EXAMINING THE DISTRIBUTIONAL EFFECTS OF MILITARY SERVICE ON EARNINGS: A TEST OF INITIAL DOMINANCE

by

Chrisopher J. Bennett and Ricardas Zitikis

Working Paper No. 11-W11

First version: September 2011

DEPARTMENT OF ECONOMICS
VANDERBILT UNIVERSITY
NASHVILLE, TN 37235

www.vanderbilt.edu/econ

Examining the Distributional Effects of Military Service on Earnings: A Test of Initial Dominance

Christopher J. Bennett[a,*] and Ričardas Zitikis[b]

[a] *Department of Economics, Vanderbilt University, VU Station B #351819, 2301 Vanderbilt Place, Nashville, TN 37235-1819, U.S.A.*

[b] *Department of Statistical and Actuarial Sciences, University of Western Ontario, London, Ontario N6A5B7, Canada*

Abstract

Existing empirical evidence suggests that the effects of Vietnam veteran status on earnings in the decade-and-a-half following service may be concentrated in the lower tail of the earnings distribution. Motivated by this evidence, we develop a formal statistical procedure which is specifically designed to test for lower tail dominance in the distributions of earnings. When applied to the same data as in previous studies, the test reveals that the distribution of earnings for veterans is indeed dominated by the distribution of earnings for non-veterans up to $12,610 (in 1978 dollars), thereby indicating that there was higher social welfare and lower poverty experienced by non-veterans in the decade-and-a-half following military service.

JEL classification: I31, I32, C12, C14

Keywords: Treatment effect; Potential outcome; Earnings; Hypothesis test; Stochastic dominance; Initial dominance; Crossing point; Causal effect.

---

*Corresponding author: tel.: +1 615 936-7262; fax: +1 615 343 8495; e-mail: chris.bennett@vanderbilt.edu

1 Introduction

Measuring and analyzing the effects of participation and treatments play important roles in evaluating the impact of various programs and policies, particularly in the health and social sciences (Heckman and Vytlacil, 2007; Morgan and Winship, 2007).

The social and economic costs of military service, for example, has drawn considerable attention from researchers and policy makers, in part due to continuing military engagements and questions surrounding adequate compensation of veterans (e.g., Angrist and Chen, 2008; Chaudhuri and Rose; 2009; and references therein). Indeed, a key policy-related question is whether military service tends to reduce earnings over the life-cycle.

In this paper, we focus specifically on the effect that military service in Vietnam had on subsequent earnings. Existing empirical evidence (Angrist, 1990; Abadie, 2002), for example, suggests that the effects of veteran status on earnings following service in Vietnam may have been concentrated in the lower tail of the earnings distribution. Aided with a new statistical inferential technique developed in this paper, we examine the effect of military service on the overall distribution of earnings and on the lower tail in particular, in the decade-and-a-half following the end of war in Vietnam.

Historically, the non-random selection for military service has posed a significant challenge for those researching and analyzing causal effects of service. To overcome the selection problem, Angrist (1990) has exploited the exogenous variation in the draft lottery to instrument for veteran status, thereby allowing for unbiased estimation of the average causal effect of service on earnings. Angrist (1990) has found, for example, that white male veterans experienced roughly a 15% average loss in earnings in the early 1980s. More recently, Angrist and Chen (2008) have reported that these losses in earnings dissipated over time and appear to be close to zero by the year 2000.

Abadie (2002) has also exploited the variation in the draft lottery to identify the causal effect of service on earnings, focusing on examining the effects of Vietnam veteran status over the entire distribution of earnings. Abadie (2002) has reported no statistically significant difference between the distributions of earnings for non-veterans and veterans. However, Abadie's (2002) empirical analysis suggests that the effects on earnings appear to be isolated to the lower quantiles of the earnings distributions.

Influenced by these findings, here we give a further look at the problem and investigate, roughly speaking, the range of income levels over which there are statistically significant differences in the distributions of potential earnings. A rigorous formulation of the new test, the corresponding statistical inferential theory, and subsequent empirical findings make up the main body of the present paper.

To give an indication of our findings, we note at the outset that the empirical evidence points to a distribution of earnings for non-veterans which stochastically dominates the corresponding distribution of earnings for veterans for all income levels up to \$12,610. This finding lends support to the aforementioned observation of Abadie (2002) concerning the isolated effects on the distribution of earnings. Furthermore, our findings suggest that, for any poverty line up to \$12,610 (in 1978 dollars), there was statistically greater poverty among Vietnam veterans in the early 1980s.

The rest of the paper is organized as follows. In Section 2, we briefly review the literature on testing functional hypotheses, and also introduce a new test and explain the intuition behind it. In Section 3, we apply the new test to re-examine the distributional effects of military service on civilian earnings. Section 4 contains concluding notes. Large-sample results supporting the new test are relegated to Appendix A. Monte Carlo simulations of a more detailed nature are given in Appendix B.

2 A test of initial dominance

As we have noted in the introduction, uncovering relationships between distributions of potential earnings for non-veterans (F) and veterans (G) is of considerable interest. For example, if $F \leq G$ on $[0, \infty)$, then rather powerful statements can be made concerning the comparative levels of poverty and social welfare among the groups. Specifically, if the aforementioned relationship between the cumulative distribution functions (cdf's) F and G were to hold, then income poverty would be greater for veterans according to *any* poverty index which is symmetric and monotonically decreasing in incomes. Similarly, social welfare as measured by *any* social welfare function that is symmetric and increasing in incomes would be greater for non-veterans than for veterans.

Such interest in relationships between cdf's has given rise to a large literature on statistical testing procedures. We next briefly review a few of the classical tests, which we shall in turn contrast with the test developed in the present paper. This will help us to shed light on the meaning and the novelty of the latter test.

To begin, consider a test of the null $F = G$ against the alternative $F \neq G$, both relationships on $[0,\infty)$. Such a test, which is often referred to as a test for homogeneity, is typically used to infer whether two distributions are equal or not; that is, a rejection of the null hypothesis would enable one to infer only that a difference exists between the cdf's under consideration, but not the location or the nature of this difference. Classical approaches to this testing problem include the use of $\sup_{x\geq 0}|G(x)-F(x)|$, $\int_0^\infty |G(x)-F(x)|dx$, more general L_p-versions of the latter integral, and generalizations thereof that include the use of weight functions (e.g., Csörgő and Horváth, 1993). The tests which emerge from such constructions are usually associated with the names of Kolmogorov, Smirnov, Cramér, Anderson, and Darling, in various combinations.

Intimately related to testing for homogeneity of two distributions is the so-called one-sided problem, that of testing the null $F \leq G$ against the alternative $F \nleq G$ on $[0,\infty)$, for which we may rely on one-sided variants of the aforementioned supremums and integrals: $\sup_{x\geq 0}(G(x)-F(x))_+$ and $\int_0^\infty (G(x)-F(x))_+ dx$. Because the null hypothesis corresponds to the first-order stochastic dominance, the one-sided formulations give rise to statistical tests of whether F dominates G over the entire support $[0,\infty)$. For recent contributions to the one-sided testing problem, we refer to Linton, Song, and Whang (2010), as well as to the list of references therein.

The use of the above tests presupposes that a restriction on the nature of the difference (or lack of difference) between the cdf's must hold over their entire supports $[0,\infty)$, or at least over a pre-specified subset of the supports. In other words, the classical formulations impose global restrictions on the nature of the difference between the two cdf's.

In many contexts, however, it is useful to learn both *if* and *where* a restriction holds, particularly when the restriction may hold only over some (unknown) subset of the supports. For example, suppose that we are interested in comparing poverty across two income distributions using the headcount ratio, which is the proportion of the population with incomes at or below a given poverty line. Because it is often difficult to reach a consensus on a specific value which will be used to demarcate the poverty line, an attractive procedure would be the one that would identify the maximal income level x_1 and hence the interval $[0,x_1]$ over which the poverty ranking implied by the headcount ranking is consistent. If, for example, x_1 is found to be sufficiently large by such a procedure, say so large as to constitute an upper bound for any reasonable choice of poverty line, then policymakers may reach a consensus as to the poverty ranking even without having reached a consensus on the poverty line.

Coming now back to our underlying example of Vietnam veterans, in the context of comparing earnings distributions for the veterans and non-veterans, one distribution may not dominate another over the entire support (Figure 2.1) and yet the relations that hold between these two cdf's may still be sufficient for establishing similarly powerful statements about poverty and social welfare rankings. For example, little if anything is

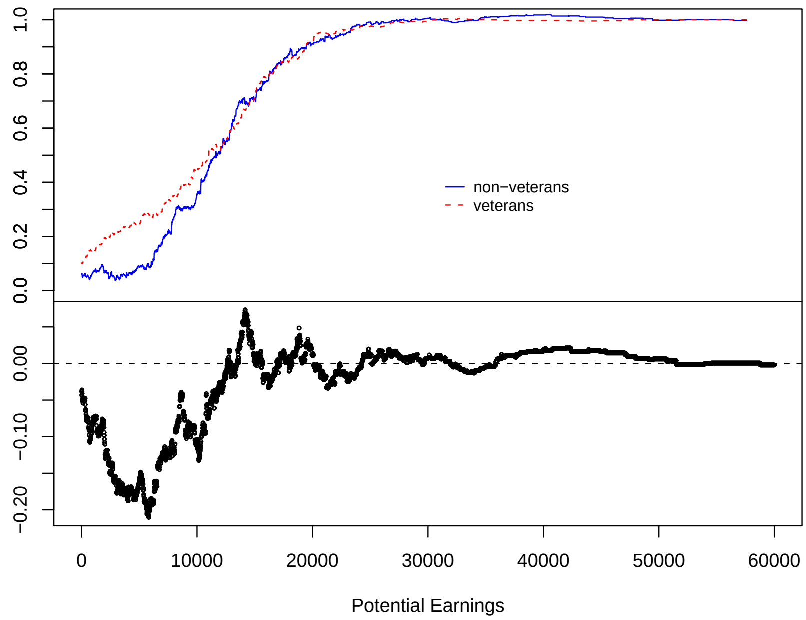


Figure 2.1: Plots of the empirical cdf's (top panel) and their difference (bottom panel) of potential earnings of veterans and non-veterans , who are compliers. (Definitions of the corresponding population cdf's F_1^C and F_0^C are given by equation (3.1) below.)

lost in the context of poverty analysis if the ordering of the distributions is violated only for "high" levels of income. Concretely, suppose that $F(x) \leq G(x)$ for all $x \in [0, x_1)$. Then for any poverty line up to x_1 (e.g., \$12,610) we have that income poverty is greater in G according to any poverty index which is symmetric and monotonically decreasing in incomes (Foster and Shorrocks, 1988). Similarly, Foster and Shorrocks (1988) have shown that when the cdf's satisfy such a relation, a 'poor focused' social welfare evaluation in which F and G are censored at x_1 suggests that social welfare is greater in F than in G for any symmetric social welfare function which is increasing in incomes. We note in this regard that the 'poor focused' evaluation is based on computing the social welfare of the censored distributions $\widetilde{F}$ and $\widetilde{G}$, where these censored distributions are generated by replacing all incomes above x_1 in F and G by x_1 itself.

To tackle such empirical questions, in the next section we develop a statistical procedure which will enable us to infer the range over which a set of dominance restrictions between two cdf's is satisfied. The approach will allow us to (i) infer dominance with probability tending to 1 as the sample size tends to infinity whenever dominance is indeed present in the population, and (ii) consistently estimate the range over which this dominance holds. Thus, our testing procedure will help us to identify situations in which F dominates G, and to also differentiate between situations in which this dominance holds over the entire support or only over some initial range.

2.1 Hypotheses

Let F and G be two cdf's, both continuous and with supports on $[0, \infty)$. Of course, we have in mind the cdf's of potential earnings for non-veterans (F) and veterans (G), but the discussion that follows is for generic F and G, unless specifically noted otherwise. We are interested in testing the null hypothesis H_0 of $F \geq G$ at least initially against the alternative H_1 of $F \leq G$ initially with possibly crossing later on. Rigorously, these hypotheses are defined as follows:

H_0 : Given a pair (F, G) of cdf's, let $x_0 \in [0, \infty]$ denote the maximal point such that $F(x) \geq G(x)$ for all $x \in [0, x_0)$. By definition, this null hypothesis consists of those pairs (F, G) for which either

(i$_0$) $x_0 = \infty$

or

(ii$_0$) $x_0 < \infty$ and there is $x_0^* \in (0, x_0)$ such that $F(x_0^*) > G(x_0^*)$.

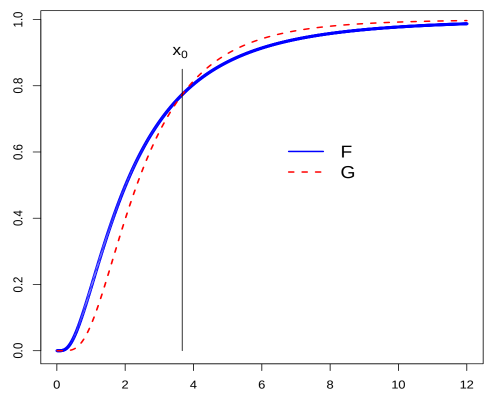
H_1 : By definition, this alternative hypothesis consists of those pairs (F, G) for which there exist $x_1 \in (0, \infty)$ and $\epsilon > 0$, depending on the pair (F, G), such that

(i$_1$) $F(x) \leq G(x)$ for all $x \in [0, x_1)$ with some $x^* \in (0, x_1)$ such that $F(x^*) < G(x^*)$,

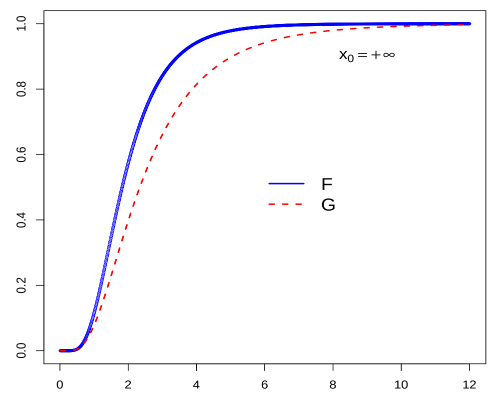
and

(ii$_1$) $F(x) > G(x)$ for all $x \in (x_1, x_1 + \epsilon)$.

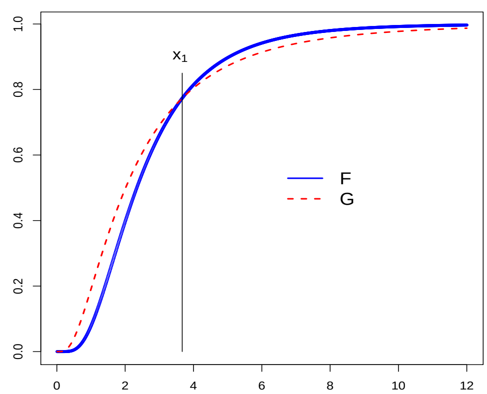
To facilitate a greater clarity of the definitions of H_0 and H_1, as well as the distinction between them, we have depicted a few scenarios in Figure 2.2.



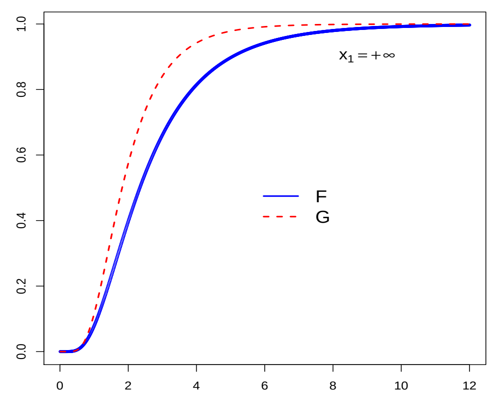
(a) Configuration in the null: crossing with no initial region of dominance



(b) Configuration in the null: G stochastically dominates F



(c) Configuration in the alternative: crossing with an initial region of dominance



(d) Configuration in the alternative: F stochastically dominantes G

Figure 2.2: Illustrations of the null H_0 and the alternative H_1.

Note that the null hypothesis H_0 includes the following sub-hypothesis

$$\mathrm{H}_0^* : \ F(x) = G(x) \text{ for all } x \in [0, \infty),$$

which is a special case of part (i_0). In fact, it is this (null) sub-hypothesis that we are particularly interested in testing against the alternative H_1, because retaining H_0^* would mean no statistical evidence for the claim that the veteran status affects incomes, whereas rejecting the sub-hypothesis in favour of H_1 would suggest the opposite, and perhaps even trigger certain policy decisions. From the technical point of view, the case $F = G$ on

$[0, \infty)$, which is at the 'boundary' between the null H_0 and the alternative H_1, will play a pivotal role when calculating critical values of the new test.

Notably, our specification of the alternative is consistent with any situation in which F dominates G on some interval $[0, x_1)$. The location of x_1, which is generally unknown, may be anywhere to the right of zero. Consequently, the alternative is consistent with strict dominance holding only over a subset of the supports of the two cdf's as well as strict dominance holding over the entire supports. When x_1 happens to be finite, so that a crossing of the two cdf's does occur, the point x_1 is just the left-most crossing point: it may or may not be the only crossing point.

2.2 Statistics and a test procedure

To construct a test for H_0 against H_1, we shall employ a one-dimensional parameter θ such that the hypotheses could be reformulated as follows:

$$\begin{aligned} \mathrm{H}_0 &: \ \theta = 0 \\ \mathrm{H}_1 &: \ \theta > 0 \end{aligned} \tag{2.1}$$

For this, we first define an auxiliary function:

$$H(y) = \int_0^y \big(F(x) - G(x)\big)_+ dx,$$

where $z_+ = 0$ when $z \le 0$ and $z_+ = z$ when $z \ge 0$. The function $H(y)$ is non-negative and non-decreasing. Define its generalized inverse by the formula

$$H^{-1}(t) = \inf\{y \ge 0 : H(y) \ge t\},$$

and then define the point

$$x_1 = \lim_{t \downarrow 0} H^{-1}(t). \tag{2.2}$$

We see that under the null H_0, the point x_1 is equal to 0, whereas under the alternative H_1, the point coincides with x_1 specified in the definition of the alternative H_1; hence, our use of the same notation. Since $x_1 = 0$ under H_0, the parameter θ defined by

$$\theta = \int_0^{x_1} \big(G(x) - F(x)\big)_+ dx$$

is equal to 0. Under the alternative H_1, however, the parameter θ is (strictly) positive. This is precisely as stated in (2.1). Note that the integrand in the definition of θ is the positive part of $G(x) - F(x)$, unlike the integrand in the definition of $H(y)$, which is the positive part of $F(x) - G(x)$.

We next construct an estimator for θ. To begin with, let $X_1, \dots, X_n$ be independent random variables from F, and let $Y_1, \dots, Y_m$ be independent random variables from G. These random variables are also assumed to be independent between the samples. Denote the corresponding empirical cdf's by F_n and G_m, and then define an estimator of $H(y)$ by the formula

$$H_{m,n}(y) = \int_0^y \big(F_n(x) - G_m(x)\big)_+ dx.$$

The corresponding generalized inverse is

$$H_{m,n}^{-1}(t) = \inf\{y \geq 0 : H_{m,n}(y) \geq t\}.$$

We are now in the position to define an estimator $x_{m,n} \equiv x_{m,n}(\delta_{m,n})$ of the point x_1, which is given by the formula

$$x_{m,n} = H_{m,n}^{-1}(\delta_{m,n}),$$

where $\delta_{m,n} > 0$ is a 'tuning' parameter such that

$$\delta_{m,n} \downarrow 0 \quad \text{and} \quad \sqrt{\frac{nm}{n+m}}\, \delta_{m,n} \to \infty \tag{2.3}$$

when $\min\{m, n\} \to \infty$. That is, the tuning parameter $\delta_{m,n}$ cannot be large nor too small. The reason why we cannot set $\delta_{m,n}$ to zero is because the function $F_n - G_m$ is 'rough' even when the population cdf's F and G are identical on $[0, \infty)$. In practice, since the sample sizes m and n are finite, the parameter $\delta_{m,n} > 0$ will be set to a small constant, but certainly not to zero. Lastly, we define an estimator $\theta_{m,n}$ of θ by the formula

$$\theta_{m,n} = \int_0^{x_{m,n}} \big(G_m(x) - F_n(x)\big)_+ dx.$$

Because the task requires some tedious mathematics, we shall develop a large-sample asymptotic theory for the above constructed estimators $x_{m,n}$ and $\theta_{m,n}$ in Appendix A. Here we only note that we retain the null H_0 when the value of $\sqrt{mn/(m+n)}\ \theta_{m,n}$ is small and reject when it is large, with 'small' and 'large' determined by a critical value calculated using the following bootstrap algorithm (cf., e.g., Horváth et al., 2006):

i) Form the pooled distribution $L_{m,n}(x) = (mG_m(x) + nF_n(x))/(m+n)$.

ii) Generate mutually independent i.i.d. samples $X_1^*, \ldots, X_n^*$ and $Y_1^*, \ldots, Y_m^*$ from the cdf $L_{m,n}$, assuming that the values $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ are fixed, as is the case in practice, given data.

iii) Compute $x_{m,n}^*$ and $\theta_{m,n}^*$, which are the bootstrap analogues of $x_{m,n}$ and $\theta_{m,n}$ based on the bootstrap samples generated in step ii).

iv) Repeat steps ii)–iii) B times and record $\{\theta_{m,n,b}^*, 1 \leq b \leq B\}$.

The nominal α level critical value, which we denote by $c_{m,n}(\alpha)$, is then computed as the $(1-\alpha)$ quantile of the distribution of bootstrap estimates $\theta_{m,n,1}^*, \ldots, \theta_{m,n,B}^*$:

$$c_{m,n}(\alpha) = \inf\left\{c : \frac{1}{B}\sum_{b=1}^{B} \mathbf{1}\left\{\sqrt{\frac{mn}{m+n}}\,\theta_{m,n,b}^* \leq c\right\} \geq 1-\alpha\right\}.$$

With this critical value $c_{m,n}(\alpha)$, the decision rule at the nominal level α is to reject H_0 and thus infer that there is (strict) dominance over the interval $[0, x_{m,n})$ whenever

$$\sqrt{\frac{mn}{m+n}}\,\theta_{m,n} > c_{m,n}(\alpha).$$

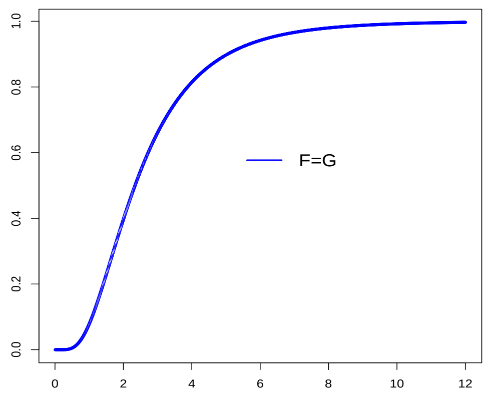
We next illustrate the performance of the proposed testing procedure, with more detailed simulation results relegated to Appendix B. For this, we generate independent log-normal samples $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ according to $X_i = \exp(\sigma_1 Z_{1i} + \mu_1)$ and $Y_i = \exp(\sigma_2 Z_{2i} + \mu_2)$, where the Z_{ki}'s are independent standard normal random variables. Various choices of the parameter-pairs (μ_i, σ_i) are explored in the simulation study, and they are specified in the left-hand panels of Figure 2.3. The figure illustrates the performance of the test under three different scenarios:

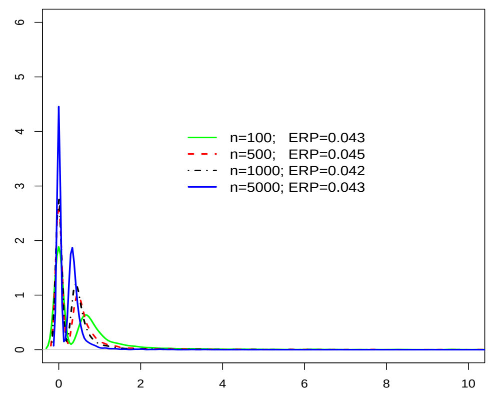
Panels (a)–(b): $F = G$.

Panels (c)–(d): F lies below G up to a point and crosses above thereafter.

Panels (e)–(f): F lies above G initially and crosses below at some later point.

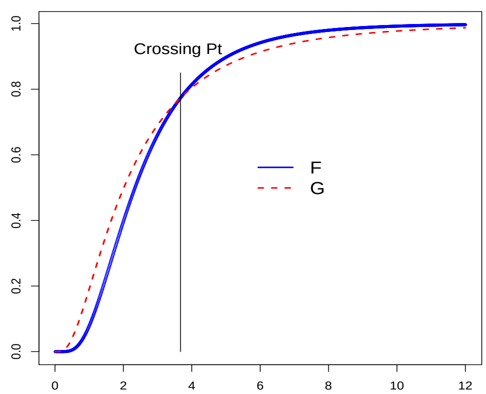
Panels (a)–(b) and (c)–(d) are in the null H_0 with the panels (a)–(b) reflecting the 'boundary' of the null. Panels (c)–(d), on the other hand, reflect a configuration of F and G in the alternative H_1. The three left-hand panels (a), (c), and (e) illustrate the relationships between the cdf's under consideration, and the right-hand panels (b), (d), and (f) illustrate the densities of the estimated crossing points for different sample sizes, along with the corresponding empirical rejection probabilities (EPR's).

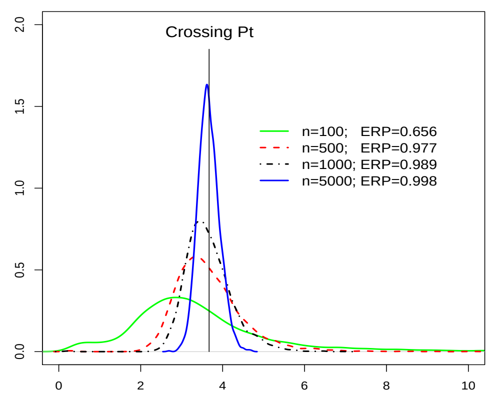




(a) $F = \text{LN}(0.85, 0.6) = G$

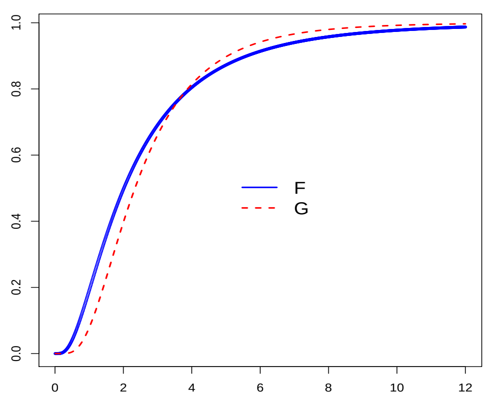
(b) Density plots of estimated crossing

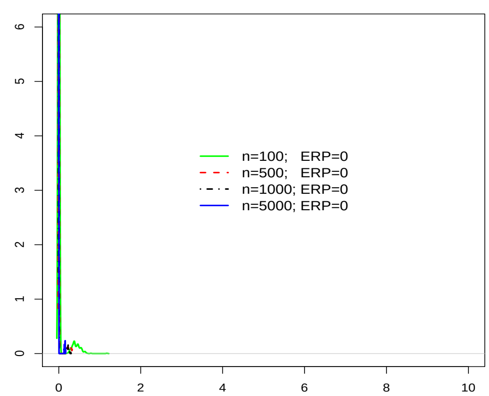




(c) $F = \text{LN}(0.85, 0.6)$, $G = \text{LN}(0.7, 0.8)$

(d) Density plots of estimated crossing points





(e) $F = \text{LN}(0.7, 0.8)$, $G = \text{LN}(0.85, 0.6)$

(f) Density plots of estimated crossing points

Figure 2.3: Monte Carlo illustrations based on parent cdf's (left-hand panels) with the corresponding density plots (right-hand panels) of estimated crossing points along with ERPs at the 5% nominal level.

We have conducted 1,000 bootstrap resamples in each of the 5,000 Monte Carlo replications, setting $\delta_{m,n} = \kappa((m+n)/(mn))^{1/2-\epsilon}$ with $\kappa = 10^{-6}$ and $\epsilon = 0.01$. The following information can be gleaned from the three sets of plots. First, the empirical rejection probabilities in the 'boundary' case $F = G$ are consistently close to the nominal size of 5% in samples ranging from $n = m = 100$ observations through $5,000$ observations, suggesting that the test maintains the desired control over the Type I error rate. The corresponding density plots demonstrate that the estimated crossing point in this case is generally close to zero, indicating that an erroneous inference of strict dominance may be coupled with an estimated interval of dominance, which is rather small. Interestingly, one might view a small interval estimate in such a case as mitigating the cost associated with the erroneous inference of strict dominance.

We see from panel (b) that the crossing-point estimator appears to be biased downward in small samples, albeit with the degree of bias dissipating in larger samples. The bias, in fact, is largely a reflection of the choice of $\delta_{m,n}$. Recall that $\delta_{m,n}$ is a parameter which, roughly speaking, controls the number of times F_n is permitted to cross above G_m before a crossing point is declared. When $\delta_{m,n}$ is large, then a substantial amount of in-sample crossing is permitted before a true crossing is declared. In contrast, when $\delta_{m,n}$ is small, then typically the first crossing in-sample also corresponds to the estimated point of crossing. In the simulations used to generate the results depicted in panel (b), the value of κ in $\delta_{m,n}$ is set to 10^{-6}, thus generating an estimator which is quite sensitive to crossings in-sample and therefore biased downward. Because this bias tends to have little impact on power (Appendix B), one may interpret the effect of the bias as producing a conservative estimate on the range over which strict dominance holds. In our view it is perhaps better to adopt such a conservative approach, particularly in small sample situations.

When there is an initial region of dominance as in panel (c) of Figure 2.3, it is desirable that our test procedure have appreciable power to detect the presence of dominance and to also return a reliable estimate of the region over which dominance holds. Panel (d) of Figure 2.3 reports that the test has considerable power to detect the initial region of dominance even in samples consisting of only 100 observations. As expected, the density plots show that in such small samples there is a great deal of variation in the estimated crossing point about the true value, and we also see that the densities of the estimated crossing point become increasingly concentrated around the true value as the sample sizes are increased.

The bottom pair of panels (e) and (f) correspond to the null configuration in which F is strictly dominated by G, at least initially. In this case, we see from the concentration of the density plots about the origin that the initial dominance of G over F is detected in virtually all of the Monte Carlo trials. More importantly, there is not a single instance in which the null is erroneously rejected.

3 Vietnam veteran data re-examined

Here we apply the new test to re-examine distributional effects of Vietnam veteran status on labour earnings. Abadie (2002) shows that the potential distributions for veterans and non-veterans can be estimated for the sub-population of compliers, and he uses Kolmogorov-Smirnov type tests to empirically examine whether the earnings distribution of non-veterans stochastically dominates that of veterans. Using the same data set, we shall see below that the new test allows us to make statistically-significant statements concerning the effect of veteran status on poverty.

3.1 The potential-outcomes model

Let Z be a binary instrument, taking on the values 1 (draft eligible) and 0 (not draft eligible) with some positive probabilities, and let $D(1)$ and $D(0)$ denote the values of the treatment indicator D that would be obtained given $Z = 1$ and $Z = 0$, respectively. Both $D(1)$ and $D(0)$ are random, taking on the values 1 (serve in the military) and 0 (do not serve in the military) with some positive probabilities. The binary nature of both the instrument and the treatment variables gives rise to four possible types of individuals in the population:

- Compliers when $D(1) = 1$ and $D(0) = 0$
- Always-takers when $D(1) = 1$ and $D(0) = 1$
- Never-takers when $D(1) = 0$ and $D(0) = 0$
- Defiers when $D(1) = 0$ and $D(0) = 1$

Because our interest centers on the causal effect of military service on earnings, let $Y(1)$ denote the potential incomes for individuals if they served ($D = 1$) and $Y(0)$ denote the potential incomes for the same individuals had they not served ($D = 0$). In practice, the analyst observes the realized outcomes

$$Y = Y(1)D + Y(0)(1 - D),$$

where $D = D(1)Z + D(0)(1 - Z)$. Within this framework of potential incomes, the distributional effect of military service in the general population is captured by the difference between the cdf's of $Y(1)$ and $Y(0)$. However, neither of the two cdf's is identified because, among other things, the data are (i) uninformative about $Y(1)$ for the sub-population of never-takers, and also (ii) uninformative about $Y(0)$ in the sub-population of always-takers. Because of this identification problem, it is common in the treatment effects literature to focus the analysis on the sub-population of compliers. In the context of examining for distributional treatment effects, this amounts to comparing the conditional cdf's F_1^C and F_0^C defined by

$$F_k^C(y) = \mathbf{P}\big[Y(k) \leq y |\ D(1) = 1, D(0) = 0\big], \tag{3.1}$$

where, naturally, we assume that $\mathbf{P}[D(1) = 1, D(0) = 0\big] > 0$, which means that compliers are present in the general population.

3.2 An underlying theory

When comparing F_1^C and F_0^C, we are particularly interested in establishing whether there exists a level of income y_1 such that $F_0^C(y) \leq F_1^C(y)$ for all $y \in [0, y_1)$ with the inequality being strict over some subset of $[0, y_1)$. Inferring the existence of such an income level would allow us to make statistically significant statements concerning poverty and social welfare orderings of the two distributions.

We note in this regard that any member of the popular FGT (Foster, Greer, and Thorbecke, 1984) class of poverty measures would indicate that poverty is at least as great or greater among those in the sub-population of compliers who served in the military for any poverty line in the interval $[0, y_1)$. Similarly, any monotonic utilitarian social welfare function when applied to the censored distributions generated by replacing incomes above y^* by y^* itself would rank the social welfare of non-veterans equal to or higher than veterans for any $y^* \in [0, y_1)$. We refer to, e.g., Foster and Shorrocks (1988) for details.

The task of testing for an initial region of dominance among the sub-population of compliers can be simplified by the fact that we do not actually need to estimate the two complier cdf's but rather examine the sign of the difference between these cdf's. This appears to be possible even under somewhat weaker assumptions than those noted by Abadie (2002, p. 285).

To substantiate the latter statement, we next show that the relationship of dominance, or lack thereof, between F_1^C and F_0^C can be investigated in terms of two other cdf's, namely F_1 and F_0 defined by

$$F_k(y) = \mathbf{P}\big[Y \leq y|\ Z = k\big]. \tag{3.2}$$

In particular, while maintaining the standard assumption of no defiers, we show that the cdf's F_1 and F_0 can be used to investigate dominance under only Condition 1(i) of Imbens and Angrist (1994), without imposing any restriction on the relative sizes of the complier and non-complier sub-populations. Specifically, we do not require that $\mathbf{P}[D(1) = 1] > \mathbf{P}[D(0) = 1]$.

The following theorem formalizes the above statements.

Theorem 3.1 *Assume that, for $z = 0$ and also for $z = 1$, the triplet $\{Y(1), Y(0), D(z)\}$ is independent of Z. If there are no defiers, then for every y we have that*

$$F_1(y) - F_0(y) = \big(F_1^C(y) - F_0^C(y)\big)\mathbf{P}\big[D(1) = 1, D(0) = 0\big]. \tag{3.3}$$

Hence, in particular, $F_1(y) \geq F_0(y)$ if and only if $F_1^C(y) \geq F_0^C(y)$.

Proof. We start with a string of elementary equations:

$$\begin{aligned}
F_1(y) - F_0(y) &= \mathbf{E}\big[1\{Y \leq y\}|Z = 1\big] - \mathbf{E}\big[1\{Y \leq y\}|Z = 0\big] \\
&= \Big(\mathbf{E}\big[1\{Y \leq y\}D|Z = 1\big] - \mathbf{E}\big[1\{Y \leq y\}D|Z = 0\big]\Big) \\
&\quad + \Big(\mathbf{E}\big[1\{Y \leq y\}(1 - D)|Z = 1\big] - \mathbf{E}\big[1\{Y \leq y\}(1 - D)|Z = 0\big]\Big) \\
&= \Big(\mathbf{E}\big[1\{Y(1)D(1) + Y(0)(1 - D(1)) \leq y\}D(1)|Z = 1\big] \\
&\qquad - \mathbf{E}\big[1\{Y(1)D(0) + Y(0)(1 - D(0)) \leq y\}D(0)|Z = 0\big]\Big) \\
&\quad + \Big(\mathbf{E}\big[1\{Y(1)D(1) + Y(0)(1 - D(1)) \leq y\}(1 - D(1))|Z = 1\big] \\
&\qquad - \mathbf{E}\big[1\{Y(1)D(0) + Y(0)(1 - D(0)) \leq y\}(1 - D(0))|Z = 0\big]\Big).
\end{aligned} \tag{3.4}$$

Now we use the assumption that, for $z = 0$ and also for $z = 1$, the triplet $\{Y(1), Y(0), D(z)\}$ is independent of Z. We thus have from equation (3.4) that

$$F_1(y) - F_0(y) = A(y) + B(y), \tag{3.5}$$

where

$$A(y) = \mathbf{E}\big[1\{Y(1)D(1) + Y(0)(1 - D(1)) \leq y\}D(1)\big] \\ - \mathbf{E}\big[1\{Y(1)D(0) + Y(0)(1 - D(0)) \leq y\}D(0)\big]$$

and

$$B(y) = \mathbf{E}\big[1\{Y(1)D(1) + Y(0)(1 - D(1)) \leq y\}(1 - D(1))\big] \\ - \mathbf{E}\big[1\{Y(1)D(0) + Y(0)(1 - D(0)) \leq y\}(1 - D(0))\big].$$

We proceed working with the expectations making up the definitions of $A(y)$ and $B(y)$ by splitting each of the expectations into four parts according to the subdivision of the sample space of all individuals into four groups, as mentioned at the beginning of this section. Then we observe that some of the resulting expectations are equal to 0 and some of them cancel out. Note in this regard that the assumption that there are no defiers in the population means that $\mathbf{P}[D(1) = 0, D(0) = 1] = 0$. In summary, we have arrived at the following equations:

$$A(y) = \mathbf{P}\big[Y(1) \leq y |\ D(1) = 1, D(0) = 0\big]\mathbf{P}\big[D(1) = 1, D(0) = 0\big] \\ - \mathbf{P}\big[Y(1) \leq y |\ D(1) = 0, D(0) = 1\big]\mathbf{P}\big[D(1) = 0, D(0) = 1\big]$$

and

$$B(y) = \mathbf{P}\big[Y(0) \leq y |\ D(1) = 0, D(0) = 1\big]\mathbf{P}\big[D(1) = 0, D(0) = 1\big] \\ - \mathbf{P}\big[Y(0) \leq y |\ D(1) = 1, D(0) = 0\big]\mathbf{P}\big[D(1) = 1, D(0) = 0\big].$$

Plugging in these expressions for $A(y)$ and $B(y)$ on the right-hand side of equation (3.5), and also recalling the definitions of F_1^C and F_0^C, we obtain equation (3.3). This concludes the proof of Theorem 3.1. ■

Hence, indeed, Theorem 3.1 says that the difference $F_1 - F_0$ is proportional to the difference $F_1^C - F_0^C$, and the proportionality coefficient is the population proportion of compliers.

3.3 The veteran data and test results

In our empirical analysis we use the CPS extract (see Angrist, 2011) that was especially prepared for Angrist and Krueger (1992). This same data set has also been used by Abadie

(2002). As described in the latter paper, the data consist of 11,637 white men, born in 1950–1953, from the March Current Population Surveys of 1979 and 1981–1985. For each individual in the sample, annual labor earnings (in 1978 dollars), Vietnam veteran status, and an indicator of draft eligibility based on the Vietnam-era draft lottery outcome are provided. Following Angrist (1990) and Abadie (2002), we use draft-eligibility as an instrument for veteran status. The construction of the draft eligibility variable is described 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 empirical counterparts of F_1 and F_0 are simpler to work with, given the data set that we are exploring, and are thus adopted in our test of initial stochastic dominance. Figure 3.1 displays the empirical distributions F_{1,n_1} and F_{0,n_0}. (Compare this figure with

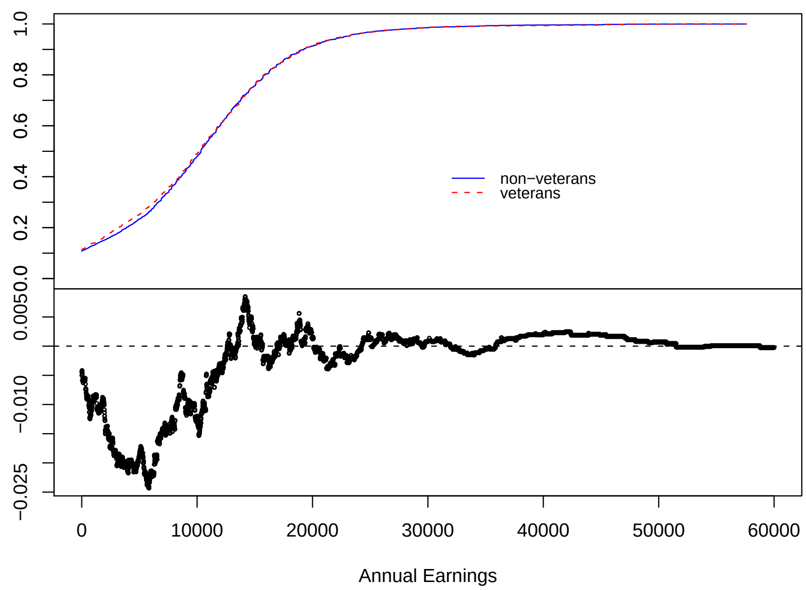


Figure 3.1: Plots of the empirical cdf's (top panel) and their difference (bottom panel) of actual earnings of draft-eligible and non-draft-eligible, who may or may not be compliers. (Definitions of the corresponding population cdf's F_1 and F_0 are in equation (3.2) above.)

Figure 2.1, where the latter depicts the empirical distributions F^C_{1,n_1} and F^C_{0,n_0}.) Applying this new test procedure, the estimated maximal point of dominance is \$12,610 with a corresponding bootstrap p-value of 0.068. Thus, for example, adopting a significance level $\alpha > 0.068$ would lead one to reject the null hypothesis and conclude that there is statistically greater poverty in the earnings distribution of Vietnam veterans for any poverty line up to $12,610.

4 Concluding remarks

In many practical situations, it is unlikely that a dominance relation holds over the entire support of the distributions under consideration. However, a number of results in the literature (e.g., Foster and Shorrocks, 1988) demonstrate that even when the dominance relation holds only on a subset of the support we may still obtain powerful orderings in terms of poverty or social welfare. With this in mind, we have developed a statistical test that enables one to infer whether there exists an initial region of stochastic dominance, and also to infer the range over which dominance holds. Notably, the test can also be used to establish dominance over the entire support. This occurs whenever dominance is indicated by our procedure *and* there is an absence of crossing in-sample.

Our work on this statistical inference problem has been inspired by existing empirical evidence (Angrist,1990; Abadie 2002) suggesting that the effects of Vietnam veteran status on earnings are concentrated in the lower tail of the earnings distribution. In particular, the inferential procedure developed in the present paper was specifically designed to test for the existence of lower tail dominance in the distributions of earnings. When applied to the same data used in previous studies, our test indicates that the distribution of earnings for veterans is dominated by the distribution of earnings for non-veterans up to $12,610 (in 1978 dollars), thereby suggesting that there was higher social welfare and lower poverty experienced by non-veterans in the decade-and-a-half following military service.

Acknowledgements

The second author thanks the Department of Economics at Vanderbilt University for hospitality during his stay at the university while working on the project, and he also gratefully acknowledges the financial support for his visit provided by the Grey fund. The support by the Natural Sciences and Engineering Research Council (NSERC) of Canada is also gratefully acknowledged.

References

[1] Abadie, A., 2002. Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models. *Journal of the American Statistical Association* 97: 1893-1905.

[2] Angrist, J.D., 2011. CPS extract [data file]. `http://econ-www.mit.edu/faculty/angrist/data1/data/angkru95`. Retrieved May 5, 2011.

[3] Angrist, J D., 1990. Lifetime Earnings and the Vietnam-Era Draft Lottery: Evidence from Social Security Administrative Records. *American Economic Review* 9: 317-323.

[4] Angrist, J.D., Chen, S., 2008. Long-Term Economic Consequences of Vietnam-Era Conscription: Schooling, Experience, and Earnings. *IZA Discussion Papers* No. 3628

[5] Angrist, J.D., Krueger, A.B., 1992. Estimating the Payoff to Schooling Using the Vietnam-Era Draft Lottery. *NBER Working Paper* No. 4067.

[6] Chaudhuri, S., Rose E., 2009. Estimating the Veteran Effect with Endogenous Schooling When Instruments are Potentially Weak. *IZA Discussion Papers* No. 4203

[7] Csörgő, M., Horváth, L., 1993. *Weighted Approximations in Probability and Statistics.* Chichester: Wiley.

[8] Foster, J., Greer, J., Thorbecke., E., 1984. A Class of Decomposable Poverty Measures. *Econometrica* 52: 761-776.

[9] Foster, J.E., Shorrocks, A.F., 1988. Poverty Orderings and Welfare Dominance. *Social Choice and Welfare* 5: 179-198.

[10] Heckman, J. J., Vytlacil, E.J., 2007. Econometric Evaluation of Social Programs, Part I: Causal Models, Structural Models and Econometric Policy Evaluation, in *Handbook of Econometrics Vol. 6*, ed. by J.J. Heckman and E.E. Leamer. New York: Elsevier, 4779-4874

[11] Horváth, L., Kokoszka, P., Zitikis, R., 2006. Testing for Stochastic Dominance using the Weighted McFadden-type Statistic. *Journal of Econometrics* 133: 191-205.

[12] Imbens, G.W., Angrist, J.D., 1994. Identification and Estimation of Local Average Treatment Effects. *Econometrica* 62: 467-475

[13] Linton, O., Song, K., Whang, Y.-J., 2010. An Improved Bootstrap Test for Stochastic Dominance. *Journal of Econometrics* 154: 186-202.

[14] Morgan, S.L., Winship, C., 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research.* Cambridge: Cambridge University Press.

A Appendix: Large sample properties

We first investigate the consistency of the estimator $x_{m,n}$ when the point x_1 is infinite (Theorem A.1) and when it is finite (Theorem A.2). In the remaining four theorems, A.3–A.6, we establish the rate of consistency of the estimator $\theta_{m,n}$ as well as its asymptotic distribution under the (null) sub-hypothesis H_0^*. Throughout, we use the notation

$$\Delta_{m,n}(x) = (F_n(x) - F(x)) - (G_n(x) - G(x)).$$

Theorem A.1 *Under* H_0^* *(i.e.,* $F = G$ *on* $[0, \infty)$*) or, more generally, when* $F \leq G$ *on* $[0, \infty)$*, we have that* $x_1 = \infty$ *and* $x_{m,n} \to_{\mathbf{P}} \infty$ *when* $\min\{m, n\} \to \infty$.

Proof. Since $F \leq G$ on $[0, \infty)$ by assumption, the function H is equal to 0 on the entire half-line $[0, \infty)$, and thus $x_1 = \infty$. We shall next show that the estimator $x_{m,n} \equiv x_{m,n}(\delta_{m,n})$ converges to ∞ in probability, which means that, for every $M > 0$,

$$\mathbf{P}[H_{m,n}^{-1}(\delta_{m,n}) \geq M] \to 1. \tag{A.1}$$

Since $H_{m,n}^{-1}(\delta_{m,n}) \geq M$ is equivalent to $\delta_{m,n} \geq H_{m,n}(M)$, we have that

$$\mathbf{P}[H_{m,n}^{-1}(\delta_{m,n}) \geq M] = \mathbf{P}\left[\sqrt{\frac{nm}{n+m}}\delta_{m,n} \geq \sqrt{\frac{nm}{n+m}}H_{m,n}(M)\right].$$

In view of assumption (2.3), statement (A.1) holds provided that

$$\sqrt{\frac{nm}{n+m}}H_{m,n}(M) = O_{\mathbf{P}}(1),$$

which is a consequence of

$$\sqrt{\frac{nm}{n+m}}\int_0^M \big|\Delta_{m,n}(x)\big|dx = O_{\mathbf{P}}(1)$$

and

$$\sqrt{\frac{nm}{n+m}}\int_0^M (F(x)-G(x))_+ dx \le 0.$$

This concludes the proof of statement (A.1) as well as that of Theorem A.1. ■

Theorem A.2 *Under the alternative* H_1*, and also under the null* H_0*, with the exception of the case when* $F = G$ *on* $[0,\infty)$ *which is covered by Theorem A.1, we have that the point* $x_1 \ge 0$ *is finite and* $x_{m,n} \to_{\mathbf{P}} x_1$ *when* $\min\{m,n\} \to \infty$*.*

Proof. Keeping in mind that the cdf's F and G are not identical throughout this proof, and irrespectively of whether we are dealing with the null H_0 or the alternative H_1, we always have $x_1 \in [0,\infty)$ and $\epsilon > 0$ such that

1) $F(x) \le G(x)$ for all $x \in (0, x_1]$, and

2) $F(x) > G(x)$ for all $x \in (x_1, x_1+\epsilon)$.

We want to show that $x_{m,n} \to_{\mathbf{P}} x_1$, which is equivalent to the statement that $\mathbf{P}[|H_{m,n}^{-1}(\delta_{m,n}) - x_1| > \gamma] \to 0$ for every $\gamma > 0$. This follows if

$$\mathbf{P}\big[H_{m,n}^{-1}(\delta_{m,n}) > x_1+\gamma\big] \to 0 \tag{A.2}$$

and

$$\mathbf{P}\big[H_{m,n}^{-1}(\delta_{m,n}) \le x_1-\gamma\big] \to 0. \tag{A.3}$$

We first establish statement (A.2):

$$\begin{aligned}
&\mathbf{P}[H_{m,n}^{-1}(\delta_{m,n}) > x_1+\gamma] \\
&\qquad = \mathbf{P}[\delta_{m,n} > H_{m,n}(x_1+\gamma)] \\
&\qquad = \mathbf{P}\left[\delta_{m,n} > \int_0^{x_1+\gamma} \Big(\Delta_{m,n}(z) + (F(z)-G(z))\Big)_+ dz\right] \\
&\qquad \le \mathbf{P}\left[\delta_{m,n} > \int_{x_1}^{x_1+\min\{\gamma,\epsilon\}} \Big(\Delta_{m,n}(z) + (F(z)-G(z))\Big)_+ dz\right] \\
&\qquad \le \mathbf{P}\left[\delta_{m,n} > -\int_{x_1}^{x_1+\min\{\gamma,\epsilon\}} \big|\Delta_{m,n}(z)\big| dz + \int_{x_1}^{x_1+\min\{\gamma,\epsilon\}} \big(F(z)-G(z)\big)_+ dz\right].
\end{aligned}$$

The right-hand side converges to 0 because $\delta_{m,n} \to 0$,

$$\int_{x_1}^{x_1+\min\{\gamma,\epsilon\}} \big|\Delta_{m,n}(z)\big| dz = o_{\mathbf{P}}(1)$$

and

$$\int_{x_1}^{x_1+\min\{\gamma,\epsilon\}} \big(F(z)-G(z)\big)_+ dz > 0.$$

This concludes the proof of statement (A.2).

To prove statement (A.3), we write:

$$\begin{aligned}
\mathbf{P}[H_{m,n}^{-1}(\delta_{m,n}) \le x_1-\gamma] & \\
&= \mathbf{P}\Bigg[\sqrt{\frac{nm}{n+m}}\delta_{m,n} \le \sqrt{\frac{nm}{n+m}} H_{m,n}(x_1-\gamma)\Bigg] \\
&= \mathbf{P}\Bigg[\sqrt{\frac{nm}{n+m}}\delta_{m,n} \le \sqrt{\frac{nm}{n+m}} \int_0^{x_1-\gamma} \big(\Delta_{m,n}(z) + (F(z)-G(z))\big)_+ dz\Bigg] \\
&\le \mathbf{P}\Bigg[\sqrt{\frac{nm}{n+m}}\delta_{m,n} \le \sqrt{\frac{nm}{n+m}} \int_0^{x_1-\gamma} \big|\Delta_{m,n}(z)\big| dz\Bigg], \qquad \text{(A.4)}
\end{aligned}$$

where the right-most bound holds because $F(z) \le G(z)$ for all $z \in (0, x_1-\gamma]$ irrespectively of the value of $\gamma > 0$. The right-hand side of bound (A.4) converges to 0 in probability because of assumption (2.3) and the fact that

$$\sqrt{\frac{nm}{n+m}} \int_0^{x_1-\gamma} \big|\Delta_{m,n}(z)\big| dz = O_{\mathbf{P}}(1).$$

This concludes the proof of statement (A.3) as well as that of Theorem A.2. ■

With the above two theorems concerning $x_{m,n}$, we are now in the position to investigate the asymptotic behaviour of the estimator $\theta_{m,n}$. Denote

$$\Theta_{m,n} = \sqrt{\frac{nm}{n+m}}\, \theta_{m,n}.$$

When $F \ge G$ on $[0,\infty)$, and in particular when $F = G$ on $[0,\infty)$, then $\theta = 0$ and so $\Theta_{m,n}$ is equal to $\sqrt{nm/(n+m)}\,(\theta_{m,n} - \theta)$. We shall use the case $F = G$ on $[0,\infty)$, which defines H_0^*, to calculate critical values of the test.

Theorem A.3 *Suppose that X and Y have $2+\kappa$ finite moments for some $\kappa > 0$, no matter how small. Under the (null) sub-hypothesis H_0^*, that is, when $F = G$ on $[0,\infty)$, we have that*

$$\Theta_{m,n} \to_d \int_0^\infty \big(\mathcal{B}(F(x))\big)_+ dx, \qquad \text{(A.5)}$$

where $\mathcal{B}$ denotes the standard Brownian bridge.

Proof. We know from Theorem A.1 that when $F = G$ on $[0,\infty)$, then $x_1 = \infty$ and $x_{m,n} \to_{\mathbf{P}} \infty$. Next we rewrite $\Theta_{m,n}$ as follows:

$$\Theta_{m,n} = -\int_{x_{m,n}}^\infty \sqrt{\frac{nm}{n+m}} \big(\Delta_{m,n}(z)\big)_+ dz + \int_0^\infty \sqrt{\frac{nm}{n+m}} \big(\Delta_{m,n}(z)\big)_+ dz.$$

Choose any (small) $\nu > 0$. We have that

$$\begin{aligned}
&\int_{x_{m,n}}^{\infty} \sqrt{\frac{nm}{n+m}} \Big(\Delta_{m,n}(z)\Big)_+ dz \\
&\quad = \int_{x_n}^{\infty} \sqrt{\frac{nm}{n+m}} \left(\frac{(F_n(z) - F(z))}{(1-F(z))^{1/2-\nu}} - \frac{(G_m(z) - G(z))}{(1-F(z))^{1/2-\nu}} \right)_+ (1-F(z))^{1/2-\nu} dx \\
&\quad = O_{\mathbf{P}}(1) \int_{x_{m,n}}^{\infty} (1-F(z))^{1/2-\nu} dz.
\end{aligned}$$

The right-hand side is of the order $o_{\mathbf{P}}(1)$ because $x_{m,n} \to_{\mathbf{P}} \infty$ and the integral $\int_0^\infty (1 - F(z))^{1/2-\nu} dz$ is finite for any sufficiently small $\nu > 0$, due to the assumption that X has $2 + \kappa$ finite moments for some $\kappa > 0$. Furthermore, since $F = G$ on $[0, \infty)$, we have that

$$\int_0^{\infty} \sqrt{\frac{nm}{n+m}} \big(\Delta_{m,n}(z)\big)_+ dz \to_d \int_0^{\infty} \big(\mathcal{B}(F(z))\big)_+ dz.$$

This concludes the proof of Theorem A.3. ■

Theorem A.4 *Under the null* H_0 *with the exception of the case* $F = G$ *on* $[0, \infty)$ *which has been covered by Theorem A.3, and assuming that* $m/(n+m) \to \eta \in (0, \infty)$ *when* $\min\{m, n\} \to \infty$, *we have that*

$$\Theta_{m,n} \to_d \int_0^{x_1} \big(\mathcal{B}(F(z))\big)_+ dz, \tag{A.6}$$

where x_1 *[calculated using formula (2.2)] is finite. The limiting integral in statement (A.6) is smaller than that in statement (A.5), which proves the least-favourable nature of the critical values calculated under the (null) sub-hypothesis* H_0^*.

Proof. From the definition of the point x_1 given by equation (2.2), we see that under the null H_0, the point x_1 is the largest x such that $F(x) = G(x)$. Note that the point x_1 is finite under the assumption of the theorem, because if it were infinite, then we would be under H_0^*, which is excluded. Hence, $x_1 < \infty$, and in this case we also have that $F(x) > G(x)$ at least for some x to the right of x_1. Furthermore, by Theorem A.2 we have that $x_{m,n} \to_{\mathbf{P}} x_1$.

With the above information, we proceed as follows:

$$\begin{aligned}\Theta_{m,n} &= \int_0^{x_{m,n}} \left\{\sqrt{\frac{nm}{n+m}}\Delta_{m,n}(z) + \sqrt{\frac{nm}{n+m}}\big(G(z)-F(z)\big)\right\}_+ dz \\ &= \mathbf{1}\{x_{m,n} \le x_1\} \int_0^{x_1} \left\{\sqrt{\frac{nm}{n+m}}\Delta_{m,n}(z)\right\}_+ dz \\ &\quad - \mathbf{1}\{x_{m,n} \le x_1\} \int_{x_{m,n}}^{x_1} \left\{\sqrt{\frac{nm}{n+m}}\Delta_{m,n}(z)\right\}_+ dz \\ &\quad + \mathbf{1}\{x_{m,n} > x_1\} \int_0^{x_1} \left\{\sqrt{\frac{nm}{n+m}}\Delta_{m,n}(z)\right\}_+ dz \\ &\quad + \mathbf{1}\{x_{m,n} > x_1\} \int_{x_1}^{x_{m,n}} \left\{\sqrt{\frac{nm}{n+m}}\Delta_{m,n}(z) + \sqrt{\frac{nm}{n+m}}\big(G(z)-F(z)\big)\right\}_+ dz.\end{aligned} \tag{A.7}$$

Hence, we have expressed $\Theta_{m,n}$ as a linear combination of four quantities. The second one converges to 0 because $x_{m,n} \to_{\mathbf{P}} x_1$ and

$$\sqrt{\frac{nm}{n+m}} \sup_{z\in[0,\infty)} \big|\Delta_{m,n}(z)\big| = O_{\mathbf{P}}(1), \tag{A.8}$$

by the classical Kolmogorov-Smirnov theorem. The fourth quantity converges to 0 because of the same reasons, plus the fact that $G(z) < F(z)$ for all $z \in (x_1, x_{m,n})$ and for all sufficiently large m and n so that $x_{m,n}$ would be sufficiently close to x_1 with as large a probability as desired. Hence,

$$\begin{aligned}\Theta_{m,n} &= \int_0^{x_1} \left\{\sqrt{\frac{nm}{n+m}}\Delta_{m,n}(z)\right\}_+ dz + o_{\mathbf{P}}(1) \\ &\to_d \int_0^{x_1} \Big(\sqrt{\eta}\,\mathcal{B}_1(F(z)) + \sqrt{1-\eta}\,\mathcal{B}_2(G(z))\Big)_+ dz,\end{aligned} \tag{A.9}$$

where $\mathcal{B}_1$ and $\mathcal{B}_2$ are two independent standard Brownian bridges. But we have $F(x) = G(x)$ for all $x \in [0, x_1)$. Hence, the processes $\{\sqrt{\eta}\,\mathcal{B}_1(F(z)) + \sqrt{1-\eta}\,\mathcal{B}_2(G(z)), z \in [0, x_1)\}$ and $\{\mathcal{B}(F(z)), z \in [0, x_1)\}$ coincide in distribution. This concludes the proof of Theorem A.4. ■

Theorem A.5 *Under the alternative* H_1*, we have* $\Theta_{m,n} \to_{\mathbf{P}} \infty$ *when* $\min\{m, n\} \to \infty$.

Proof. By Theorem A.2, we have that x_1 is finite and $x_{m,n} \to_{\mathbf{P}} x_1$. Next we write the bounds

$$\begin{aligned}
\Theta_{m,n} &= \int_0^{x_{m,n}} \left\{ \sqrt{\frac{nm}{n+m}} \Delta_{m,n}(z) + \sqrt{\frac{nm}{n+m}} \big(G(z) - F(z)\big) \right\}_+ dz \\
&\geq \int_0^{x_{m,n}} \left\{ \sqrt{\frac{nm}{n+m}} \big(G(z) - F(z)\big) \right\}_+ dz - \int_0^{x_{m,n}} \sqrt{\frac{nm}{n+m}} \big|\Delta_{m,n}(z)\big| dz \\
&\geq \int_0^{x_{m,n}} \left\{ \sqrt{\frac{nm}{n+m}} \big(G(z) - F(z)\big) \right\}_+ dz - x_{m,n} \sqrt{\frac{nm}{n+m}} \sup_{z \in [0,\infty)} \big|\Delta_{m,n}(z)\big| \qquad \text{(A.10)}
\end{aligned}$$

that hold for every pair (F, G). By statement (A.8), we are only left to verify that

$$\int_0^{x_{m,n}} \left\{ \sqrt{\frac{nm}{n+m}} \big(G(z) - F(z)\big) \right\}_+ dz \to_{\mathbf{P}} \infty. \qquad \text{(A.11)}$$

Recall that, under H_1, we have $F(x) \leq G(x)$ for all $x \in [0, x_1)$ with some $x^* \in (0, x_1)$ such that $F(x^*) < G(x^*)$. Hence, for a sufficiently small $\tau > 0$, we have that

$$\int_0^{x_1 - \tau} \big(G(z) - F(z)\big)_+ dz > 0. \qquad \text{(A.12)}$$

Since $x_{m,n} \to_{\mathbf{P}} x_1$, we can restrict ourselves to the case when $x_{m,n} > x_1 - \tau$. Hence, quantity (A.12) when multiplied by $\sqrt{nm/(n+m)}$ tends to ∞. Statement (A.11) follows. The proof of Theorem A.5 is finished. ■

Theorem A.6 *Suppose that X and Y have $2 + \kappa$ finite moments for some $\kappa > 0$, no matter how small. When $F \leq G$ but $F \neq G$ on $[0, \infty)$, then $\Theta_{m,n} \to_{\mathbf{P}} \infty$ when $\min\{m, n\} \to \infty$.*

Proof. We start with a bound:

$$\begin{aligned}
\Theta_{m,n} &= \int_0^{x_{m,n}} \left\{ \sqrt{\frac{nm}{n+m}} \Delta_{m,n}(z) + \sqrt{\frac{nm}{n+m}} \big(G(z) - F(z)\big) \right\}_+ dz \\
&\geq \int_0^{x_{m,n}} \left\{ \sqrt{\frac{nm}{n+m}} \big(G(z) - F(z)\big) \right\}_+ dz - \int_0^{\infty} \sqrt{\frac{nm}{n+m}} \big|\Delta_{m,n}(z)\big| dz, \qquad \text{(A.13)}
\end{aligned}$$

which holds for every pair (F, G). Due to the assumption that X and Y have $2 + \kappa$ finite moments for some $\kappa > 0$, we have that

$$\sqrt{\frac{nm}{n+m}} \int_0^{\infty} \big|\Delta_{m,n}(z)\big| dz = O_{\mathbf{P}}(1). \qquad \text{(A.14)}$$

Since $x_{m,n} \to_{\mathbf{P}} \infty$, the first integral on the right-hand side of bound (A.13) tends to ∞ in probability provided that, for a sufficiently large $M < \infty$,

$$\int_0^{M} \left\{ \sqrt{\frac{nm}{n+m}} \big(G(z) - F(z)\big) \right\}_+ dz \to \infty. \qquad \text{(A.15)}$$

This is true because $F \leq G$ but not $F = G$ on $[0, \infty)$, and so the right-continuity of the cdf's implies that $\int_0^M (G(z) - F(z))_+ dz > 0$ for all sufficiently large $M < \infty$. This concludes the proof of Theorem A.6. ■

B Appendix: Detailed simulation results

Table B.1: Estimated bias and variance of crossing point estimators based on $\delta_{m,n} = \kappa((m+n)/(mn))^{1/2-\epsilon}$ with $n = m$ and $\epsilon = 0.1$. Estimates correspond to different choices of the constant κ and are computed using 5,000 Monte Carlo replications with $B = 1,000$ bootstrap samples.

	$n = 100$	$n = 500$	$n = 1,000$	$n = 5,000$
	$\kappa = 10^{-2}$			
Bias	0.05	0.12	0.08	0.02
Variance	6.21	0.87	0.39	0.07
	$\kappa = 10^{-3}$			
Bias	-0.30	-0.09	-0.09	-0.08
Variance	5.01	0.96	0.51	0.08
	$\kappa = 10^{-4}$			
Bias	-0.41	-0.14	-0.14	-0.11
Variance	5.19	0.95	0.46	0.09
	$\kappa = 10^{-5}$			
Bias	-0.39	-0.19	-0.14	-0.13
Variance	5.05	0.94	0.47	0.10
	$\kappa = 10^{-6}$			
Bias	-0.43	-0.18	-0.15	-0.13
Variance	5.14	0.96	0.44	0.10

Table B.2: Estimated bias and variance of crossing point estimators based on $\delta_{m,n} = \kappa((m+n)/(mn))^{1/2-\epsilon}$ with $n = m$ and $\epsilon = 0.01$. Estimates correspond to different choices of the constant κ and are computed from 5,000 Monte Carlo replications with $B = 1,000$ bootstrap samples.

	$n = 100$	$n = 500$	$n = 1,000$	$n = 5,000$
		$\kappa = 10^{-2}$		
Bias	-0.02	0.15	0.07	0.01
Variance	5.66	0.91	0.37	0.06
		$\kappa = 10^{-3}$		
Bias	-0.38	-0.09	-0.07	-0.07
Variance	4.54	0.99	0.42	0.07
		$\kappa = 10^{-4}$		
Bias	-0.38	-0.16	-0.14	-0.11
Variance	4.98	1.03	0.46	0.09
		$\kappa = 10^{-5}$		
Bias	-0.36	-0.22	-0.16	-0.12
Variance	4.78	0.94	0.50	0.10
		$\kappa = 10^{-6}$		
Bias	-0.39	-0.19	-0.15	-0.12
Variance	5.16	0.91	0.45	0.10

Table B.3: Estimated size of dominance tests at the 5% nominal level when $F = G = LN(0.85, 0.6)$ based on $\delta_{m,n} = \kappa((m+n)/(mn))^{1/2-\epsilon}$ with $n = m$. Empirical rejection probabilities are reported for various combinations of κ and ϵ, and are computed from 5,000 Monte Carlo replications with $B = 1,000$ bootstrap samples.

κ	$n = 100$	$n = 500$	$n = 1,000$	$n = 5,000$
		$\epsilon = 0.1$		
10^{-2}	0.041	0.042	0.047	0.052
10^{-3}	0.044	0.041	0.044	0.049
10^{-4}	0.040	0.043	0.043	0.041
10^{-5}	0.042	0.039	0.042	0.042
10^{-6}	0.039	0.041	0.043	0.041
		$\epsilon = 0.01$		
10^{-2}	0.043	0.042	0.043	0.050
10^{-3}	0.042	0.044	0.043	0.043
10^{-4}	0.043	0.042	0.042	0.043
10^{-5}	0.042	0.040	0.042	0.041
10^{-6}	0.041	0.041	0.042	0.042

Table B.4: Estimated power of dominance tests at the 5% nominal level against $F = LN(0.85, 0.6)$ and $G = LN(0.7, 0.8)$ based on $\delta_{m,n} = \kappa((m+n)/(mn))^{1/2-\epsilon}$ with $n = m$. Empirical rejection probabilities are reported for various combinations of κ and ϵ, and are computed from 5,000 Monte Carlo replications with $B = 1,000$ bootstrap samples.

κ	$n = 100$	$n = 500$	$n = 1,000$	$n = 5,000$
		$\epsilon = 0.1$		
10^{-2}	0.398	0.978	1.000	1.000
10^{-3}	0.603	0.979	0.993	1.000
10^{-4}	0.634	0.981	0.988	0.998
10^{-5}	0.664	0.981	0.989	0.997
10^{-6}	0.658	0.979	0.989	0.997
		$\epsilon = 0.01$		
10^{-2}	0.397	0.984	1.000	1.000
10^{-3}	0.603	0.979	0.993	1.000
10^{-4}	0.657	0.979	0.989	0.998
10^{-5}	0.661	0.977	0.987	0.997
10^{-6}	0.662	0.980	0.989	0.997