ belongs to $\mathcal{H}_1$ under the null hypothesis. Also, $\tilde{h}_2^{\theta_1,\theta_2} = \lambda \cdot \tilde{h}_{2,\theta_1} + (1-\lambda)\tilde{h}_{2,\theta_2}$. Note that under pointwise asymptotics, $\tilde{h}_2^{\theta_1,\theta_2}$ is the covariance kernel of the limiting Gaussian processes in the integral in Proposition 4.4.

As in Lemma A2 of Andrews and Shi (2013), we can show that for any $\tilde{\delta} > 0$,

$$\limsup_{n\to\infty} \sup_{(\theta_1,\theta_2)\in\{\Theta_0^2 | h_{2,\theta_1}, h_{2,\theta_2} \in \mathcal{H}_{2,cpt}\}} P\Big(\widehat{T}^1 > c_0\big(h_{1,n}^{\theta_1,\theta_2}, \tilde{h}_2^{\theta_1,\theta_2}, 1-\alpha\big) + \delta\Big) \leq \alpha, \quad (24)$$

Also, as in Lemma A3 of Andrews and Shi (2013), we can show that for all $\alpha > 0$

$$\limsup_{n\to\infty} \sup_{(\theta_1,\theta_2)\in\{\Theta_0^2 | h_{2,\theta_1}, h_{2,\theta_2} \in \mathcal{H}_{2,cpt}\}} P\Big(c_0\big(\sqrt{\frac{n_1 n_2}{n_1+n_2}}\phi_n, \tilde{h}_2^{\theta_1,\theta_2}, 1-\alpha\big) < c_0(h_{1,n}^{X,Y}, h_2^{\theta_1,\theta_2}, 1-\alpha)\Big) = 0. \quad (25)$$

Note that the discontinuity issue can be handled in the same fashion as in the proof of Theorem 4.6.

To complete the proof, we can follow Donald and Hsu (2013) to show that for all $0 < \tilde{\delta} < \eta$,

$$\limsup_{n\to\infty} \sup_{(\theta_1,\theta_2)\in\{\Theta_0^2 | h_{2,\theta_1}, h_{2,\theta_2} \in \mathcal{H}_{2,cpt}\}} P\Big(\hat{c}^{1,u}_{\eta,n} < c_0(\sqrt{\frac{n_1 n_2}{n_1+n_2}}\phi_n, \tilde{h}_2^{\theta_1,\theta_2}, 1-\alpha) + \tilde{\delta}\Big) = 0. \quad (26)$$

The result in Theorem 4.9 can then be shown by combining (24), (25) and (26). The proof for the second part is identical to that for the second part of Theorem 6.1 of Donald and Hsu (2013).

Proof of (13): Note that uniformly over $p \in [0,1]$,

$$
\begin{aligned}
&\sqrt{n_j}(\hat{P}_j^R(p,\hat{z}_j) - P_j^R(p,z_j)) \\
=&\sqrt{n_j}\Big(\frac{\hat{P}_j(p,\hat{z}_j)}{\hat{z}_j} - \frac{P_j(p,z_j)}{z_j}\Big) \\
=&\sqrt{n_j}\Big(\big(\frac{1}{\hat{z}_j} - \frac{1}{z_j}\big)\hat{P}_j(p,\hat{z}_j) + \frac{1}{z_j}\big(\hat{P}_j(p,\hat{z}_j) - P_j(p,z_j)\big)\Big) \\
=&\sqrt{n_j}\frac{1}{z_j}\big(\hat{P}_j(p,\hat{z}_j) - P_j(p,\hat{z}_j)\big) + \sqrt{n_j}P_j(p,z_j)\big(\frac{1}{\hat{z}_j} - \frac{1}{z_j}\big) + o_p(1) \\
=&\sqrt{n_j}\frac{1}{z_j}\big(\hat{P}_j(p,\hat{z}_j) - P_j(p,z_j)\big) - \sqrt{n_j}\frac{P_j(p,z_j)}{z_j^2}\big(\hat{z}_j - z_j\big) + o_p(1) \\
=&\sqrt{n_j}\frac{1}{z_j}\big[\hat{P}_j(p,\hat{z}_j) - P_j(p,z_j) - P_j^R(p,z_j)\big(\hat{z}_j - z_j\big)\big] + o_p(1),
\end{aligned}
$$

where the third equality holds because $\sup_{p\in[0,1]} |\hat{P}_j(p,\hat{z}_j) - P_j(p,z_j)| = o_p(1)$, the fourth equality is obtained by applying the delta method on $1/\hat{z}_j$ and the last equality holds because $P_j^R(p,z_j) = P_j(p,z_j)/z_j$. Then by the same argument for Proposition 4.4, the result follows.

References

Andrews, D. W. and X. Shi (2013), "Inference Based on Conditional Moment Inequalities," *Econometrica*, 81, 609-666.

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

Barrett, G. F. and S. G. Donald (2003), "Consistent Tests for Stochastic Dominance," *Econometrica*, 71, 71-104.

Barrett, G. F. and S. G. Donald (2009), "Statistical Inference with Generalized Gini Indices of Inequality, Poverty and Welfare," *Journal of Business and Economic Statistics*, 27, 1-17.

Barrett, G. F., S. G. Donald and D. Bhattacharya (2014), "Consistent Nonparametric Tests for Lorenz Dominance," *Journal of Business and Economic Statistics*, 32, 1-13.

Bhattacharya, D. (2005), "Asymptotic Inference from Multi-stage Samples," *Journal of Econometrics*, 126, 145-171.

Corbae, D., M. B. Stinchcombe, and J. Zeman (2009). *An Introduction to Mathematical Analysis for Economic Theory and Econometrics,* Princeton University Press.

Dardanoni, V. and A. Forcina (1999), "Inference for Lorenz curve 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.

Donald, S. G. and Y.-C. Hsu (2013), "Improving the Power of Tests of Stochastic Dominance," forthcoming in *Econometric Reviews.*

Jenkins, S. P. (1991), "Aggregation Issues in Earnings Discrimination Measurement," Economics Discussion Paper ,University of Bath.

Jenkins, S. P. (1994), "Earnings Discrimination Measurement: a Distributional Approach ," textitJournal of Econometrics, 61, 81-102.

Jenkins, S. P. and P. J. Lambert (1997), "Three 'I's of Poverty Curves, With an Analysis of UK Poverty Trends," *Oxford Economic Papers*, 49, 317-327.

Jenkins, S. P. and P. J. Lambert (1998a), “Ranking Poverty Gap Distributions: Further TIPs for Poverty Analysis,” *Research on Economic Inequality*, 8, 31-38.

Jenkins, S. P. and P. J. Lambert (1998b), “Three ‘I’s of Poverty Curves and Poverty Dominance: TIPs for Poverty Analysis,” *Research on Economic Inequality*, 8, 39-56.

Linton, O., E. Maasoumi and Y.-J. Whang (2005), “Consistent Testing for Stochastic Dominance: A Subsampling Approach,” *Review of Economic Studies*, 72, 735-765.

Linton, O., K. Song and Y.-J. Whang (2010), “An Improved Bootstrap Test of Stochastic Dominance,” *Journal of Econometrics*, 154, 186-202.

McFadden, D. (1989), “Testing for Stochastic Dominance” in *Studies in the economics of uncertainty: In honor of Josef Hadar,* Fomby,T-B, and Seo,T-K, eds. New York; Berlin; London and Tokyo: Springer, 113-34.

Ravallion, M. (1994), *Poverty Comparisons*, Harwood Academic Publishers, Chur Switzerland.

Sen, A. (1976), “Poverty: an Ordinal Approach to Measurement,” *Econometrica,* 44, 219-231.

Shorrocks, A. F. (1995), “Revisiting the Sen Poverty Index,” *Econometrica*, 63 1225-1230.

Shorrocks, A. F. (1998), “Deprivation Profiles and Deprivation Indices” in *The distribution of welfare and household production: International perspectives,* Jenkins, S. P., A. Kapteyn and B. van Praag eds. Cambridge, Cambridge University Press.

Van der Vaart, A. W. and J. A. Wellner (1996), *Weak Convergence and Empirical Processes: With Applications to Statistics*, Springer-Verlag New York.

Yitzhaki, S. (1999). “Necessary and Sufficient Conditions for Dominance Using Generalized Lorenz curves,”In: *Advances in Econometrics*, Income Distribution and Scientific Methodology: Essays in Honor of Camilo Dagum ,D .J . Slottje (eds.), Physica-Verlag: Heidelberg.

Figure 3A: Aus HES Consumption Poverty Gap Profile

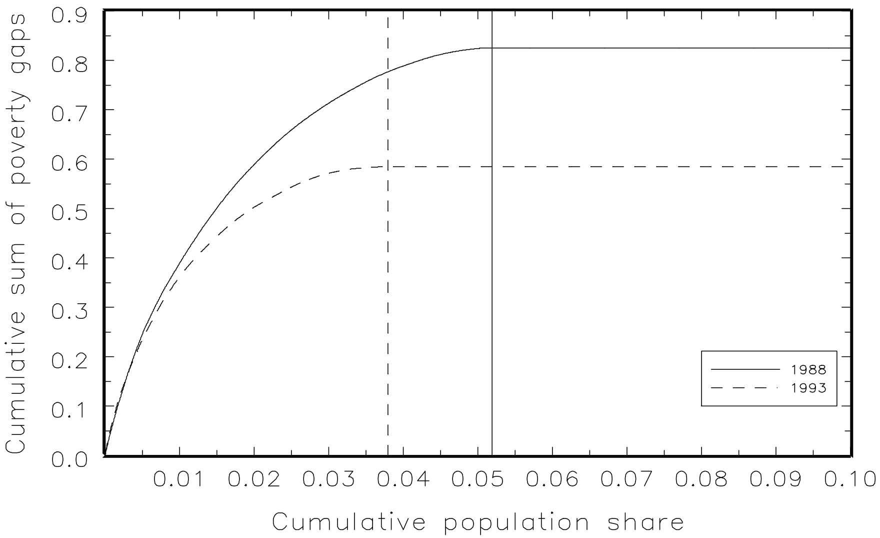


Figure 4A: Aus HES Consumption Normalized Poverty Gap Profile

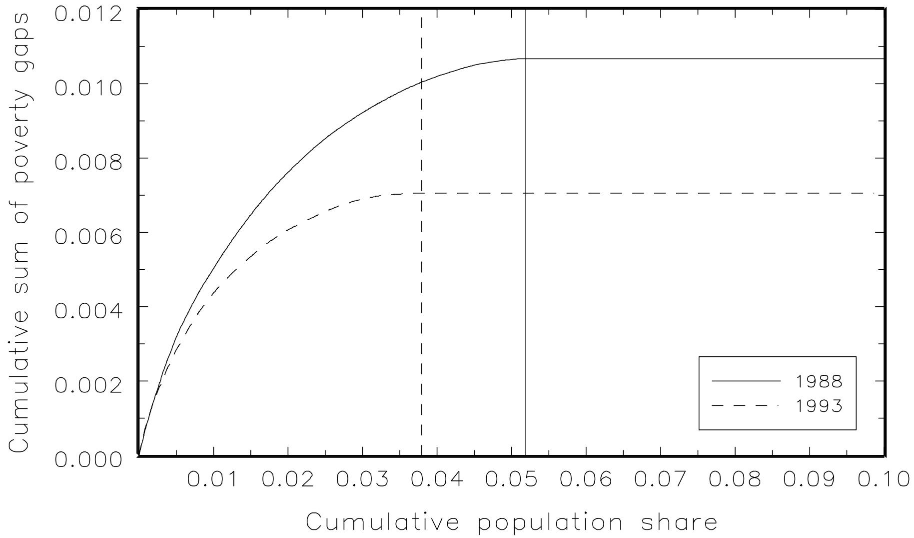


Figure 3B: Difference in Poverty Gap Profiles

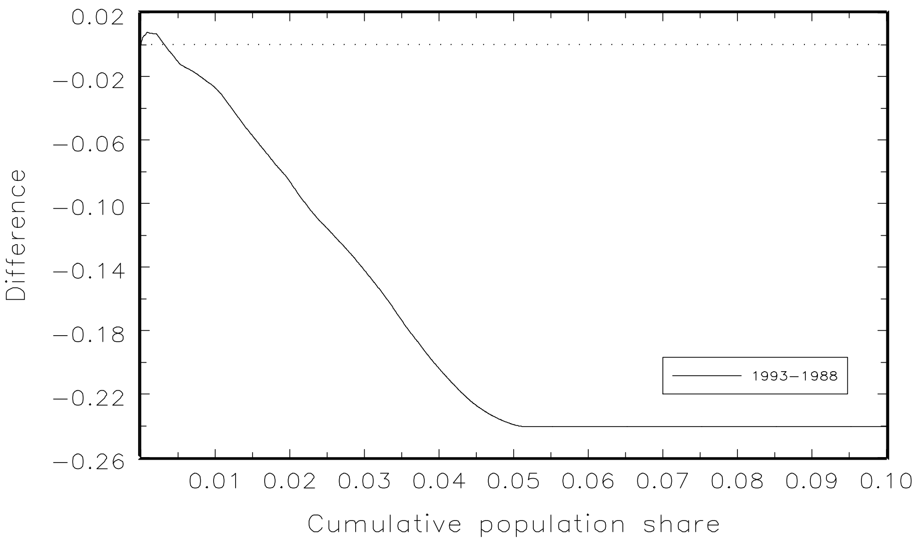


Figure 4B: Difference in Normalized Poverty Gap Profiles

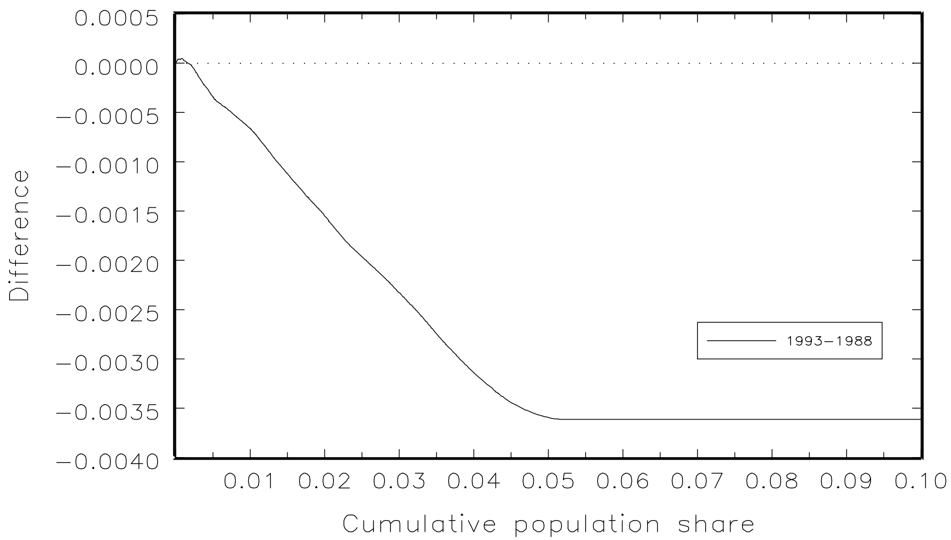


Figure 5A: Aus HES Consumption Poverty Gap Profile

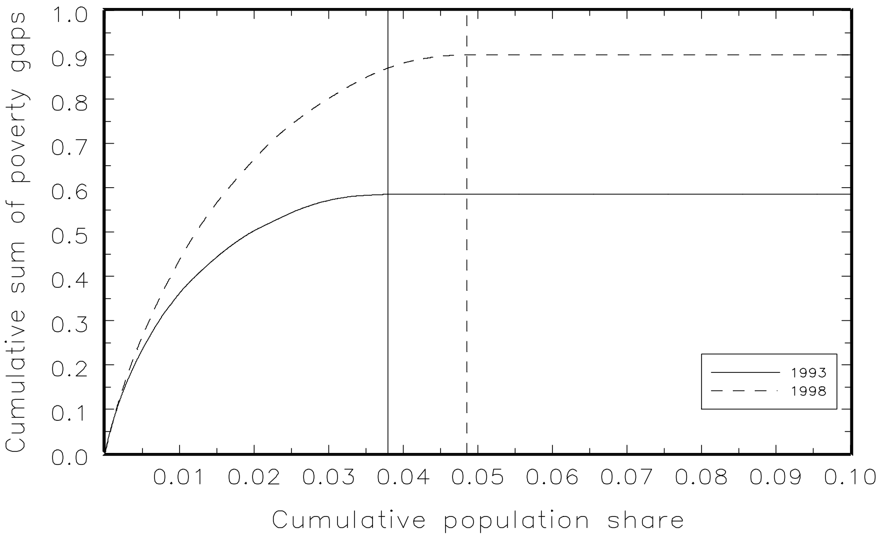


Figure 6A: Aus HES Consumption Normalized Poverty Gap Profile

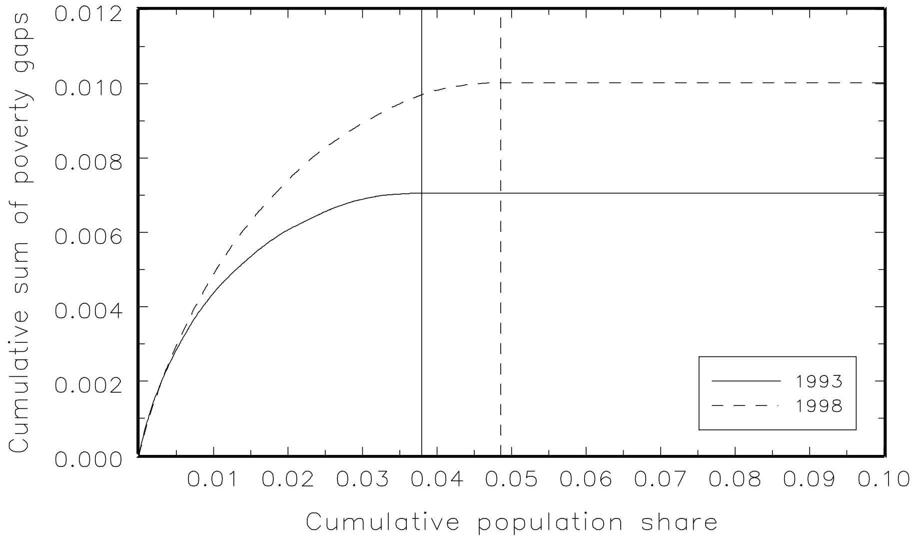


Figure 5B: Difference in Poverty Gap Profiles

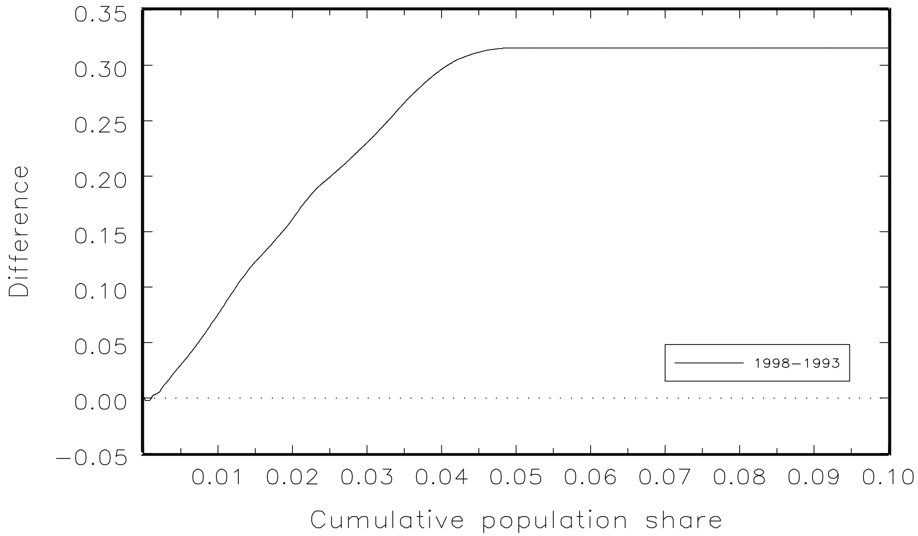


Figure 6B: Difference in Normalized Poverty Gap Profiles

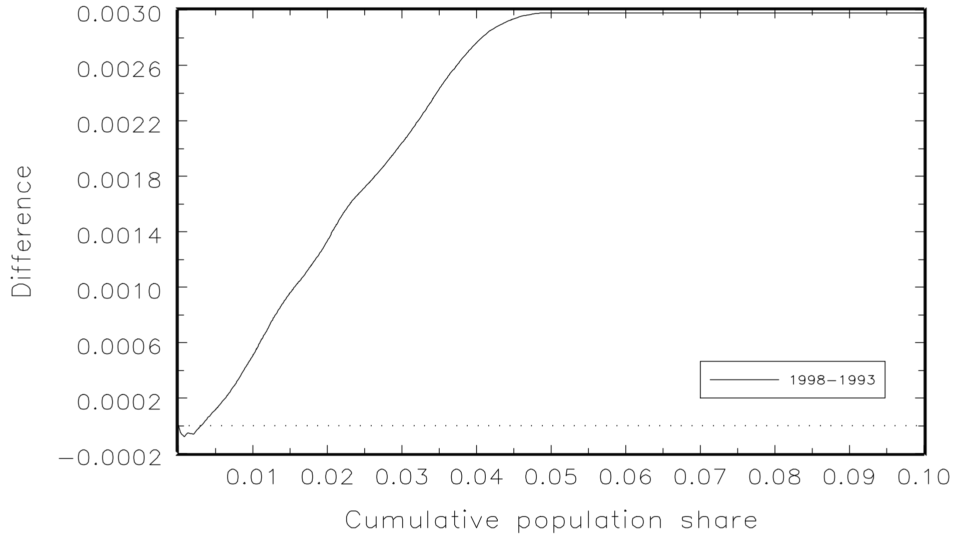


Figure 7A: Aus HES Consumption Poverty Gap Profile

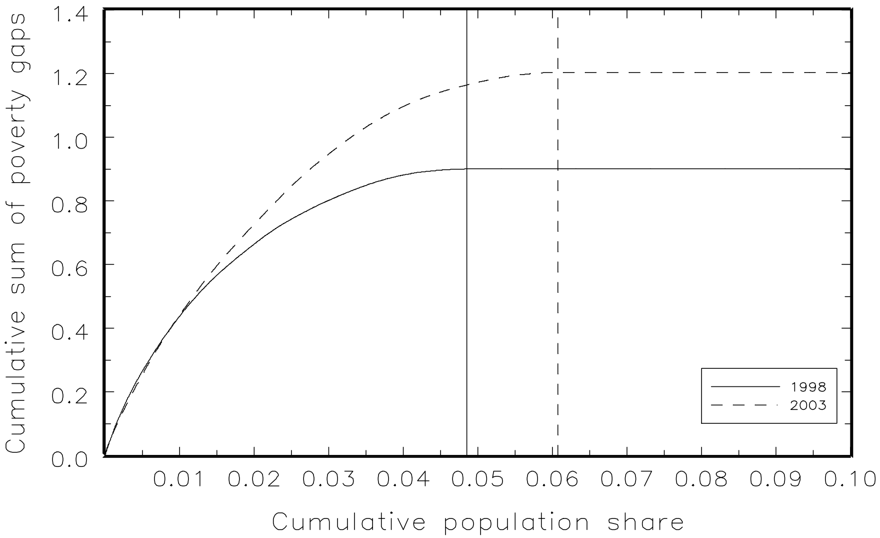


Figure 8A: Aus HES Consumption Normalized Poverty Gap Profile

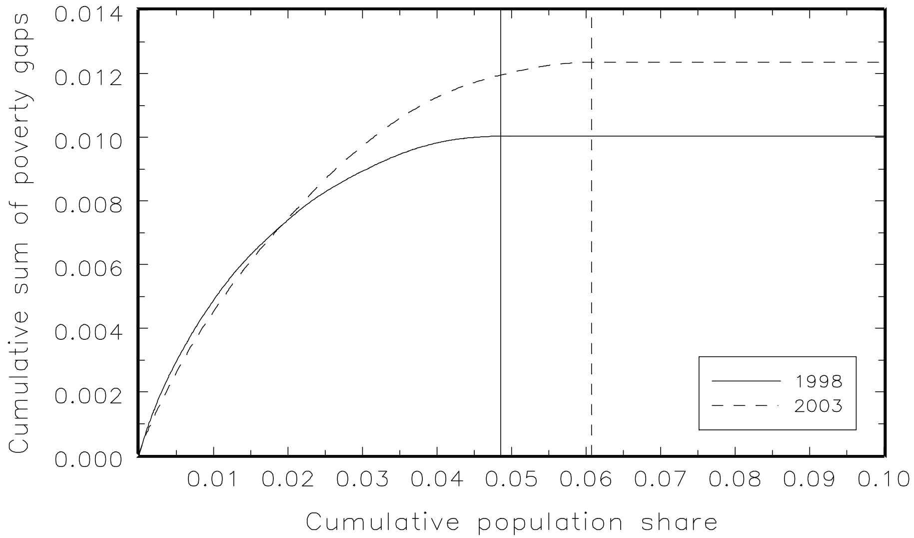


Figure 7B: Difference in Poverty Gap Profiles

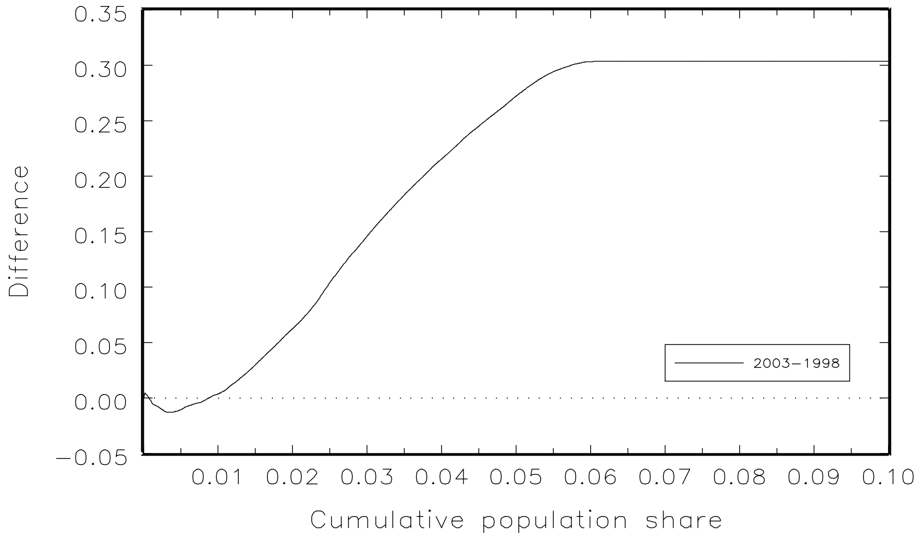


Figure 8B: Difference in Normalized Poverty Gap Profiles

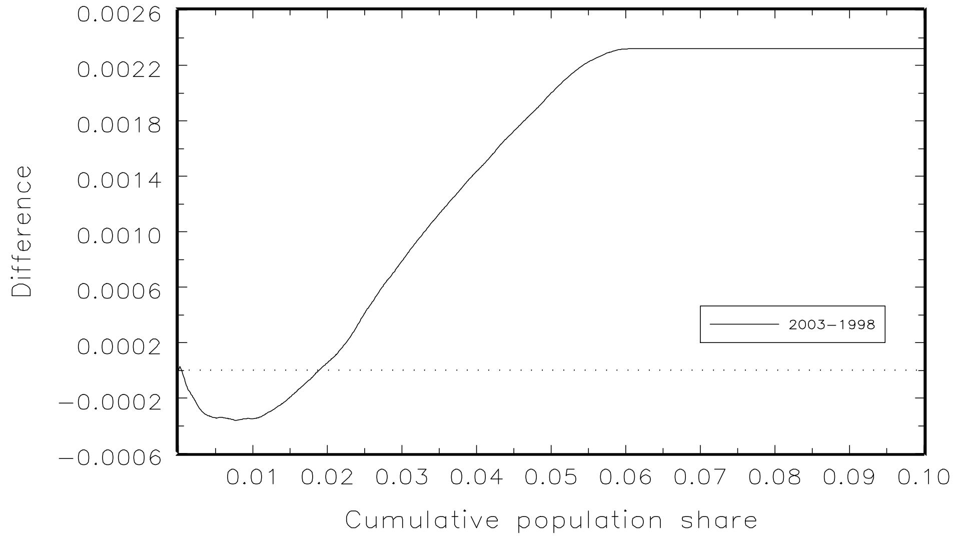


Figure 9A: Aus HES Consumption Poverty Gap Profile

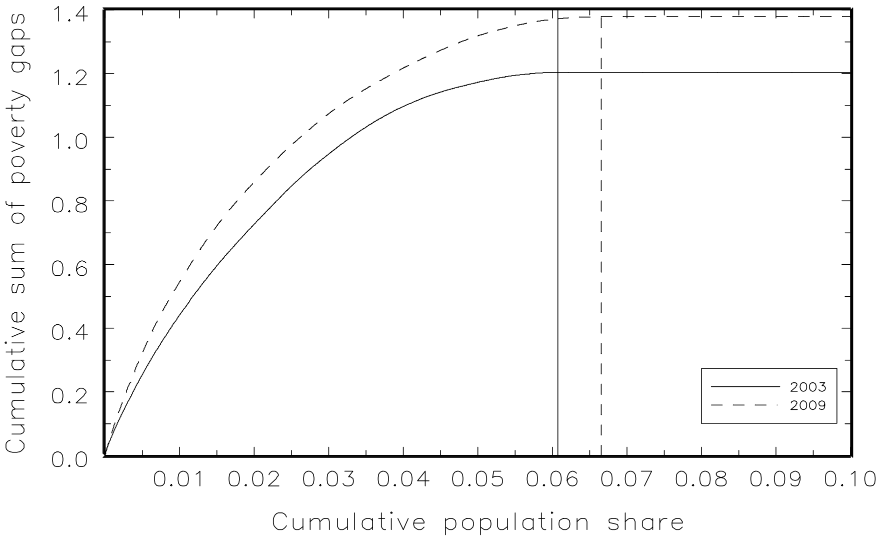


Figure 10A: Aus HES Consumption Normalized Poverty Gap Profile

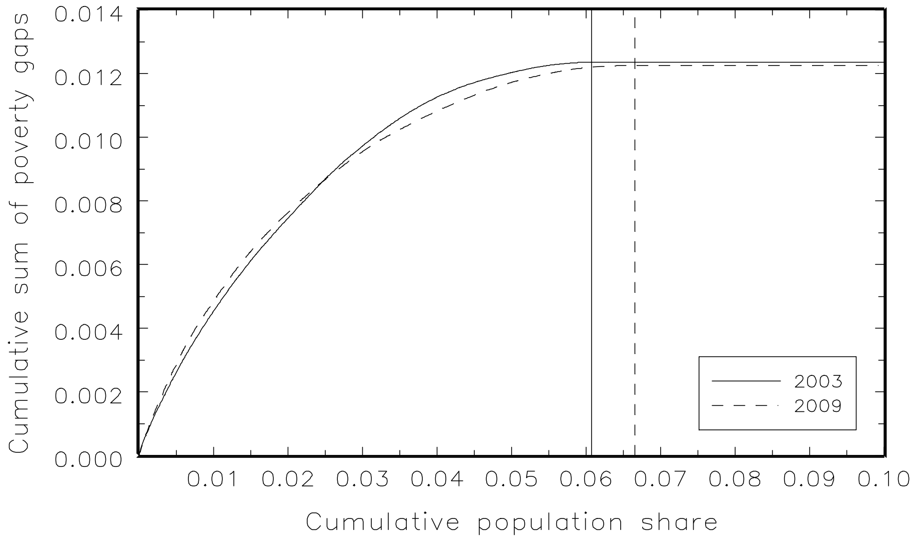


Figure 9B: Difference in Poverty Gap Profiles

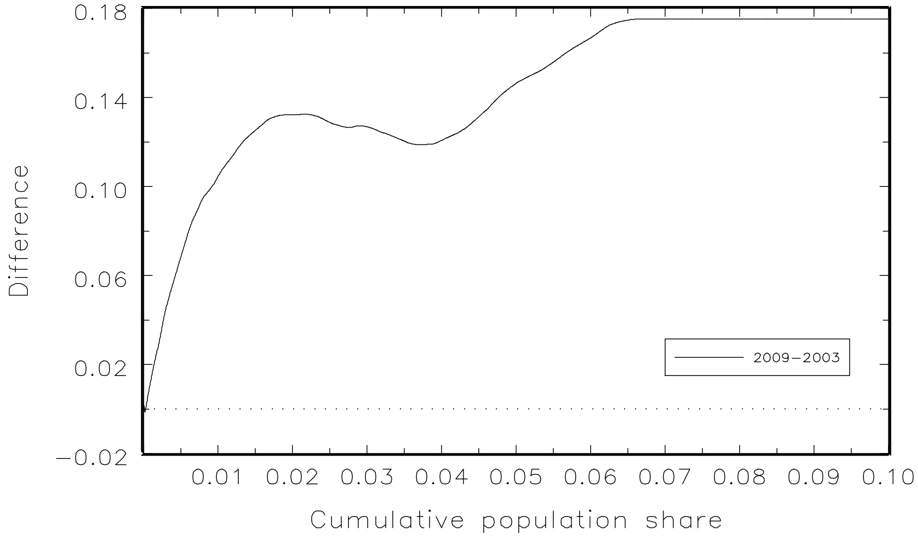
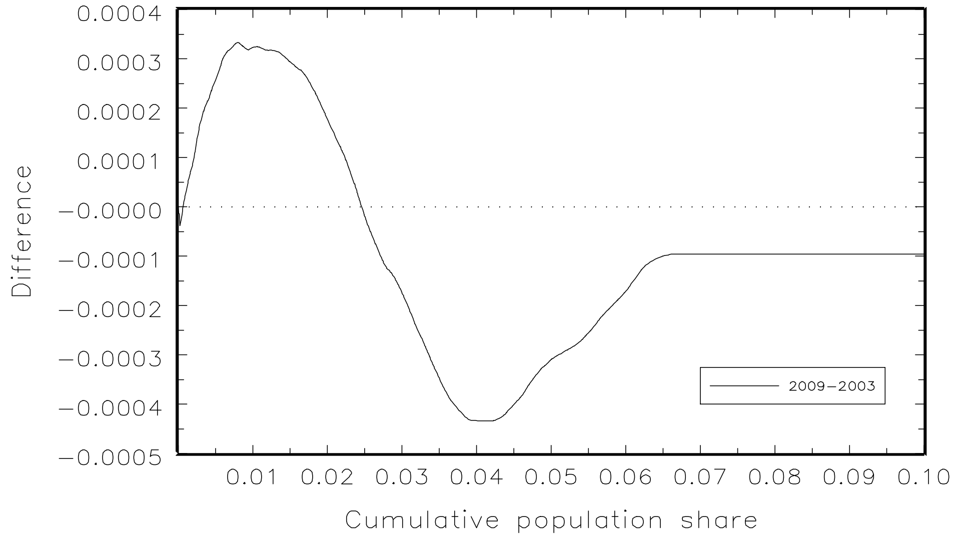


Figure 10B: Difference in Normalized Poverty Gap Profiles



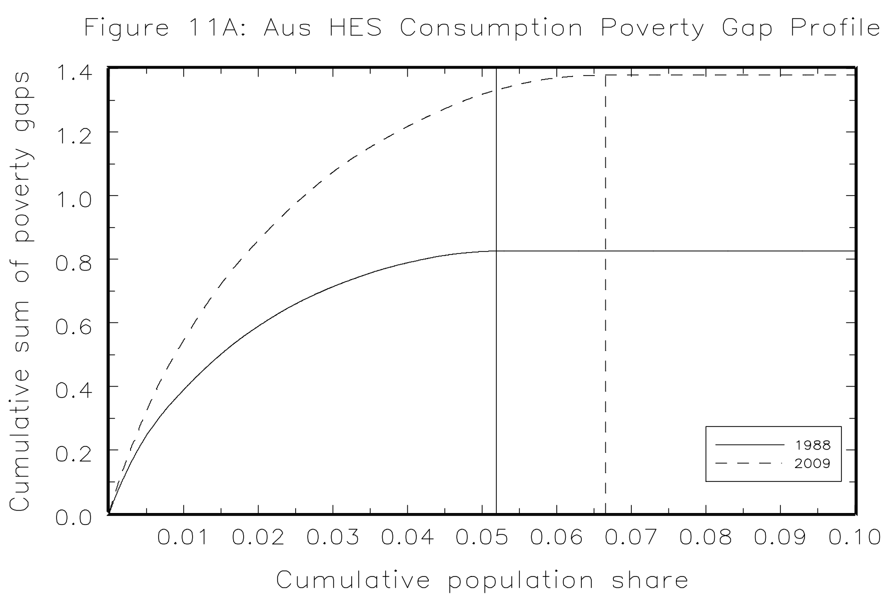


Figure 11A: Aus HES Consumption Poverty Gap Profile
Cumulative sum of poverty gaps
Cumulative population share
1988
2009

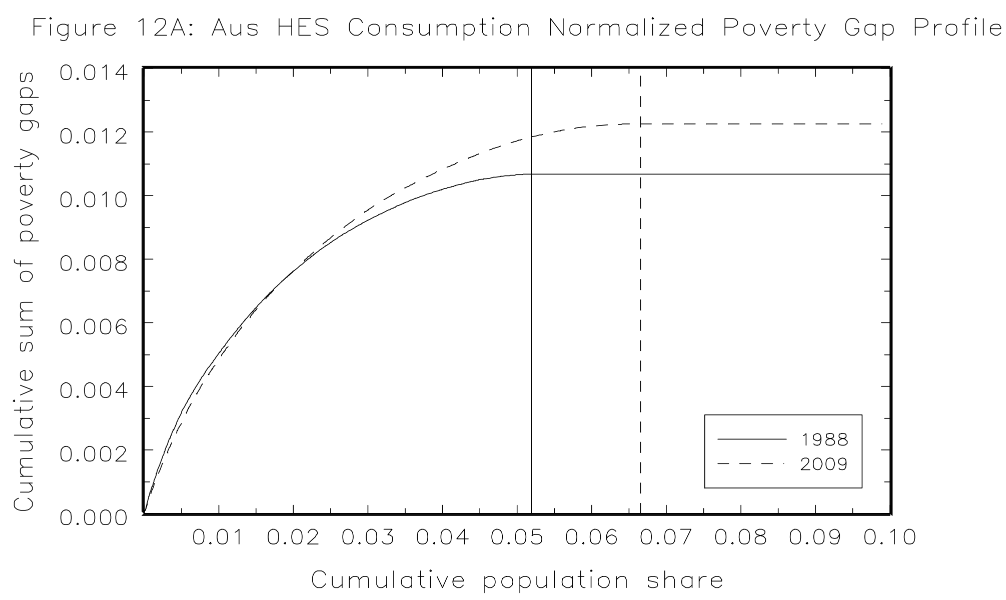


Figure 12A: Aus HES Consumption Normalized Poverty Gap Profile
Cumulative sum of poverty gaps
Cumulative population share
1988
2009

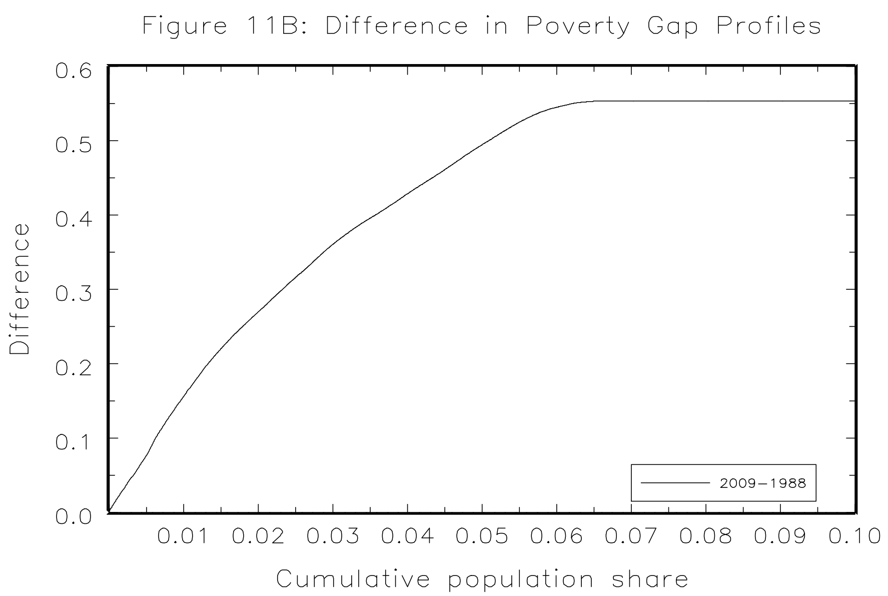


Figure 11B: Difference in Poverty Gap Profiles
Difference
Cumulative population share
2009–1988

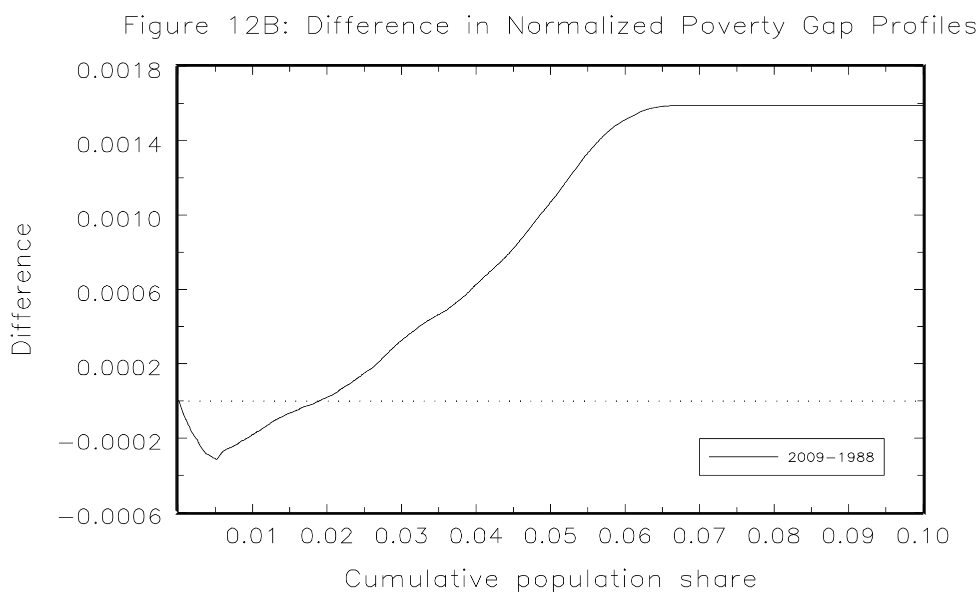


Figure 12B: Difference in Normalized Poverty Gap Profiles
Difference
Cumulative population share
2009–1988