CONSISTENT AND ASYMPTOTICALLY UNBIASED MINP TESTS OF MULTIPLE INEQUALITY MOMENT RESTRICTIONS

by

Christopher J. Bennett

Working Paper No. 09-W08

July 2009

DEPARTMENT OF ECONOMICS
VANDERBILT UNIVERSITY
NASHVILLE, TN 37235

www.vanderbilt.edu/econ

Consistent and Asymptotically Unbiased MinP Tests of Multiple Inequality Moment Restrictions*

Christopher J. Bennett
Department of Economics
Vanderbilt University, Nashville, TN 37240
E-mail: `chris.bennett@vanderbilt.edu`

July 7, 2009

Abstract

This paper considers the general problem of testing multiple inequality moment restrictions against an unrestricted alternative. We first introduce a test based on a maximum statistic and show how, via a partially recentered bootstrap scheme, we may obtain a testing procedure that delivers, at least asymptotically, an exact α-level test for *any* configuration of the parameters on the boundary of the null hypothesis. We prove that this bootstrap test is asymptotically unbiased and that it weakly dominates analogous testing procedures based on the canonical (fully centered) bootstrap. Building on these results we introduce a computationally inexpensive minimum p-value test. The minimum p-value test enjoys the asymptotic unbiasedness property of the underlying partially recentered bootstrap test. Additionally, the minimum p-value test delivers balance of power among the individual moment inequalities under test without studentization, and also allows users to gauge the strength of the evidence against the individual moment inequalities. To illustrate the use of our proposed testing procedure we examine the distributional effects of Vietnam veteran status on earnings. In particular, the results from our procedure when applied to testing for stochastic dominance and normalized stochastic dominance demonstrate that there is unambiguously greater poverty *and* greater relative inequality in earnings for veterans.

KEYWORDS: Bootstrap; Moment Inequalities; Asymptotic Bias, Stochastic Dominance
JEL CLASSIFICATION: C12, C14, I32

*I would like to thank participants of the 2009 Econometric Society summer meetings and the OPHI Workshop on Robustness Methods for Multidimensional Welfare Analysis for helpful comments. I also would like to thank Yanqin Fan, James Foster, and Zhiao Shi for their comments and support.

1. INTRODUCTION

In this paper we propose a new minimum p-value (MinP) test for testing multiple inequality moment restrictions against an unrestricted alternative. Our proposed testing procedure is endowed with a number of attractive properties which, when taken together, are not possessed by either minimum distance or existing MinP testing procedures. First, under mild regularity conditions our test is shown to be asymptotically valid uniformly on the boundary of the null and to be asymptotically unbiased. These properties are associated with the enhanced ability of our testing procedure to detect violations of the moment inequalities, especially when the number of moment inequalities under test is large. Second, our procedure delivers a balanced test without the need for studentizing the individual moment restrictions. Third, our testing procedure yields multiplicity-adjusted p-values which allow users to gauge the strength of the evidence against individual moment restrictions without having to pre-specify a target Type I error rate α. Lastly, and perhaps most importantly from a practical standpoint, our testing procedure is far less computationally demanding than the conventional bootstrap implementation of MinP-type tests in which a double bootstrap is generally required.

The large class of the problems to which our proposed testing procedure applies can be described as follows: Given a measurable $r \times 1$ vector valued function $\psi : \mathbb{R}^m \times \Theta \to \mathbb{R}^r$, the hypothesis to be tested takes the form[1]

$$
\begin{gathered}
H_0 : E_P[\psi_{(i)}(X, \theta)] \leq 0 \text{ for all } 1 \leq j \leq r \\
\text{against} \\
H_1 : E_P[\psi_{(i)}(X, \theta)] > 0 \text{ for some } 1 \leq j \leq r,
\end{gathered}
\tag{1}
$$

where X denotes a m-dimensional random vector generated according to the true probability mechanism P, and the true value of the parameter $\theta(P) \in \Theta \subset \mathbb{R}^k$ is unknown but may be

[1] Our focus here is on multiple one-sided tests, however the MinP procedure generalizes in a straightforward manner to accommodate both one-sided and two-sided tests.

consistently estimated. Since our interest centers on tests of the moment conditions and not θ *per se*, our testing problem differs from the recent literature on inferential procedures for θ when θ is point or partially identified by inequality moment conditions; see, for example, Andrews and Soares (2007), Andrews and Jia (2008), Chernozhukov, Hong and Tamer (2007), and references therein. As the following examples are intended to illustrate, the class of testing problems under consideration is nevertheless general and encompasses a wide range of potential applications in econometrics.

Example 1 (Moon and Schorfheide (2008)). *Economic theory may suggest a number of overidentifying moment inequality restrictions. A test of $E_P[\psi(X,\theta)] \leq 0$, where θ is replaced by $\hat{\theta}$—an estimator of θ that does not impose the inequality restrictions—may be viewed as an LM-type test of the moment inequality restrictions.*

Example 2 (Zheng, Formby, Smith and Chow (2000)). *Let P denote the probability measure associated with $X = (Y, Z)$. Also, assume that Y and Z have marginals F and G with common support $[0, \infty)$, and let $\theta = (\mu(F), \mu(G))$. The system of inequalities*

$$E_P[\psi_i(X,\theta)] = \int_0^{t_i} [F(\mu_F t) - G(\mu_G t)]\, dt \leq 0$$

for a fixed grid $(t_1, \ldots, t_r)$ constitute a set of necessary conditions for (weak) second-order normalized *stochastic dominance (SND) of F over G.*[2]

Example 3 (Dufour (1989)). *Consider the translog production function*

$$Y_i = \theta_0 + \sum_{j=1}^{2} \theta_j Z_{ij} + \theta_3 Z_{i1}^2 + \theta_4 (2 Z_{i1} Z_{i2}) + \theta_5 Z_{i2}^2 + U_i \ \ (i = 1, \ldots, n)$$

Testing for concavity of the production function amounts to testing

$$H_0 : \theta_3 \leq 0, \theta_5 \leq 0, \ \text{and } \theta_4^2 - \theta_3\theta_5 \leq 0$$

[2]Second-order normalized stochastic dominance is Lorenz consistent in that it yields the same partial ordering on the space of distributions; see, e.g., Foster and Sen (1999, pp. 142-148) for further details.

against an unrestricted alternative. In this case we may define

$$\psi(X,\theta) = \begin{pmatrix} \theta_3 \\ \theta_5 \\ \theta_4^2 - \theta_3\theta_5 \end{pmatrix}.$$

Example 4 (White (2000))**.** *Let $f_0(X,\theta)$ denote a benchmark model and $f_i(X,\theta)$ for $1 \leq i \leq r$ denote competing model specifications. For a given loss function $L(\cdot)$ a test of superior performance of a competitor to the benchmark may be framed as*

$$H_0 : E_P[L(f_0(X,\theta)) - L(f_i(X,\theta))] \leq 0$$

for $1 \leq i \leq r$. In this case by defining

$$\psi_i(X,\theta) = L(f_0(X,\theta)) - L(f_i(X,\theta))$$

we see that testing for model superiority with respect to $L(\cdot)$ can also be treated within our setup.

Variants of the testing problem defined in (1) have received significant attention in the literature. Wolak (1987) proposes a test in the context of a linear regression model with linear equality and inequality constraints on the model parameters. Kodde and Palm (1986), Wolak (1989), and Wolak (1991), treat the more general case of testing nonlinear inequality restrictions in nonlinear models. In each of these cases, the tests are based on (or are asymptotically equivalent to) a statistic which is computed as the minimum distance between the estimated parameter and the set containing parameter configurations satisfying the null hypothesis. Generally the asymptotic null distributions are a complex mixture of χ^2 distributions for any parameter configuration on the boundary of the null; see Perlman (1969) for a development of the underlying theory. Due to the composite nature of the tests, the critical values are determined by the asymptotic distribution under the least favourable configuration (LFC). In the linear model, the derivation of the LFC asymptotic null distribution and the computation of the relevant critical value is, at least in principle, feasible.[3] However, more serious complications can arise whenever nonlinearities are present in either the inequalities or the model. In particular, unlike the linear case where the LFC is uniquely defined by the configuration in which all of the inequalities are binding, the nonlinear cases are plagued by situations in which the binding of inequalities does uniquely identify a parameter configuration, and in which the LFC may occur on the boundary of the null at a point where one or more inequalities do not bind. These complications, among others, and their consequences for the practical implementation of minimum distance tests are discussed in detail in Wolak (1991) and Silvapulle and Sen (2004).

An obvious alternative to the minimum distance tests is the application of multiple testing procedures; see, e.g., Savin (1984). Since H_0 is true if and only if all of the individual inequality restrictions are true, the hypothesis in (1) is equivalent to

$$H_0 : \max_{1 \leq i \leq r} \Psi_{(i)} \leq 0 \text{ against } H_1 : \max_{1 \leq i \leq r} \Psi_{(i)} > 0, \tag{2}$$

where $\Psi_{(i)}$ denotes the ith element $E_P[\psi(X, \theta)]$. Formulated in this way it is a simple exercise to obtain bounds on the p-value for an asymptotic test. For example, letting t_i denote a studentized version of $\Phi_{n,(i)}$—the empirical counterpart of $\Phi_{(i)}$—we have, under mild regularity conditions and via the Bonferonni inequality, that

$$\lim_{n \to \infty} P\left(\max_{1 \leq i \leq r} t_i \leq z_\delta\right) \geq 1 - \delta r$$

where z_δ is the $(1 - \delta)$ quantile of the standard normal distribution. Then choosing $\delta = \alpha/r$

[3] See Wolak (1989) p. 215 for a discussion of the computational difficulties associated with computing the covariance-dependent weights of the Chi-bar distribution. Note that Dufour (1989) provides a bounds test that allows one to avoid explicit calculation of the probabilities.

implies that the asymptotic familywise error rate is bounded by α, i.e.

$$Pr(\text{Reject as least one } H_i | \text{ all } H_i \text{ are true}) \leq \alpha.$$

Although the use of probability inequalities is simple, the resulting tests are conservative. Moreover, in the presence of inequality constraints, the extent of under-rejection under the null increases with the number of non-binding constraints. To see this, suppose that H_0 is true but q of the r inequalities do not bind. In this case

$$\begin{aligned}\lim_{n\to\infty} P\left(\max_{1\leq i\leq r} t_i \leq z_\delta\right) &\geq 1-\delta(r-q) \\ &= 1-\alpha\frac{(r-q)}{r}\end{aligned}$$

and it becomes clear that the degree of conservatism is increasing in q.[4]

To avoid the use of probability inequalities, bootstrap methods can be used to deliver, at least asymptotically, an exact α-level test in the LFC. This approach is taken, for instance, by White (2000) in the context of testing for superior performance of a given model relative to a set of benchmark models. The basic idea when applied in our context entails obtaining the bootstrap p-value from

$$P^*(T_n^* > T_n)$$

where $T_n = \sqrt{n}\max_{1\leq r} \Psi_{n,(i)}$, T_n^* denotes a bootstrap version of T_n, i.e.

$$T_n^* = \sqrt{n}\max_{1\leq i\leq r}(\Psi^*_{n,(i)} - \Psi_{n,(i)}),$$

and P^* denotes the probability distribution associated with T_n^*. Hereafter, we refer to such tests as full recentered bootstrap tests, or (FRB) for short.

[4]This problem is well documented and has motivated the development of alternative multiplicity tests in which the familywise error rate is relaxed in favour of controlling other false discovery rates; see e.g. Romano, Shaikh and Wolf (2007) and references therein.

All of the aforementioned tests can be characterized as LFC-tests in that they attempt to bound the probability of a Type I error for all configurations under the null. A consequence is that each of the LFC-tests is conservative if any of the inequalities are non-binding. Additionally, each of these tests is biased, i.e. each will exhibit power less than size for some sequence of Pitman alternatives approaching the boundary of the null. The issue of unbiasedness of composite tests is the focus of the recent paper by Hansen (2003), where he shows, among other things, that contrary to popular belief the LFC approach is not necessary in order to construct asymptotically exact tests.

In this paper we extend the work of Bennett (2009) by introducing a new MinP-type test for jointly testing multiple moment inequalities. Our testing procedure makes novel use of the bootstrap distribution in order to dramatically reduce the computational burden which is typically associated with MinP tests. Moreover, we modify the typical fully recentered bootstrap by recentering only those inequality restrictions which are estimated to be in the alternative and those strictly in the null but within a small neighbourhood of the boundary. This subtle innovation applied to the resampling procedure has a dramatic impact on the resampling distribution; namely, the distribution mimics the asymptotic null distribution for *any* parameter configuration on the boundary of the null. Consequently, we are able to establish that our test delivers asymptotic sizes that are valid uniformly over the null hypothesis, and that our test delivers asymptotic sizes equal to the nominal level uniformly over the boundary of the null hypothesis. As is demonstrated by Andrews and Guggenberger (2007) and Mikusheva (2007), for example, uniformity is required for good finite sample approximations whenever a test statistic has a discontinuity in its limit distribution, as is the case here when testing multiple moment inequalities.

The related literature on inference with moment inequalities is large and growing rapidly. Examples of recent contributions in this area include Linton, Song and Whang (2008), Hansen (2005), Moon and Schorfheide (2008), Chernozhukov et al. (2007), Andrews and Soares (2007), Andrews and Guggenberger (2007), Rosen (2006), Fan and Park (2007). And indeed other procedures exist that may be adapted to our environment. For example the methods prescribed by Andrews and Jia (2008) can be adapted and applied in the context considered herein. However, our MinP procedure is the *only* testing procedure which is (i) consistent; (ii) balanced; (iii) invariant to studentization; (iv) asymptotically unbiased; and (v) yields multiplicity-adjusted p-values that enable the user to gauge the strength of the evidence against the individual moment inequalities. Users thus are not only able to identify which moment inequalities which are responsible for rejection of the null but also to validly infer their sign. Furthermore, in situations where hypotheses are not considered equally important users may introduce weight functions to be applied directly to the adjusted p-values—as opposed to the test statistics—thus enabling application of a transparent re-weighting of the statistical evidence.

The remainder of the paper is organized as follows. Section 2 describes the estimation of the moment inequalities and develops the associated asymptotics. Section 3 develops the background for the the minimum p-value test which is introduced in Section 4. In Section 5 we discuss the asymptotic power of our test against fixed and local alternatives, and then we illustrate the finite sample performance via simulation in Section 6. Section 7 presents an empirical analysis of the distributional effects of Vietnam veteran status on earnings. Section 8 contains concluding remarks.

2. ESTIMATION AND ASYMPTOTIC THEORY

Central to the testing procedures discussed below is the estimation of $\Psi(P) = \mathbb{E}_P[\psi(X_i, \theta(P))]$. Implicit in our assumptions will be the availability of a consistent estimator $\theta_n := \theta(\mathbb{P}_n)$ of $\theta(P)$ which is used to obtain the "plug-in" estimator $\Psi_n(\mathbb{P}_n) = \mathbb{P}_n[\psi(X_i, \theta(\mathbb{P}_n)]$ of $\Psi(P)$. Formally, let X denote a random vector with support $\mathcal{X}$ and let Θ_0 denote a bounded subset of $\mathbb{R}^k$. Given a measurable function $\psi : \mathcal{X} \times \Theta_0 \to \mathbb{R}^r$ let $\mathcal{F}$ denote the class of functions

$$\{\langle\psi(\cdot;\theta), h\rangle : \theta \in \mathbb{R}^{dim(\Theta_0)}, h \in \mathbb{R}^k, \|h\| \leq 1\} \tag{3}$$

where $\langle \cdot, \cdot \rangle$ denotes the standard inner product on $\mathbb{R}^n$. Notice that in all of the examples in the previous section, the centered random vectors

$$\sqrt{n}\left(\mathbb{P}_n \psi(X_i, \theta(\mathbb{P}_n)) - P\psi(X_i, \theta(P))\right) \tag{4}$$

may be recovered from an empirical process of the form

$$\sqrt{n}(\mathbb{P}_n f_{\theta_1(\mathbb{P}_n),h_1} - P f_{\theta_1(P),h_1}) - \sqrt{n}(\mathbb{P}_n f_{\theta_2(\mathbb{P}_n),h_2} - P f_{\theta_2(P),h_2}), \tag{5}$$

where $f_{\theta,h} \in \mathcal{F}$, either by setting $h_1 = h_2$ equal to the standard basis vectors in $\mathbb{R}^r$, or by setting h_1 equal to the standard basis vectors in $\mathbb{R}^r$ and setting $h_2 = 0_{r\times 1}$.

In this section we outline a set of primitive conditions on the data generating process and the map ψ and show that these assumptions are sufficient for establishing that the empirical process defined in (5) is Donsker uniformly in a class of measures which we will denote by $\mathcal{P}$. We will then obtain as a corollary the convergence of (4) to a multivariate normal random vector $\mathcal{Z} \sim N(0, \Omega(P))$ uniformly in $P \in \mathcal{P}$.

We begin by providing a formal statement of our assumptions.[5] In particular, denote by $\mathcal{P}$ the collection of of all potential distributions of X that satisfy the conditions of Assumption 1 below.

Assumption 1.

i. $X_1, \ldots, X_n$ *are i.i.d. copies of a vector random variable* X *with distribution* $P \in \mathcal{P}$

ii. $\sup_{P\in\mathcal{P}} \sup_{\theta\in\Theta_0} P\left[\|\psi_\theta\|_2\right]^{2+\delta} < \infty$ *for some* $\delta > 0$

iii. $\|\psi_{\theta_1} - \psi_{\theta_2}\| \leq \|\theta_1 - \theta_2\| L(x)$ *for all* $x \in \mathcal{X}$.

iv. $\sup_{P\in\mathcal{P}} PL^{2+\delta} < \infty$ *for some* $\delta > 0$

[5] In our discussion we make frequent use of operator notation, i.e. $Pf := \int f dP$.

v. For some $\delta > 0$, there exists a function $D_P(\theta - \theta_0)$ such that

$$|Pf_{\theta,h} - Pf_{\theta_0,h} - h \cdot D_P(\theta - \theta_0)| \leq C\|\theta - \theta_0\|$$

for $\theta \in B_\delta(\theta_0)$, where C is a constant that does not depend on P.

vi. For each $\epsilon > 0$

$$\limsup_{n\to\infty} \sup_{P\in\mathcal{P}} \mathbb{P}_P \left\{ \left| \sqrt{n} D_P(\theta - \theta(P)) - M(\theta(P)) \frac{1}{\sqrt{n}} \sum_{i=1}^{n} S(X_i, \theta(P)) \right| > \epsilon \right\} = 0$$

where $E_P[S(X_i, \theta(P))] = 0$ and $\sup_{P\in\mathcal{P}} P[\|M(\theta(P))S(\cdot, \theta(P))\|_2]^{2+\delta}$ for some $\delta > 0$.

Assumption 1(i) is rather straightforward. Assumption 1(ii) requires the moments of the map ψ to be uniformly bounded up to order $2 + \delta$. Parts (iii) and (iv) require ψ to be Lipschitz continuous with Lipschitz function having bounded moments up to order $2+\delta$. Parts (v) and (vi) of Assumption 1 together require $E_P[f_{\theta,h} - f_{\theta_0,h}]$ to admit an asymptotic linear representation uniformly in $\mathcal{P}$. Note that many common estimators of θ_0, including GMM, OLS, and IV estimators will satisfy this condition under appropriate differentiability conditions on $E_P[f_{\theta,h}]$ at $\theta = \theta_0$.

For the statement of our main result of this section we introduce the following additional notation. Define $\mathcal{G} = \{\langle h, M(\theta(P))S(\cdot, \theta(P))\rangle : h \in \mathbb{R}^r, \|h\|_2 \leq 1\}$, and denote by Φ the class of functions

$$\{(f_{\theta_1,h_1} - f_{\theta_2,h_2}) + (g_{\theta_1,h_1} - g_{\theta_2,h_2}) : (f_{\theta_1,h_1}, f_{\theta_2,h_2}, g_{\theta_1,h_1}, g_{\theta_2,h_2}) \in \mathcal{F}^2 \times \mathcal{G}^2\}.$$

Theorem 1. *Suppose the conditions of Assumption 1 are satisfied. Then*

$$\sqrt{n}(\mathbb{P}_n f_{\theta_1(\mathbb{P}_n),h_1} - Pf_{\theta_1(P),h_1}) - \sqrt{n}(\mathbb{P}_n f_{\theta_2(\mathbb{P}_n),h_2} - Pf_{\theta_2(P),h_2}) =$$
$$\mathbb{G}_{n,P}[(f_{\theta_1(P),h_1} - f_{\theta_2(P),h_2}) + (g_{\theta_1(P),h_1} - g_{\theta_2(P),h_2})] + o_P(1)$$

uniformly in $P \in \mathcal{P}$ and

$$\mathbb{G}_{n,P} \rightsquigarrow \mathbb{G}_P$$

in $\ell^\infty(\Phi)$ uniformly in $P \in \mathcal{P}$, where $\mathbb{G}_P$ denotes a mean-zero Gaussian process.

Theorem 1 implies that for any finite set $\{\phi_1, \ldots, \phi_r, \phi_i \in \Phi\}$

$$(\mathbb{G}_{n,P}\phi_1, \ldots, \mathbb{G}_{n,P}\phi_r)$$

converges in distribution to a multivariate normal distribution uniformly in $P \in \mathcal{P}$. The connection to our testing problem is made upon setting h_i equal to a standard basis vector in $\mathbb{R}^r$. We therefore obtain the desired convergence result as an immediate corollary to Theorem 1.

Corollary 1. *Suppose the conditions of Assumption 1 are satisfied. Then,*

$$\sqrt{n}(\mathbb{P}_n\psi(X_i, \theta(\mathbb{P}_n)) - P\psi(X_i, \theta(P)) \Rightarrow \mathcal{Z}_P$$

uniformly in $P \in \mathcal{P}$ where $\mathcal{Z}_P$ denotes a $r \times 1$ Gaussian random variable with covariance matrix $\Omega(P)$

3. MODIFIED MAXT TEST

The hypothesis in (1) is formulated as a joint hypothesis, the intersection of r hypotheses. A straightforward statistic for this testing problem is given by

$$T_n = \sqrt{n} \max_{1 \leq i \leq r} \Psi_{n,(i)}(\mathbb{P}_n) \tag{6}$$

where the subscript "i" denotes the ith element of the $r \times 1$ vector $\Psi_n(\mathbb{P}_n)$.[6]

[6]Presumably due to the shape of the acceptance boundaries, (Goldberger 1992) refers to a test that rejects when the maximum of two normalized sample means exceeds a suitably chosen critical value as a BOX test.

Notice that we have chosen to define T_n without studentizing the component statistics $\Psi_{n,(i)}$. In general we might expect tests based on the studentized versions to perform better in finite samples—the studentized $\Psi_{n,(i)}$ would then be asymptotically pivotal and invariant to scalar transformations.[7] Our choice to work here with the basic statistics is merely for notational convenience. Since the minimum p-value test that we introduce is invariant to studentization of the component statistics, focusing here on the basic statistics will allow us to maintain consistent notation throughout the paper.

Before presenting the asymptotic distribution of T_n we first introduce some additional notation. Let $\mathcal{P}_0 \subset \mathcal{P}$ denote the collection of probability measures for which $\Psi_{(i)}(P) \leq 0$ for $1 \leq i \leq r$, and $\mathcal{P}_{00} \subset \mathcal{P}$ denote the collection of probability measures for which the inequality is strict for every i. Additionally, denote by $\mathcal{I}(P)$ and $\mathcal{I}_1(P)$ the set of indices $i \in \{1, \ldots, r\}$ for which $\Psi_{(i)}(P) = 0$ and $\Psi_{(i)}(P) \geq 0$ under P. The asymptotic distribution of T_n may now be stated as follows:

Theorem 2. *Suppose that Assumption 1 is satisfied. Then,*

$$T_n \to^d \begin{cases} \max_{i \in \mathcal{I}(\mathcal{P})} \mathcal{Z}_{P,(i)} & \text{if } P \in \mathcal{P}_0 \backslash \mathcal{P}_{00} \\ -\infty & \text{if } P \in \mathcal{P}_{00} \\ \infty & \text{if } P \in \mathcal{P} \backslash \mathcal{P}_0 \end{cases} \tag{7}$$

uniformly in $P \in \mathcal{P}$ where $\mathcal{Z}_P$ is the $k \times 1$ Gaussian random variable defined in Corollary 1.

The nature of the asymptotic distribution outlined in Proposition 2 suggests that a procedure for estimating an α-level critical value that exploits the information content in a given sample concerning the set $\mathcal{I}(P)$ may be used to obtain smaller asymptotically valid critical values, and hence more powerful tests than those based on the least favourable configuration. This is the basic intuition behind the testing procedures of Hansen (2003),

[7]Note that irrespective of whether $\Psi_{n,(i)}$ or a studentized version enters T_n, T_n is not asymptotically pivotal.

Hansen (2005), Linton et al. (2008), and also our work herein.[8] By adopting such a strategy we seek to develop tests which satisfy both

$$\lim_{n\to\infty} \sup_{P\in\mathcal{P}_0\backslash\mathcal{P}_{00}} \mathbb{P}_P\left(T_n \leq \hat{\tau}_n\right) = 1-\alpha \tag{8}$$

and

$$\lim_{n\to\infty} \inf_{P\in\mathcal{P}_0\backslash\mathcal{P}_{00}} \mathbb{P}_P\left(T_n \leq \hat{\tau}_n\right) = 1-\alpha. \tag{9}$$

Tests that satisfy (9) are said to be asymptotically valid uniformly on the boundary of the null hypothesis. A test satisfying condition (9) is said to be asymptotically similar on the boundary of the null hypothesis. Note that the latter condition is necessary for asymptotic unbiasedness, and also that these conditions are considerably stronger than the requirement that $\hat{\tau}_n$ satisfy

$$\sup_{P\in\mathcal{P}_0} \lim_{n\to\infty} \mathbb{P}_P(T_n \leq \hat{\tau}_n) = 1-\alpha;$$

i.e. the least favourable configuration α-level critical value.

3.1 Bootstrap Critical Values

Standard implementations of the MaxT test employ critical values obtained in the least favourable case where all of the inequalities bind. As was pointed out in the previous section, this approach leads to conservative testing procedures. In this section we describe a bootstrap procedure involving the basic statistics that delivers $\hat{\tau}_n$ satisfying both (8) and (9).

Let $\Psi^*_{n,(i)}$ denote a bootstrap version of $\Psi_{n,(i)}(\mathbb{P}_n)$. Typically bootstrap critical values are obtained from the recentered bootstrap distribution, i.e.

$$\Psi^{*FC}_{n,(i)} = \Psi^*_{n,(i)} - \Psi_{n,(i)}(\mathbb{P}_n).$$

[8] Similar tools are also employed in the literature on partial identification.

In this paper we consider a slight modification of the fully recentered bootstrap. Specifically, we define the partially recentered bootstrap statistics that enter the scaled max function by

$$\Psi^{*PC}_{n,(i)} = \Psi^{*}_{n,(i)} - \max\{\Psi_{n,(i)}(\mathbb{P}_n), -\delta_n\} \tag{10}$$

for $1 \leq i \leq r$, where δ_n is sequence satisfying both $\delta_n \to 0$ and $\sqrt{n}\delta_n \to \infty$. Notice that the recentering in (10) differs from the standard implementation of the bootstrap; namely only those inequality restriction estimates which fall in the alternative or are in the null but within δ_n of equality are recentered. As is demonstrated Theorem 3 below, this modification enables the bootstrap distribution to mimic the asymptotic distribution for any parameter configuration under the null.

Theorem 3. *Suppose that the conditions of Assumption 1 are satisfied. Then, for any positive sequence δ_n such that $\delta_n \to 0$ and $\sqrt{n}\delta_n \to \infty$, we have*

$$T^*_n := \sqrt{n} \max_{1 \leq i \leq r} \Psi^{*PC}_{n,(i)} \to^d \begin{cases} \max_{i \in \mathcal{I}(P)} \mathcal{Z}_{P,(i)} & \text{if } P \in \mathcal{P}_0 \backslash \mathcal{P}_{00} \\ -\infty & \text{if } P \in \mathcal{P}_{00} \\ \max_{i \in \mathcal{I}_1(P)} \mathcal{Z}_{P,(i)} & \text{if } P \in \mathcal{P} \backslash \mathcal{P}_0 \end{cases} \tag{11}$$

and the convergence is uniform over $P \in \mathcal{P}$, .

Theorem 3 above demonstrates that the partially recentered bootstrap, with an appropriately chosen sequence δ_n, is indeed sufficient to obtain consistency. Moreover, the theorem establishes that consistency holds for any configuration of the parameters under the null hypothesis. Theorem 3 also provides the justification for the using the partially recentered bootstrap distribution to approximate the sampling distribution of T_n and in turn using this approximation to estimate asymptotically valid critical values. A formal statement of the decision rule associated with the modified MaxT test is given by

$$\text{Reject } H_0 \text{ if } T_n > \max\{0, \hat{\tau}_n(\alpha)\},$$

where

$$\hat{\tau}_n(\alpha) = \inf\{\tau : P^*(T_n^* \leq \tau) \leq 1 - \alpha\}.$$

The validity of the testing procedure is the content of Theorem 4:

Theorem 4. *Suppose conditions of Theorem 3 are satisfied. Then, for any positive sequence δ_n such that $\delta_n \to 0$ and $\sqrt{n}\delta_n \to \infty$, we have for $\alpha < 1/2$*
(i) Uniformly over $P \in \mathcal{P}_0$

$$\lim_{n\to\infty} \mathbb{P}_P(T_n \leq \hat{\tau}_n(\alpha)) \leq 1 - \alpha;$$

and
(ii) Uniformly over $P \in \mathcal{P}_{00} \backslash \mathcal{P}_0$

$$\lim_{n\to\infty} \mathbb{P}_P(T_n \leq \hat{\tau}_n(\alpha)) = 1 - \alpha.$$

Part (i) of the theorem states that the bootstrap test is asymptotically valid uniformly over $P \in \mathcal{P}_0$. Part (ii) of the theorem demonstrates that the test delivers exact size uniformly over the boundary of the null hypothesis. It is precisely this latter property of asymptotic similarity that will enable the test to maintain greater power against alternatives in which the number of inequality restrictions is large and a portion of the inequality restrictions are strictly in the null. Notably, none of the LFC tests—this includes the minimum distance tests—share this property and as a result will generally have low power in such situations. These issues are discussed in further detail in Section 5.

An attractive feature of the bootstrap test proposed herein that is also not shared by the minimum distance tests is that conclusive statements may be made about the sign of individual restrictions when the null is rejected. Specifically, for all $c > 0$ we have

$$\lim_{n\to\infty} \sup_{P\in\mathcal{P}} \mathbb{P}_P\left(\sqrt{n}\Psi_{n,(i)} \leq c\right) \geq \lim_{n\to\infty} \sup_{P\in\mathcal{P}} \mathbb{P}_P(T_n \leq c) \tag{12}$$

Combining (12) with, say, Theorem 4(i) implies that the confidence set

$$C_n(1-\alpha) = (-\infty, \hat{\tau}_n(\alpha)]$$

has asymptotic coverage probability at least $1-\alpha$ for all $\Psi_{(i)}$, $1 \leq i \leq r$.

4. MINIMUM P-VALUE TEST

The component statistics entering the maximum in T_n will generally have different distributions. Consequently, it may be desirable to reformulate the tests so as to eliminate the potentially excessive influence of any single component on T_n. In principle this can be accomplished by resampling the p-values of the individuals statistics and then basing our test on the minimum of these p-values. In this section we show how the partially recentered resampling schemes which were discussed previously can used to develop computationally efficient minimum p-value tests that preserve the desirable properties of consistency and asymptotic similarity.

The minimum p-value approach to multiple testing problems has recently been applied in Godfrey (2005) in the context of controlling the overall significance level of a battery of least squares diagnostic tests; see also MacKinnon (2007) and references therein for a general discussion of this approach. In contrast to previous approaches which require a double bootstrap, we propose a novel bootstrap procedure that requires only two individual bootstrap procedures. Denoting the number of bootstrap replications as B_1 and B_2 in the first and second stages, our procedure requires only $B_1 + B_2$ bootstrap replications in contrast to $B_1 + B_1 B_2$ replications required of the standard double bootstrap.[9] A second computational advantage of our procedure is that the bootstrap procedure in the second-stage does not involve re-estimation but instead involves only resampling with replacement from the bootstrap distribution obtained in the first-stage. Avoiding re-estimation can result in significant computational savings, especially in nonlinear environments. Remarkably, our procedure retains the attractive features of consistency, and asymptotic similarity on the boundary of the null.

In our description of the minimum p-value tests we make use of the following notation. First, we denote the jth of B_1 resampled vectors by

$$\Psi_j^* = (\Psi_{j,(i)}^*, \ldots, \Psi_{j,(r)}^*),$$

where the $\Psi_{j,(i)}^*$, $1 \leq i \leq r$ denote the uncentered bootstrap statistics corresponding to the jth bootstrap sample (to ease the notational burden we have dropped the subscript corresponding to the sample size). For $1 \leq j \leq B_1$ we record both the fully recentered and partially recentered bootstrap statistics Ψ_j^{*PC} and Ψ_j^{*FC}, and denote by $\mathcal{H}^{*FC}$ and $\mathcal{H}^{*PC}$ the respective samples. We also define the ith marginal empirical distribution associated with $\mathcal{H}^{*FC}$ by

$$H_{(i)}^*(x_{(1)}, \ldots, x_{(r)}) = \ B_1^{-1} \textstyle\sum_{j=1}^{B_1} \mathbb{1}\{n^{1/2} \Psi_{j,(i)}^{*FC} \leq x_{(i)}\}, \quad 1 \leq i \leq r \ .$$

Finally by defining the function $\rho : \mathbb{R}^r \to [0, 1]$ by

$$\rho(x) = \min_{1 \leq i \leq r} \left\{ \left(1 - H_{(i)}^*(x)\right) \right\}$$

[9] There are several papers which seek to minimize the computational burden of the double bootstrap. See, for example, MacKinnon (2007) for a discussion of the "fast double bootstrap," and Nankervis (2005) for the use of various stopping rules. It is also worth noting that our bootstrap procedure is valid under weaker conditions than those required by the "fast double bootstrap."

we may write the minimum p-value test statistic for testing (2) concisely as

$$\hat{p}_{min} = \rho(\sqrt{n}\Psi_n).$$

The limiting null distribution of $\hat{p}_{min}$ is the content of Proposition 1 below:

Proposition 1. *Suppose that the conditions of Assumption 1 are satisfied, and let U denote a $r \times 1$ random vector with uniform marginals. Then,*

$$\hat{p}_{min} \rightarrow^d \begin{cases} \min_{i \in \mathcal{I}(P)} U_{(i)} & \text{if } P \in \mathcal{P}_0 \backslash \mathcal{P}_{00} \\ 1 & \text{if } P \in \mathcal{P}_{00} \\ 0 & \text{if } P \in \mathcal{P} \backslash \mathcal{P}_0 \end{cases} \tag{13}$$

uniformly in $P \in \mathcal{P}$

Small p-values are associated with estimated values in the upper tail of the marginal distributions, and hence rejection of the null should occur for $\hat{p}_{min}$ sufficiently small, i.e. for for $\hat{p}_{min} < p_{crit}$. Estimating p_{crit} so as to obtain an asymptotic α-level test requires estimating the sampling distribution of $\hat{p}_{min}$. We now describe a second-stage sampling procedure for estimating the sampling distribution $\hat{p}_{min}$ that delivers an exact α-level test uniformly over the boundary of the null hypothesis. Specifically, letting $\{\Psi_1^{**}, \ldots, \Psi_{B_2}^{**}\}$ denote B_2 randomly sampled observations from $\mathcal{H}^{*PC}$, an asymptotically exact α-level critical value for the minimum $MinP$ test may be estimated from

$$\hat{p}_{crit}(\alpha) = \inf \left\{ t : B_2 \sum_{j=1}^{B_2} \mathbb{1} \left\{ \rho(\Psi_j^{**}) \leq t \right\} \geq \alpha \right\}$$

Notice that for each term in the sum the minimum is evaluated at a randomly sampled vector from the joint (resampled) distribution of the statistics. This is necessary in order to capture the dependence between the individual statistics. Theorem 5 below summarizes several basic asymptotic properties of the $MinP$ value tests.

Theorem 5. *Suppose that the conditions of Assumption 1 are satisfied. Then, for any positive sequence δ_n such that $\delta_n \to 0$ and $\sqrt{n}\delta_n \to \infty$, we have for $\alpha < 1/2$*
(i) Uniformly over $P \in \mathcal{P}_0$

$$\lim_{n\to\infty} \mathbb{P}_P(\hat{p}_{min} \leq \hat{p}_{crit}(\alpha)) \leq 1 - \alpha;$$

and
(ii) Uniformly over $P \in \mathcal{P}_{00} \backslash \mathcal{P}_0$

$$\lim_{n\to\infty} \mathbb{P}_P(\hat{p}_{min} \leq \hat{p}_{crit}(\alpha)) = 1 - \alpha.$$

Theorem 5 reveals that the $MinP$ tests is uniformly asymptotically valid, and delivers asymptotic size equal to the nominal size uniformly on the boundary of the null hypothesis.

Similar to the $MaxT$ test, the $MinP$ test enjoys the desirable property that a rejection of the null hypothesis permits inferences to be made about the signs of a subset of the inequality restrictions being tested. In particular, when the null is rejected we may infer the sign of the ith moment to be equal to the sign of the estimate of restriction i for any i such that

$$1 - H_i^*(n^{1/2}\Psi_n) < \hat{p}_{crit}(\alpha)$$

The validity of this procedure follows from the fact that

$$\lim_{n\to\infty} \mathbb{P}_P\left(1 - H_i^*(n^{1/2}\Psi_n) < \hat{p}_{crit}(\alpha)\right) \leq \alpha$$

for $1 \leq i \leq r$ and all $P \in \mathcal{P}$.

5. POWER

In this section we study the power of our proposed tests against both fixed and $\sqrt{n}$-Pitman alternatives. One of the more noteworthy findings that emerges from this analysis is the asymptotic unbiasedness of both the $MaxT$ and $MinP$ tests. In addition, we establish the weak dominance of these testing procedures over their analogues based on the canonical fully recentered bootstrap.

We begin this section by studying the power of our proposed tests against a fixed alternative. Letting $P \in \mathcal{P} \backslash \mathcal{P}_0$, we have, by definition, $\Psi_{(i)} > 0$ for some $1 \leq i \leq r$. Since $T_n^{PC} = O_P(1)$, whereas $T_n \to \infty$ in this case, it should not be surprising that $\mathbb{P}_P(T_n > \hat{\tau}_n^{PC}) \to 1$ as $n \to \infty$. Similarly, $\hat{p}_{min} \to 0$, whereas $\hat{p}_{crit}(\alpha)$ converges to the αth percentile of a continuous distribution with support on the unit interval. Under a fixed alternative, it follows that $\mathbb{P}_P(\hat{p}_{min} > \hat{p}_{crit}(\alpha)) \to 1$ as $n \to \infty$. These heuristic arguments, which are developed more formally in the proof of Theorem 6 below, establish consistency of both testing procedures.

Theorem 6. *Let $P \in \mathcal{P} \backslash \mathcal{P}_0$ and suppose Assumption 1 holds. Then,*

(i)

$$\lim_{n \to \infty} \mathbb{P}_P(T_n > \hat{\tau}_n^{PC}(\alpha)) = 1;$$

and

(ii)

$$\lim_{n \to \infty} \mathbb{P}_P(\hat{p}_{min} < \hat{p}_{crit}(\alpha)) = 1.$$

Consistency against fixed alternatives is a minimal requirement for any testing procedure. We now explore the power of our tests against $\sqrt{n}$-Pitman alternatives. Following Linton et al. (2008) we restrict attention to $\{P_n\}$ such that $P_n \in \mathcal{P} \backslash \mathcal{P}_0$ for all n and

$$\Psi(P_n) = \Psi(P) + n^{-1/2}\xi \tag{14}$$

where $\xi \in \mathbb{R}^r$ and $P \in \mathcal{P}_0$. Thus, in a slight abuse of notation, $\Psi(P)$ is used here to denote a configuration on the boundary of the null hypothesis, and $\Phi(P_n)$ denotes a configuration in which $\Psi_{(i)}(P_n)$ is strictly positive for some $i \in \{1, \ldots, r\}$.

Theorem 7. *Suppose Assumption 1 is satisfied. Additionally, assume $\Psi(P_n)$ is a sequence of alternatives satisfying (14). Then,*

(i)

$$\lim_{n \to \infty} \mathbb{P}_{P_n}(T_n > \tau_n^{PC}) \geq \alpha$$

and

(ii)

$$\lim_{n \to \infty} \mathbb{P}_{P_n}(\hat{p}_{min} < \hat{p}_{crit}(\alpha)) \geq \alpha$$

Parts (i) and (ii) of Theorem 7 establish that both testing procedures are asymptotically unbiased; i.e. both procedures have power at least as great as the nominal size of the test for any sequence of $\sqrt{n}$-Pitman alternatives. On the other hand, it is well known that the classical minimum distance tests are biased see e.g. Wolak (1991) p. 220. That these procedures weakly dominate analogous tests based on the fully recentered bootstrap follow immediately from the fact that $\tau_n^{PC} \leq \tau_n^{FC}$ and $\hat{p}_{crit} \geq \hat{p}_{crit}^{FC}$, where the superscript "FC" denotes the critical value obtained via a full recentering scheme.

6. SIMULATION EXPERIMENTS

Here we present simulation evidence concerning the size and power properties of our proposed tests. We have chosen to conduct our simulation experiment in the context of testing for *normalized* stochastic dominance (NSD). The development of powerful testing procedures in this context is a nontrivial extension of testing for stochastic dominance and also has important implications. For example, establishing NSD at second order is equivalent to establishing a Lorenz ordering, whereas NSD at third-order is consistent with the inequality ordering implied by any Lorenz consistent inequality measure that satisfies the transfer sensitivity axiom of Shorrocks and Foster (1987); see Foster and Sen (1999) for further discussion. We point out that an alternative test for normalized stochastic dominance already exists and may be found in Zheng et al. (2000); however, by construction their testing procedure is conservative. Additionally, their test is designed to test the null of equality against the alternatives of crossing, or dominance in either direction.[10] A direct comparison of power of the two tests in the context of testing one-sided dominance is therefore not presented here.[11]

In our analysis we consider dominance comparisons among various parametrizations of two hypothetical income distributions: the lognormal distribution, and the Singh-Maddala distribution.[12] For ease of reference the parametrizations as well as the specific pairwise comparisons are described in Table 1 below.

[Table 1 about here.]

In each of the experimental designs described in Table 1 pseudo random samples of size $n = 500, 1000, 1500$, and 2000 are generated from the respective distributions and used to test the hypothesis

$$H_0 : F^{(s)}(x) \leq G^{(s)}(x) \text{ for all } x \in \mathcal{X}$$

against

$$H_1 : F^{(s)}(x) > G^{(s)}(x) \text{ for some } x \in \mathcal{X}$$

for $s = 2, 3, 4$, where $F^{(2)}(x) = \int_0^x F(\mu_F y)dy$ and $F^{(s)}(x) = \int_0^x F^{(s-1)}(y)dy$ for $s > 2$, with

[10]We note that our test may be easily adapted to testing the null of equality against multiple alternatives, though a formal discussion of such a procedure is beyond the scope of the current paper.

[11]Our original intention was to adopt the experimental designs of Zheng et al. (2000) and offer a direct comparison. However, while we have, for instance, maintained Case I from their simulations of power against crossings, we do not report the results corresponding to Case II and III since we found our test to have empirical rejection probabilities of 1 for samples as small as 100. In contrast, the test of Zheng et al. (2000) rejects at most 16% of the 1000 null case simulations with a sample size of 1000.

[12]The three parameter Singh-Maddala distribution, denoted SM(a,b,q), is given by

$$F(x) = 1 - \left[1 + \left(\frac{x}{b}\right)^a\right]^{-q} \text{ with } a, b, q \geq 0.$$

$G^{(s)}$ defined analogously. In each case $\mathcal{X}$ consists of 10 equally spaced points between the the 5th and 95th percentiles of F, and the results for both the fully-recentered and partially recentered bootstrap MinP tests are reported. The value of δ_n we use in the simulations is specific to each hypothesis under test and is determined by $\delta_n = 0.1\hat{\sigma}_{(i)}\sqrt{\log(\log(n))}/\sqrt{n}$, where $\hat{\sigma}_{(i)}$ is the standard deviation of the ith statistic as estimated from the first-stage bootstrap.

Experiments (a) and (b) are designed to illustrate the empirical size of the test. The simulation results are reported in Table 2. For all samples sizes and both experiments we see that the $MinP$ test delivers rejection probabilities close to the nominal size of 5%. This is true irrespective of whether the fully-recentered or partially-recentered test is employed. There is, however, some evidence in design (b) of a slight decline in the size of the tests as the order of dominance increases, though this decline is less pronounced as the sample size is increased.[13]

[Table 2 about here.]

Designs (c) and (d) are designed to illustrate the finite sample power of our test against crossings . In particular, both cases involve relatively minute crossings of the respective population curves with case (d) being the most difficult to detect. It is in these situations that we should expect to observe the partially-recentered bootstrap test significantly outperforming the fully-recentered counterpart. Indeed, the partially-recentered bootstrap test dominates in every case and identifies up to 23% more false hypotheses than the fully-recentered counterpart.

Designs (e) and (f) illustrate the performance of the tests in the case of dominance. All of the inequalities are in the alternative and hence both implementations of the MinP test should and do yield similar results. In every case there is little difference between the rejection probabilities of the two tests and in samples as small as $n = 2000$ both tests reject

[13]Qualitatively, this finding is consistent with the simulations of Zheng et al. (2000)

in a significant proportion of trials.

Before concluding this section we note that further simulation results on the size and power of the MinP test, albeit in the context of testing the difference between vectors of simple population means (i.e. when θ is known and thus does not require estimation), can be found in Bennett (2009). While qualitatively the results reported here are similar, Bennett (2009) considers a broader range of designs—including varying the number of grid points (ranging from 10 to 100)—and also reports on the proportion of false hypotheses identified from the collection of hypotheses under test.

7. EMPIRICAL ILLUSTRATION

In this section we apply our test to expand on the analysis of Abadie (2002) where randomization in the draft lottery is exploited to examine the distributional effects of Vietnam veteran status on labour earnings. Abadie shows that the potential distributions for veterans and non-veterans can be estimated for the subpopulation of compliers, and uses the Kolmogorov-Smirnov (KS) type tests of McFadden (1989) and Barrett and Donald (2003) to empirically examine whether the earnings distribution of non-veterans stochastically dominates that of veterans. Using the exact data employed by Abadie, we apply our testing procedure to test for both stochastic dominance and *normalized* stochastic dominance. As we demonstrate below, our test allows us to make unambiguous statements concerning the effect of veteran status on both poverty and inequality.

7.1 Preliminaries

Let Z_i be a binary instrument. Let the pair $D_i(0)$ and $D_i(1)$ denote the values of the treatment for individual i that would be obtained given the instrument $Z_i = 0$ and $Z_i = 1$ respectively. If $D_i(0) = 0$ and $D_i(1) = 1$ individual i is called a *complier*. Let $Y_i(0)$ be the potential outcome for individual i without treatment and $Y_i(1)$ be the potential outcome for the same individual with treatment. In practice, the analyst observes only the realized treatment and realized outcome, i.e.

$$D_i = D_i(1)Z_i + D_i(0)(1 - Z_i),$$

and

$$Y_i = Y_i(1)D_i + Y_i(0)(1 - D_i).$$

Define the cumulative distribution functions (cdfs) of the potential outcomes for compliers

$$F_1^C(y) = \mathbf{E}[1\{Y_i(1) \leq y\}|D_i(1) = 1, D_i(0) = 0] \tag{15}$$

and

$$F_0^C(y) = \mathbf{E}[1\{Y_i(0) \leq y\}|D_i(1) = 1, D_i(0) = 0] \tag{16}$$

Under a given set of assumptions (Abadie 2002, p. 285), Imbens and Rubin (1997) prove that the potential distributions are identified, and derive estimators for their associated probability density functions. Extending these results, Abadie (2002) provides a simple method for estimating the cdfs (15) and (16). In particular, Abadie shows that

$$F_1^C(y) = \frac{\mathbf{E}[1\{Y_i(1) \leq y\}D_i|Z_i = 1] - \mathbf{E}[1\{Y_i(1) \leq y\}D_i|Z_i = 0]}{\mathbf{E}[D_i|Z_i = 1] - \mathbf{E}[D_i|Z_i = 0]}, \tag{17}$$

and

$$F_0^C(y) = \frac{\mathbf{E}[1\{Y_i(1) \leq y\}(1 - D_i)|Z_i = 1] - \mathbf{E}[1\{Y_i(1) \leq y\}(1 - D_i)|Z_i = 0]}{\mathbf{E}[(1 - D_i)|Z_i = 1] - \mathbf{E}[(1 - D_i)|Z_i = 0]}; \tag{18}$$

and thus the potential earnings distributions F_1^C and F_0^C in (15) and (16) can be estimated using empirical counterparts of equations (17) and (18) respectively.

7.2 Data and Tests

Our interest centers on examining the impact of veteran status on the distribution of earnings for the subpopulation of compliers. In our study we use the same data set employed by Abadie (2002). As described in his paper, the data consist of 11,637 white men, born in 1950-1953, form the March Current Population Surveys of 1979 and 1981-1985.[14] For each individual in the sample, annual labor earnings, Vietnam veteran status and an indicator of draft-eligibility based on the Vietnam draft lottery outcome are provided. Following Abadie (2002), we use draft-eligibility as an instrument for veteran status.[15]

[Figure 1 about here.]

Figure 1 plots the potential earnings distributions for veterans and non-veterans as estimated from equations (17) and (18). We see that the empirical earnings distribution for the veterans lies below the distribution for non-veterans everywhere in the empirical support, except for a crossing that occurs around the 95th percentile. The descriptive nature of the empirical distributions is consistent with the claim that Vietnam veteran status was detrimental in terms of both inequality and poverty. To formally test the first of these hypotheses we present results for second- and third-order normalized stochastic dominance using our proposed minimum p-value test; i.e. we test

$$H_0 : F_1^C \; NSD_j \; F_0^C$$

$$H_1 : \neg(F_1^C \; NSD_j \; F_0^C)$$

where $\neg(A)$ denotes the negation of the statement A, and j denotes the order of dominance, i.e. $j = 2, 3$. In the tests we use a grid size of 10 equally spaced points ranging from

[14] The data set is maintained by Joshua Angrist and available on his webpage; see http://econ-www.mit.edu/faculty/angrist/data1/data.

[15] The construction of the draft eligibility variable is described in detail in Appendix C of Abadie (2002). Additionally, a discussion of the validity of draft eligibility as an instrument for veteran status may be found in Angrist (1990).

the lower bound of 1000 to the upper bound of 38,000, which is approximately the 95% of the empirical support. Additionally, we set the number of bootstrap samples to $B_1 = 3,999$ and $B_2 = 2,999$, respectively. The adjusted p-values corresponding to both tests and at each of the grid points is plotted in Figure 1. Notably, the plot reveals both the strength of the evidence against the hypothesis as well as the origin of the evidence. For both second- and third-order normalized dominance, the p-values drop below the 5% nominal level, with the minimum p-values being 0.049 and 0.036, respectively. Thus at conventional levels we would reject the null hypothesis and conclude—due to the direct correspondence between second-order normalized stochastic dominance and the Lorenz ordering—that there is unambiguously greater relative inequality in the earnings distribution of veterans.

In order to formally test the second claim regarding poverty[16], we first define

$$F_1(y) = \mathbf{E}[1\{Y_i \leq y\}|Z_i = 1], \tag{19}$$

and

$$F_0(y) = \mathbf{E}[1\{Y_i \leq y\})|Z_i = 0]. \tag{20}$$

It can be shown that F_1 stochastically dominates F_0 if and only if F_1^C stochastically dominates F_0^C; see Proposition 2.1 of Abadie (2002) for details.[17] The empirical counterparts of F_1 and F_0 are simpler to work with and are thus adopted in our tests for stochastic dominance. Figure 2 plots both the estimated distributions along with the p-values corresponding to tests of first- and second-order stochastic dominance (here we use the same set of grid points as well as the same number of bootstrap replications). Again, we find that the p-values drop below the 5% nominal level. More specifically, the exact minimum p-values are 0.023 and 0.044, respectively. The p-values suggest that there is sufficient evidence to reject the

[16]See also Anderson (1996), Davidson and Duclos (2000), and Barrett and Donald (2003), for example, on tests for stochastic dominance.

[17]Due to the lack of continuity, strictly speaking our theoretical analysis requires a slight modification in order to accommodate tests for first-order dominance. For the sake of expositional simplicity, however, we have elected not to pursue a separate and formal treatment of such conditions in this paper.

claim that the distribution of veterans stochastically dominates that of non-veterans. The rejection of dominance in this direction together with the non-negative individual statistics allow us to conclude that there is unambiguously greater poverty in the earnings distribution of Vietnam veterans.

[Figure 2 about here.]

8. CONCLUSIONS

This paper introduces a computationally inexpensive minimum p-value procedure for joint tests of multiple inequality moment restrictions. Our test is quite generally applicable and is shown to be asymptotically unbiased and more powerful than tests based on the conventional (fully recentered) bootstrap or classical minimum distance tests. Our simulation results also confirm the good finite sample size and power properties of our test.

A key feature of our test is that it provides a more comprehensive picture of the evidence against the null hypothesis. As we have illustrated in our test of the distributional effects of Vietnam veteran status, our test yields adjusted p-values corresponding to each of the inequalities under test. This feature allows users to spot the source(s) of rejection when the claim under the null hypothesis is rejected. In some cases, users of the test may wish to place more or less weight on various moment inequalities. For example, when testing for first-order stochastic dominance users may require, say, "stronger" evidence in the tails due to the increased prevalence of measurement error. This can be done easily and transparently with our test since the user may simply apply a differential α-level across the various moment conditions.

The results of this paper suggest that it may be worth exploring an extension of the MinP testing procedure to situations in which we wish to test a continuum of moment conditions. For example, in our empirical application we test only a finite set of necessary conditions for stochastic dominance. While in our particular application the comparison at a finite set of grid points appears satisfactory, in general in order to ensure consistency against the full set of conditions implied by the null of dominance, a comparison of the distributions should be made at all points in the support. Extending our test in this direction presents some interesting theoretical and computational challenges, and research in this direction is currently underway.

9. APPENDIX

Lemma 1. *Let $\mathcal{F}$ denote the collection of real-valued functions defined by*

$$\mathcal{F} = \{\langle \psi_\theta, h\rangle : \theta \in \Theta,\ h \in \mathbb{R}^r_+,\ \|h\|_2 \leq 1\},$$

and let f_{θ_1,h_1} and f_{θ_2,h_2} denote arbitrary elements in $\mathcal{F}$. Then, under the conditions of Assumption 1, we have

$$|f_{\theta_1,h_1} - f_{\theta_2,h_2}| \leq K(x)\|(\theta_1,h_1) - (\theta_2,h_2)\|$$

where $K : \mathcal{X} \to \mathbb{R}$ satisfies $\sup_{P\in\mathcal{P}} PK^2 < \infty$.

Proof of Lemma 1. For a fixed $x \in \mathcal{X}$

$$\begin{aligned}
|f_{\theta_1,h_1} - f_{\theta_2,h_2}| &= \left|\sum_{i=1}^n h_{(i)1}\psi_{(i)\theta_1} - \sum_{i=1}^n h_{(i)2}\psi_{(i)\theta_2}\right| \\
&= \left|\sum_{i=1}^n h_{(i)1}(\psi_{(i)\theta_1} - \psi_{(i)\theta_2}) - \sum_{i=1}^n (h_{(i)2} - h_{(i)1})\psi_{(i)\theta_2}\right| \\
&\leq \sum_{i=1}^n |\psi_{(i)\theta_1} - \psi_{(i)\theta_2}| + m(x)\sum_{i=1}^n |h_{(i)2} - h_{(i)1}| \\
&\leq m(x)\|(\psi_{\theta_1},h_1) - (\psi_{\theta_2},h_2)\|_1 \\
&\leq m(x)C\|(\psi_{\theta_1},h_1) - (\psi_{\theta_2},h_2)\|_2
\end{aligned}$$

where $m(x) = \max_i \sup_{\theta\in\Theta} \psi_{(i)\theta}(x)$ and C is a finite constant. Using the fact that ψ is Lipschitz then permits us to write

$$|f_{\theta_1,h_1} - f_{\theta_2,h_2}| \leq L(x)K(x)C\|(\theta_1,h_1) - (\theta_2,h_2)\|_2$$

Defining $K(x) = CL(x)m(x)$, and noting that $\sup_{P\in\mathcal{P}} PK^2$ follows from the Cauchy-Schwartz inequality together with the uniform in $\mathcal{P}$ square integrability assumption for both $m(x)$ and $L(x)$, completes the proof. $\square$

Lemma 2. *Let* $\mathcal{H} = \{h \in \mathbb{R}^r : \|h\|_2 \leq 1\}$*. Under the conditions of Assumption 1 there exists a constant* C*, depending on* $\Theta \times \mathcal{H}$ *and* $d = dim(\Theta \times \mathcal{H})$ *only, such that the bracketing numbers satisfy*

$$N_{[]}(\epsilon \|K\|_{P,2}, \mathcal{F}, L_2(P)) \leq C \left(\frac{diam(\Theta \times \mathcal{H})}{\epsilon} \right)^d$$

for all $0 < \epsilon < diam(\Theta \times \mathcal{H})$*.*

Proof of Lemma 2. See van der Vaart (1998), p. 271. $\square$

Lemma 3. *Let* $\mathcal{H} = \{h \in \mathbb{R}^r : \|h\|_2 \leq 1\}$ *and denote by* $\mathcal{G}$ *the class of functions*

$$\{h' M(\theta(P)) S(\cdot, \theta(P)) : h \in \mathcal{H}\}.$$

Under the conditions of Assumption 1 there exists a constant C*, depending on* $\mathcal{H}$ *and* $d = dim(\mathcal{H})$ *only, such that the bracketing numbers satisfy*

$$N_{[]}(\epsilon \|F\|_{P,2}, \mathcal{F}, L_2(P)) \leq C \left(\frac{diam \mathcal{H}}{\epsilon} \right)^d$$

for all $0 < \epsilon < diam\mathcal{H}$ *where* $F(x) = \|M(\theta(P)) S(x, \theta(P))\|_2$*.*

Proof of Lemma 3. The Lipschitz property of $g \in \mathcal{G}$ is obvious, and so the result can be proven in an analogous manner to the proof of Lemma 2. $\square$

Lemma 4. *Suppose the conditions of Assumption 1 are satisfied. Then, for any* $\epsilon > 0$*,*

$$\limsup_{n \to \infty} \sup_{P \in \mathcal{P}} \mathbb{P}_P \left(\sup_h \|G_{n,P}(\theta(\mathbb{P}_n), h) - G_{n,P}(\theta(P), h)\| > \epsilon \right) = 0,$$

where $\mathbb{G}_{n,P}(\theta, h) := \sqrt{n}(\mathbb{P}_n - P) f_{\theta,h}$*.*

Proof of Lemma 4. Introduce the events

$$A_n = \{\sup_h \|G_{n,P}(\theta(\mathbb{P}_n), h) - G_{n,P}(\theta(P), h)\| > \epsilon\}$$

and

$$B_n = \{\|\theta(\mathbb{P}_n) - \theta(P)\| < \delta\}.$$

We may then write

$$\begin{aligned}\limsup_{n\to\infty} \sup_{P\in\mathcal{P}} \mathbb{P}_P(A_n) &= \limsup_{n\to\infty} \sup_{P\in\mathcal{P}} \mathbb{P}_P[(A_n \cap B_n) \cup (A_n \cap B_n^c)] \\ &\leq \limsup_{n\to\infty} \sup_{P\in\mathcal{P}} \mathbb{P}_P(A_n \cap B_n) + \limsup_{n\to\infty} \sup_{P\in\mathcal{P}} \mathbb{P}_P(A_n \cap B_n^c)\end{aligned} \tag{21}$$

Since (21) holds for all $\delta > 0$ we have

$$\begin{aligned}\limsup_{n\to\infty} \sup_{P\in\mathcal{P}} \mathbb{P}_P(A_n) &\leq \lim_{\delta\downarrow 0} \limsup_{n\to\infty} \sup_{P\in\mathcal{P}} \mathbb{P}_P(A_n \cap B_n) + \lim_{\delta\downarrow 0} \limsup_{n\to\infty} \sup_{P\in\mathcal{P}} \mathbb{P}_P(A_n \cap B_n^c) \\ &\leq \lim_{\delta\downarrow 0} \limsup_{n\to\infty} \sup_{P\in\mathcal{P}} \mathbb{P}_P\left(\sup_{h, \|\theta-\theta(P)\|<\delta} \|G_{n,P}(\theta) - G_{n,P}(\theta(P))\| > \epsilon\right) \\ &= 0\end{aligned} \tag{22}$$

Note that the second inequality follows from the fact that $\limsup_{n\to\infty} \sup_{P\in\mathcal{P}} \mathbb{P}_P(B_n^c)$ is equal to zero for all $\delta > 0$. The third line follows from the uniform in $\mathcal{P}$ Donsker property of $\mathcal{F}$—see Lemma 1 together with Assumption 1(ii) and Theorem 2.8.4 of van der Vaart and Wellner (1996)—which in turn implies that the sequence $\mathbb{G}_{n,P}$ is asymptotically equicontinuous uniformly in $P \in \mathcal{P}$.

□

Lemma 5. *Define* $\Delta_1 = \{f - g : (f, g) \in \mathcal{F} \times \mathcal{F}\}$, $\Delta_2 = \{f - g : (f, g) \in \mathcal{G} \times \mathcal{G}\}$, *and*

$$\Phi = \{f + g : (f, g) \in \Delta_1 \times \Delta_2\}.$$

Also, let $e(x) = \sup_{\phi\in\Phi} |\phi(x)|$. Then, If the conditions of Assumption 1 are satisfied,

$$\int_0^\infty \sup_{P\in\mathcal{P}} \sqrt{\log N_{[]}(\epsilon\|e\|_{P,2}, \Phi, L_2(P))} d\epsilon < \infty$$

Proof. By repeated application of Lemma 9.25 of Kosorok (2008) we have

$$N_{[]}(\epsilon\|\phi\|_{P,2}, \Phi, L_2(P)) \leq \left[N_{[]}\left(\frac{\epsilon}{4}\|e\|_{P,2}, \mathcal{F}, L_2(P)\right)\right]^2 \left[N_{[]}\left(\frac{\epsilon}{4}\|e\|_{P,2}, \mathcal{G}, L_2(P)\right)\right]^2$$

and thus

$$\begin{aligned}\sup_{P\in\mathcal{P}} \log N_{[]}(\epsilon\|e\|_{P,2}, \Phi, L_2(P)) \leq 2 \sup_{P\in\mathcal{P}} \log N_{[]}\left(\frac{\epsilon}{4}\|e\|_{P,2}, \mathcal{F}, L_2(P)\right) \\ + 2 \sup_{P\in\mathcal{P}} \log N_{[]}\left(\frac{\epsilon}{4}\|e\|_{P,2}, \mathcal{G}, L_2(P)\right)\end{aligned} \tag{23}$$

Since Lemmas 2 and 3 imply that each term on the right is of the order $\log\left(\frac{1}{\epsilon}\right)$ the finiteness of the bracketing integral is established.

□

Proof of Theorem 1. Following p. 234 of van der Vaart and Wellner (2007) we consider the decomposition

$$\begin{aligned}&\sqrt{n}(\mathbb{P}_n f_{\theta_1(\mathbb{P}_n),h_1} - P f_{\theta_1(P),h_1}) - \sqrt{n}(\mathbb{P}_n f_{\theta_2(\mathbb{P}_n),h_2} - P f_{\theta_2(P),h_2}) \\ &\quad = \mathbb{G}_{n,P}(f_{\theta_1(\mathbb{P}_n),h_1} - f_{\theta_1(P),h_1}) - \mathbb{G}_{n,P}(f_{\theta_2(\mathbb{P}_n),h_2} - f_{\theta_2(P),h_2}) \\ &\quad + \sqrt{n}P(f_{\theta_1(\mathbb{P}_n),h_1} - f_{\theta_1(P),h_1}) - \sqrt{n}P(f_{\theta_2(\mathbb{P}_n),h_2} - f_{\theta_2(P),h_2}) \\ &\quad + \mathbb{G}_{n,P}(f_{\theta_1(P),h_1} - f_{\theta_2(P),h_2})\end{aligned} \tag{24}$$

By Lemma 4 the terms in the first line following the equality are $o_P(1)$ uniformly in $P \in \mathcal{P}$. Moreover, by v. and vi. of Assumption 1 we have

$$\sqrt{n}P(f_{\theta_1(\mathbb{P}_n),h_1} - f_{\theta_1(P),h_1}) - \sqrt{n}P(f_{\theta_2(\mathbb{P}_n),h_2} - f_{\theta_2(P),h_2}) = \mathbb{G}_{n,P}(g_{\theta_1(P),h_1} - g_{\theta_2(P),h_2}) + o_P(1),$$

again uniformly in $P \in \mathcal{P}$ with $g_{\theta,h} \in \mathcal{G}$ as defined in Lemma 3. Collecting together these results permits us to write

$$\begin{aligned} &\sqrt{n}(\mathbb{P}_n f_{\theta_1(\mathbb{P}_n),h_1} - P f_{\theta_1(P),h_1}) - \sqrt{n}(\mathbb{P}_n f_{\theta_2(\mathbb{P}_n),h_2} - P f_{\theta_2(P),h_2}) \\ &\qquad = \mathbb{G}_{n,P}(f_{\theta_1(P),h_1} - f_{\theta_2(P),h_2}) + \mathbb{G}_{n,P}(g_{\theta_1(P),h_1} - g_{\theta_2(P),h_2}) + o_P(1) \end{aligned} \tag{25}$$

Recognizing that

$$(f_{\theta_1,h_1} - f_{\theta_2,h_2}) + (g_{\theta_1,h_1} - g_{\theta_2,h_2}) \in \Phi$$

we may obtain the desired result by establishing the uniform in $\mathcal{P}$ weak convergence of

$$\{\mathbb{G}_{n,P}\phi : \phi \in \Phi\}$$

Since

$$\lim_{M\to\infty} \sup_{P\in\mathcal{P}} Pe^2\{e > M\} = 0$$

follows from the fact that $e(x) \leq 2\sup_{f\in\mathcal{F}} |f(x)| + 2\sup_{g\in\mathcal{G}} |g(x)|$ and the Lindeberg condition is satisfied by the right-hand-side (by Assumptions 1(ii) and 1(ii)), and Lemma 5 establishes the finiteness of the bracketing integral, we have established, via Theorem 2.8.4 of van der Vaart and Wellner (1996), that $\mathbb{G}_{n,P} \rightsquigarrow \mathbb{G}$ in $\ell^\infty(\Phi)$ uniformly in $P \in \mathcal{P}$. □

Proof of Theorem 2. Writing

$$\mathcal{Z}_n = \sqrt{n}[\Phi_n(\theta(\mathbb{P}_n)) - \Phi(P)] + \sqrt{n}\Phi(P)$$

we have

$$\mathcal{Z}_n \rightsquigarrow \mathcal{Z} + \lim_{n\to\infty} n^{1/2}\Phi(\theta(P))$$

uniformly in $P \in \mathcal{P}$ from Corollary 1. The desired result is then a consequence of the continuity of the max function. □

Proof of Theorem 3. Writing

$$\mathcal{Z}_n^* = \sqrt{n}[\Phi_n(\theta(\mathbb{P}_n^*)) - \Phi_n(\theta(\mathbb{P}_n))] + \sqrt{n}[\Phi_n(\theta(\mathbb{P}_n)) - \max\{\Phi_n(\theta(\mathbb{P}_n)), \delta_n\}]$$

we have

$$\mathcal{Z}_n^* \rightsquigarrow \mathcal{Z} + \lim_{n\to\infty} n^{1/2} \min\{\Phi(\theta(P)), 0\}$$

uniformly in $P \in \mathcal{P}$. The result follows from Corollary 1 together with the fact that,

$$\max\{\Phi_{n,(i)}(\theta(\mathbb{P}_n)), \delta_n\} \to 0$$

if $\Phi_{n,(i)}(P) < 0$ and

$$\max\{\Phi_{n,(i)}(\theta(\mathbb{P}_n)), \delta_n\} \to \Phi_{n,(i)}(P)$$

if $\Phi_{n,(i)}(P) \geq 0$ with probability one uniformly in $P \in \mathcal{P}$. The desired result is then a consequence of the continuity of the max function. □

Proof of Theorem 4. Noting that $\hat{\tau}_n^{PC} \geq 0$ and $T_n^{PC} \to -\infty$ uniformly over $P \in \mathcal{P}_{00}$, part (i) will be proved if we can establish the validity of (ii). Part (i) of Proposition 3 establishes the convergence of the partially centered bootstrap version of T_n to the same limiting distribution as T_n uniformly over $P \in \mathcal{P}_0$. Letting T denote the limiting random variable, i.e. $T_n \to^d T$, we are required to show that for $\forall \epsilon > 0$ $\exists N$ such that

$$\sup_{P\in\mathcal{P}_0} |\mathbb{P}_P(T_n \leq \hat{\tau}_n(\alpha)) - \mathbb{P}_P(T \leq \tau(\alpha; P))| < \epsilon$$

for all $n \geq N$, where $\tau(\alpha; P) = \inf\{t : \mathbb{P}_P(T \leq t) = 1 - \alpha\}$. From Proposition 3 it follows that, for all $\epsilon > 0$ there exits an integer N such that

$$\sup_{P\in\mathcal{P}_0} |\mathbb{P}_P^*(T_n^{PC} \leq x) - \mathbb{P}_P(T \leq x)| < \epsilon$$

for all $n \geq N$ and all continuity points x. Since $\tau(\alpha; P)$ is a continuity point (by Proposition 2 and the assumption that $\alpha < 0.5$) it must be the case that

$$\sup_{P \in \mathcal{P}_0} |\hat{\tau}_n(\alpha) - \tau(\alpha, P)| \to 0. \tag{26}$$

Fix $\delta > 0$. There exists N_δ such that for all $n \geq N_\delta$

$$\begin{aligned}
&\sup_{P \in \mathcal{P}_0} |\mathbb{P}_P(T_n \leq \hat{\tau}_n(\alpha)) - \mathbb{P}_P(T \leq \tau(\alpha, P))| \\
&\quad \leq \sup_{P \in \mathcal{P}_0} |\mathbb{P}_P(T_n \leq \tau(\alpha, P) + \delta) - \mathbb{P}_P(T \leq \tau(\alpha, P))| \\
&\qquad + \sup_{P \in \mathcal{P}_0} |\mathbb{P}_P(T_n \leq \tau(\alpha, P) - \delta) - \mathbb{P}_P(T \leq \tau(\alpha, P))| \\
&\quad \leq \sup_{P \in \mathcal{P}_0} |\mathbb{P}_P(T_n \leq \tau(\alpha, P) + \delta) - \mathbb{P}_P(T \leq \tau(\alpha, P) + \delta)| \qquad (27) \\
&\qquad + \sup_{P \in \mathcal{P}_0} |\mathbb{P}_P(T \leq \tau(\alpha, P) - \delta) - \mathbb{P}_P(T \leq \tau(\alpha, P))| \\
&\qquad + \sup_{P \in \mathcal{P}_0} |\mathbb{P}_P(T_n \leq \tau(\alpha, P) + \delta) - \mathbb{P}_P(T \leq \tau(\alpha, P) + \delta)| \\
&\qquad + \sup_{P \in \mathcal{P}_0} |\mathbb{P}_P(T \leq \tau(\alpha, P) + \delta) - \mathbb{P}_P(T \leq \tau(\alpha, P))|
\end{aligned}$$

For small enough δ, $\tau(\alpha, P) - \delta$ is a continuity point of the limiting distribution. It follows that δ can be chosen such that each of the four terms in the sum is less than $\epsilon/4$. □

Proof of Theorem 5. Define

$$\begin{aligned}
&H_{(i)}(x_{(1)}, \ldots, x_{(r)}) \\
&\qquad = \lim_{n \to \infty} P\left(\sqrt{n}(\Phi_{n,(i)} - \Phi_{(i)}(P)) \leq x_{(i)}\right), \ 1 \leq i \leq r \qquad (28) \\
&\qquad = \mathbb{P}\left(Z_{(i)} \leq x_{(i)}\right), \ 1 \leq i \leq r
\end{aligned}$$

where the convergence is understood to hold uniformly in $P \in \mathcal{P}$. Letting $Z_n = \sqrt{n}(\Phi_n -$

$\Phi(P))$ we have

$$\begin{aligned} H^*_{(i)}(Z_n + \Phi(P)) &= H_{(i)}(Z_n + \Phi(P)) + H^*_{(i)}(Z_n + \Phi(P)) \\ &\quad - H_{(i)}(Z_n + \Phi(P)) \\ &\to^d H_{(i)}(Z + \lim_{n\to\infty} \Phi(P)) \end{aligned}$$

since $\rho_\infty(H^*_{(i)}, H_{(i)}) \to 0$ in probability conditional on $\mathcal{X}_n$ uniformly in $P \in \mathcal{P}$ (follows from the pointwise convergence of $H^*_{(i)}$ to $H_{(i)}$, together with the continuity of $H_{(i)}$ and Pólya's Theorem, cf. (Serfling 1981)) . Using the continuity of the minimum, it follows that

$$\hat{p}_{min} \to^d \min_{1\le i\le r} \Big(1 - H_{(i)}(Z + \lim_{n\to\infty} \Phi(P)\Big) \tag{29}$$

conditional on $\mathcal{X}_n$ uniformly in $P \in \mathcal{P}$. Using an analogous argument we have that

$$H^*_{(i)}(\Phi^{**}) \to^d H_{(i)}(Z + \tilde{\Delta})$$

where

$$\tilde{\Delta}_{(i)} = \begin{cases} 0, & \Phi_{(i)}(P) \geq 0, 1 \leq i \leq r \\ -\infty & \Phi_{(i)}(P) < 0, 1 \leq i \leq r \end{cases}$$

Consequently, we have via the continuity of the minimum that

$$\hat{p}^*_{min} \to^d \min_{1\le i\le r} \Big(1 - H_{(i)}(Z + \tilde{\Delta})\Big) \tag{30}$$

uniformly in $P \in \mathcal{P}$. The proof is then completed by replicating the proof of Theorem 4 using the results of equations (29) and (30).

□

Proof of Theorem 6. The proof of (i) is immediate upon noting that $\hat{\tau}^{PC}_n(\alpha) = O_p(1)$ whereas T_n diverges to infinity under the alternative. As for (ii) note that $\hat{p}_{min}$ converges in probability to zero under the alternative whereas $\hat{p}_{crit}$ converges to the αth quantile from the continuous distribution in (30) whose support is the interval [0,1].

$\square$

Proof of Theorem 7. Under the sequence of alternatives

$$\Psi(P_n) = \Psi(P) + n^{-1/2}\xi$$

as defined in (14), we have

$$\mathcal{Z}_n \rightsquigarrow \mathcal{Z} + \xi,$$

whereas the bootstrap version $\mathcal{Z}_n^*$ under the sequence of local alternative converges to

$$\mathcal{Z}_n^* \rightsquigarrow \mathcal{Z} + min\{\xi, 0\}.$$

By the continuous mapping theorem we have

$$T_n \rightsquigarrow \max_{1 \leq i \leq r} [\mathcal{Z} + \xi]$$

and

$$\mathcal{T}_n^* \rightsquigarrow \max_{1 \leq i \leq r} [\mathcal{Z} + min\{\xi, 0\}].$$

Clearly, for large enough n, T_n is stochastically larger than T_n^*. Consequently, for n large enough

$$\mathbb{P}_{P_n}(T_n > \tau_n^{PC}) \geq \alpha.$$

The proof of part (ii) of the theorem is analogous. $\square$

REFERENCES

Abadie, A. (2002), "Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models," *Journal of the American Statistical Association*, 97(457), 1893–1905.

Anderson, G. (1996), "Nonparametric Tests of Stochastic Dominance in Income Distributions," *Econometrica*, 64(5), 1183–1193.

Andrews, D. W., and Guggenberger, P. (2007), The Limit of Finite-Sample Size and a Problem with Subsampling,, Cowles Foundation Discussion Papers 1605, Cowles Foundation, Yale University.

Andrews, D. W., and Jia, P. (2008), Inference for Parameters Defined by Moment Inequalities: A Recommended Moment Selection Procedure,, Cowles Foundation Discussion Papers 1676, Cowles Foundation, Yale University.

Andrews, D. W., and Soares, G. (2007), Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection,, Cowles Foundation Discussion Papers 1631, Cowles Foundation, Yale University.

Angrist, J. D. (1990), "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records," *The American Economic Review*, 80(5), 1284–1286.

Barrett, G. F., and Donald, S. G. (2003), "Consistent Tests for Stochastic Dominance," *Econometrica*, 71(1), 71–104.

Bennett, C. J. (2009), p-Value Adjustments for Asymptotic Control of the Generalized Familywise Error Rate,, Working Paper 09-W05, Department of Economics, Vanderbilt University.

Chernozhukov, V., Hong, H., and Tamer, E. (2007), "Estimation and Confidence Regions for Parameter Sets in Econometric Models," *Econometrica*, 75(5), 1243–1284.

Davidson, R., and Duclos, J.-Y. (2000), "Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality," *Econometrica*, 68(6), 1435–1464.

Dufour, J.-M. (1989), "Nonlinear Hypotheses, Inequality Restrictions, and Non-nested Hypotheses: Exact Simultaneous Tests in Linear Regressions," *Econometrica*, 57(2), 335–55.

Fan, Y., and Park, S. (2007), Confidence Sets for Some Partially Identified Parameters,, Working paper.

Foster, J. E., and Sen, A. (1999), "On Economic Inequality After a Quarter Century," in *On Economic Inequality*, Oxford, U.K.: Clarendon Press, pp. 107–219.

Godfrey, L. G. (2005), "Controlling the overall significance level of a battery of least squares diagnostic tests," *Oxford Bulletin of Economics and Statistics*, 67, 267.

Goldberger, A. (1992), "One-sided and inequality tests for a pair of means," in *Contributions to Consumer Demand and Econometrics: Essays in Honour of Henri Theil*, New York: St. Martins Press, chapter 8, pp. 1–6.

Hansen, P. (2003), Asymptotic Tests of Composite Hypotheses,, Working Papers 2003-09, Brown University, Department of Economics.

Hansen, P. R. (2005), "A Test for Superior Predictive Ability," *Journal of Business & Economic Statistics*, 23, 365–380.

Kodde, D. A., and Palm, F. C. (1986), "Wald Criteria for Jointly Testing Equality and Inequality Restrictions," *Econometrica*, 54(5), 1243.

Kosorok, M. R. (2008), *Introduction to Empirical Processes and Semiparametric Inference*, New York: Springer.

Linton, O., Song, K., and Whang, Y.-J. (2008), Bootstrap tests of stochastic dominance with asymptotic similarity on the boundary,, CeMMAP working papers CWP08/08, Centre for Microdata Methods and Practice, Institute for Fiscal Studies.

MacKinnon, J. G. (2007), Bootstrap Hypothesis Testing,, Working Papers 1127, Queen's University, Department of Economics.

McFadden, D. (1989), "Testing for Stochastic Dominance," in *Studies in the Economics of Uncertainty. In Honour of Josef Hadar*, eds. T. B. Fomby, and T. K. Seo, New York: Springer Verlag, pp. 113–134.

Mikusheva, A. (2007), "Uniform Inference in Autoregressive Models," *Econometrica*, 75(5), 1411–1452.

Moon, H. R., and Schorfheide, F. (2008), Estimation with Overidentifying Inequality Moment Conditions,. Unpublished working paper, Department of Economics, University of Southern California.

Nankervis, J. C. (2005), "Computational algorithms for double bootstrap confidence intervals," *Computational Statistics & Data Analysis*, 49(2), 461–475.

Perlman, M. (1969), "One-Sided Testing Problems in Multivariate Analysis," *Annals of Mathematical Statistics*, 40, 549.

Romano, J. P., Shaikh, A. M., and Wolf, M. (2007), Control of the False Discovery Rate under Dependence using the Bootstrap and Subsampling,, IEW - Working Papers iewwp337, Institute for Empirical Research in Economics - IEW.

Rosen, A. (2006), Confidence sets for partially identified parameters that satisfy a finite number of moment inequalities,, CeMMAP working papers CWP25/06, Centre for Microdata Methods and Practice, Institute for Fiscal Studies.

Savin, N. (1984), "Multiple hypothesis testing," in *Handbook of Econometrics*, eds. Z. Griliches, and M. D. Intriligator, Vol. 2 of *Handbook of Econometrics* Elsevier, chapter 14, pp. 827–879.

Serfling, R. J. (1981), *Approximation Theorems of Mathematical Statistics (Wiley Series in Probability and Statistics)* Wiley-Interscience.

Shorrocks, A. F., and Foster, J. E. (1987), "Transfer Sensitive Inequality Measures," *Review of Economic Studies*, 54(3), 485–97.

Silvapulle, M. J., and Sen, P. K. (2004), *Constrained Statistical Inference: Inequality, Order, and Shape Restrictions*, 1 edn Wiley-Interscience.

van der Vaart, A. W. (1998), *Asymptotic Statistics* Cambridge University Press.

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

van der Vaart, A. W., and Wellner, J. A. (2007), "Empirical processes indexed by estimated functions," *IMS Lecture Notes Monograph Series*, 55, 234.

White, H. (2000), "A Reality Check for Data Snooping," *Econometrica*, 68(5), 1097.

Wolak, F. (1987), "An Exact Test for Multiple Inequality and Equality Constraints in the Linear Regression Model," *Journal of the American Statistical Association*, 82, 782.

Wolak, F. (1989), "Local and Global Testing of Linear and Nonlinear Inequality Constraints in Nonlinear Econometric Models," *Econometric Theory*, 5, 1.

Wolak, F. (1991), "The Local Nature of Hypothesis Tests Involving Inequality Constraints in Nonlinear Models," *Econometrica*, 59(4), 981.

Zheng, B., Formby, J. P., Smith, W. J., and Chow, V. K. (2000), "Inequality Orderings, Normalized Stochastic Dominance, and Statistical Inference," *Journal of Business & Economic Statistics*, 18(4), 479–88.

List of Figures

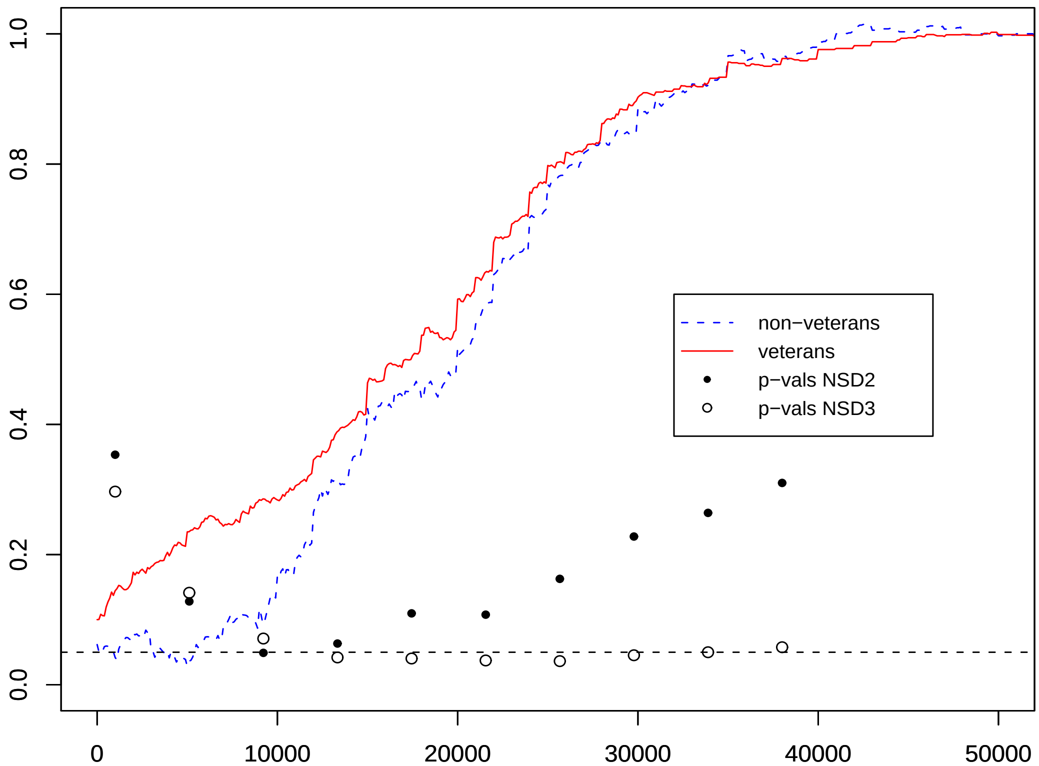

Figure 1: Plots of the empirical distributions of earnings F_1^C and F_0^C together with the adjusted p-values associated with tests of second- and third-order normalized stochastic dominance. The horizontal dashed line marks the 5% level of significance.

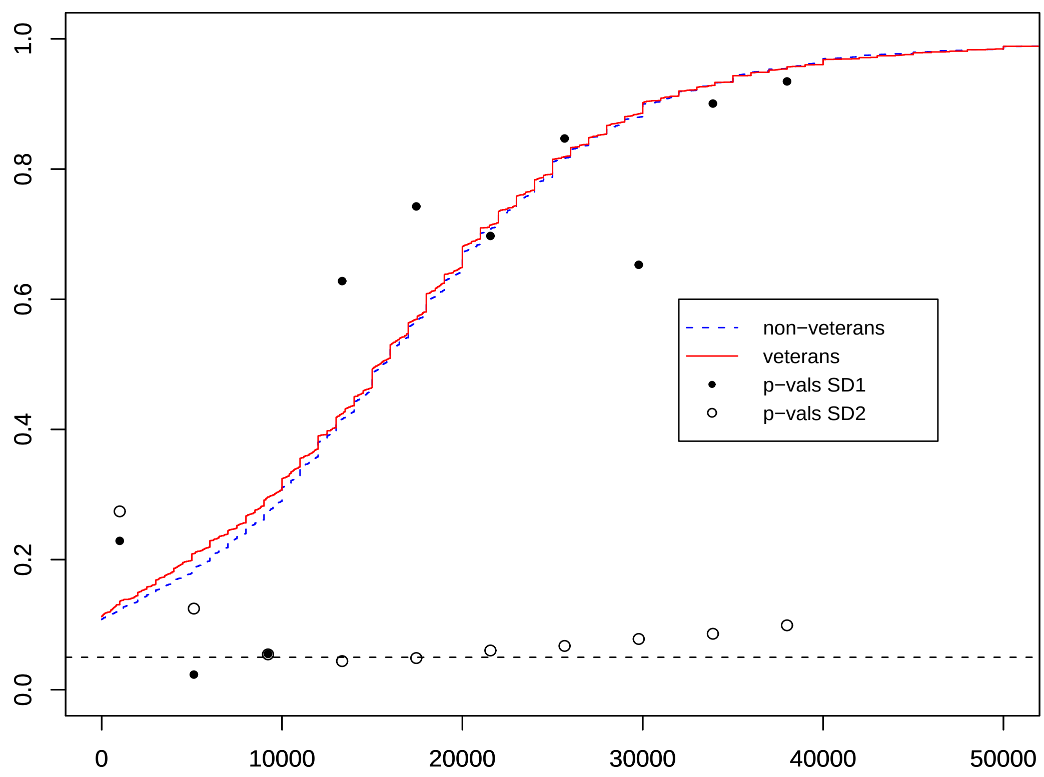

Figure 2: Plots of the empirical distributions of earnings F_{1,n_1} and F_{0,n_0} together with the adjusted p-values associated with tests of first- and second-order stochastic dominance. The horizontal dashed line marks the 5% level of significance.

List of Tables

Table 1: Experimental Designs

Design	Distributions		NSD		
	F	G	2^{nd}	3^{rd}	4^{th}
(a)	SM(1.69,1,8.37)	SM(1.69,1,8.37)	Equal	·	·
(b)	LN(60,100)	LN(60,100)	Equal	·	·
(c)	SM(1.69,1,8.37)	SM(2.15,1,1.81)	·	·	Cross
(d)	SM(3.11,1,9.30)	LN(20,100)	·	·	Cross
(e)	LN(70,100)	LN(60,100)	Dominance	·	·
(f)	SM(1.69,1,8.37)	SM(1.84,1,8.37)	Dominance	·	·

Table 2: Empirical Rejection Probabilities

Design	n	$NSD1$	$NSD1_p$	$NSD2$	$NSD2_p$	$NSD3$	$NSD3_p$
(a)	500	0.055	0.059	0.042	0.043	0.041	0.042
	1000	0.053	0.054	0.044	0.046	0.044	0.045
	2000	0.047	0.053	0.050	0.057	0.047	0.050
(b)	500	0.050	0.055	0.040	0.042	0.035	0.035
	1000	0.046	0.046	0.051	0.051	0.046	0.047
	2000	0.053	0.054	0.050	0.052	0.046	0.046
(c)	500	0.393	0.443	0.427	0.448	0.431	0.441
	1000	0.600	0.668	0.668	0.696	0.676	0.688
	1500	0.838	0.870	0.860	0.872	0.872	0.880
	2000	0.887	0.933	0.907	0.917	0.903	0.920
(d)	500	0.072	0.184	0.140	0.244	0.144	0.216
	1000	0.104	0.296	0.344	0.508	0.404	0.520
	1500	0.226	0.424	0.560	0.722	0.634	0.780
	2000	0.311	0.534	0.726	0.843	0.800	0.877
(e)	500	0.442	0.446	0.414	0.414	0.382	0.382
	1000	0.687	0.687	0.695	0.695	0.678	0.678
	1500	0.884	0.884	0.892	0.892	0.860	0.860
	2000	0.920	0.920	0.921	0.921	0.904	0.904
(f)	500	0.373	0.381	0.410	0.422	0.390	0.399
	1000	0.648	0.652	0.668	0.668	0.632	0.640
	1500	0.840	0.841	0.846	0.846	0.835	0.835
	2000	0.880	0.880	0.886	0.889	0.877	0.877

Table reports the empirical rejection probabilities for the fully recentered and partially recentered (subscript "p") MinP tests at the nominal %5 level as estimated from 1000 Monte Carlo simulations