

Consistent Nonparametric Tests for Bivariate Stochastic Dominance with Applications to Labor Market Econometrics

Luciano A. Perez ¹

¹Instituto de Economía (INECO)
Universidad Argentina de la Empresa (UADE)
Lima 775, C1073 Buenos Aires, Argentina
luperez@uade.edu.ar

ABSTRACT

In this paper, we develop fully consistent nonparametric tests for bivariate stochastic dominance, generalizing the weak convergence approach. We overcome known problems with multivariate Kolmogorov-Smirnov statistics using empirical processes theory and bootstrap techniques to obtain fully consistent non parametric tests for bivariate 1st and 2nd order stochastic dominance over several modularity classes of test functions. We introduce theorems showing the weak convergence of our test statistics to appropriate transformations of Brownian bridge processes and use this to obtain data depending p-values. We also make a bootstrap power assessment of the tests, showing them to be in line with top standards in the statistical literature. As we don't assume absolute continuity our approach is useful even for discrete distributions. We finally apply our tests to construct bivariate economic welfare rankings for Argentina in the aftermath of the 2001 crisis.

Keywords: Stochastic dominance, multidimensional inequality, empirical processes, bootstrap, Kolmogorov-Smirnov statistics.

2010 Mathematics Subject Classification: 62G30, 62F40, 62P20.

JEL classification: D81, C01, D63.

1 Introduction

The econometric assessment of the effects of active labor market policies on aggregate social welfare usually focuses on its effects on income distribution. The standard approach to multidimensional distributional analysis has been mostly reduced to the use of generalized inequality indexes that combine several features (e.g. income, education or labor status) into a one dimensional measure whose concentration is then analyzed using the usual tools for income distribution, like the Lorenz curve and its associated indexes (Tsui, 1995).

This puts some limitation on the statistical approach for multidimensional inequality analysis since the asymptotic distribution of such indexes is not easily derived and consistent tests for multidimensional inequality are then scarce (Brufman and Perez, 2011). Another limitation

is the inexistence of a multidimensional analogue to the Lorenz curve or even a generally accepted definition for the quantiles of a multivariate distribution (Perez, 2015).

Barret and Donald 2003 introduced fully consistent tests for one dimensional inequality based on the concept of stochastic dominance, comparing distributions in a point-to-point basis rather than using concentration measures. This tests are based on the Kolmogorov-Smirnov (KS) type tests introduced by McFadden 1989 and as a consequence are not directly generalizable to the multidimensional case because multivariate KS statistics are not pivotal (Simpson, 1951).

Recent works on stochastic dominance, as those surveyed in Whang 2019, focus mainly on univariate distributions, and when dealing with multivariate stochastic dominance, they avoid the statistical approach that is needed to apply the tests of Atkinson and Bourguignon 1982 to household survey data.

In this paper we develop the statistical approach necessary to test bivariate stochastic dominance of 1st and 2nd order, thus extending both Atkinson and Bourguignon 1982 and Barret and Donald 2003 . Our approach relies on empirical processes theory and allows us to use bootstrapping techniques to obtain critical values dependent on data (as in van der Vaart W. and Wellner 2000).

In general terms, the statistical approach to multidimensional economic inequality, and multivariate stochastic dominance in particular, has been unsatisfactory, for the reasons outlined above and detailed in Perez 2015. In this paper, we intend to fill the gap by providing powerful multivariate tests, well founded in mathematical theory, but also useful in practical terms and with straightforward computational implementation.

In Section 2 we introduce the concept of stochastic dominance as developed in mathematical risk theory. In Section 3 we introduce the multidimensional modularity classes of test functions, and present the sufficient conditions for 1st and 2nd order dominance in the form of Perez (2018), that generalizes both Atkinson and Bourguignon 1982 and Hadar and Russell 1974 discarding the assumption of absolute continuity of the underlying distribution. This is important as it allows us to apply our tests directly on discrete distributions, without the need of an intermediate KDE estimation for the distribution that would imply potential loss in consistency. In Section 4 we develop the mathematical basis for our tests in the form of theorems proving the weak convergence of certain classes of tests statistics of the KS type. In Section 5 we outline the practical implementation of our approach and show its application to the construction of welfare rankings for bivariate income-human capital distributions in Argentina. In Section 6 we asses some statistical properties of our tests, like power, comparing to alternatives. Section 7 presents concluding remarks and introduce some potential further developments.

2 Stochastic Dominance

In general terms, stochastic dominance is a mathematical rule for ordering random prospects based on their expected values over certain classes of test functions. For example, given random variables X and Y with cdfs F_X and F_Y respectively, with common compact support $\bar{U}$, and a class Φ of real valued functions, we say there is stochastic dominance of X over Y

for the class Φ , if for every $\phi \in \Phi$ we have:

$$\int_{\mathbb{R}} \phi(t) dF_X \geq \int_{\mathbb{R}} \phi(t) dF_Y \quad (2.1)$$

This can be interpreted as X having a higher expected von Neumann and Morgenstern utility than Y , and thus being a dominant lottery. Interpretation as point dominance of associated quantile couplings can be achieved with the help of Strassen theorem but only in the one dimensional case, as shown in Chapter 3 of Perez 2015.

The most important classes of functions in the univariate case are the class Φ_1 of $\mathcal{C}^1$ increasing functions and Φ_2 of $\mathcal{C}^2$ increasing and concave functions. If X dominates Y over Φ_1 we say that there is first order stochastic dominance, and we denote it by XSD_1Y . If X dominates Y over Φ_2 we say that there is second order stochastic dominance, denoted XSD_2Y .

Hadar and Russell 1969 provided sufficient conditions for first and second order dominance. If we assume (without loss of generality) that F_X and F_Y have support $\bar{U} = [0, 1]$ we have:

$$F_X(s) \leq F_Y(s) \quad \forall s \in [0, 1] \Rightarrow XSD_1Y \quad (2.2)$$

$$\int_0^t F_X(s) ds \leq \int_0^t F_Y(s) ds \quad \forall t \in [0, 1] \Rightarrow XSD_2Y \quad (2.3)$$

As conditions 2.2 and 2.3 are obviously also necessary, we have equivalence of those conditions and stochastic dominance. The importance of this conditions lies on the fact that they involve a single inequality for cdfs (or its integrals) as opposed to the definition of stochastic dominance which involves an infinite set of inequalities (one for each test function ϕ).

This in turn allows for an easy definition of univariate stochastic dominance of arbitrary order using the following sequence of integral operators:

$$\mathcal{T}_1(z, F_X) = F_X(z) \quad \mathcal{T}_{j+1}(z, F_X) = \int_0^z \mathcal{T}_j(t, F_X) dt \quad (2.4)$$

We say that there is j -th order stochastic dominance of X over Y if $\mathcal{T}_j(z, F_X) \leq \mathcal{T}_j(z, F_Y)$.

Additionally, and perhaps more important in practice, conditions 2.2 and 2.3 involve only cdfs as inputs, which makes easier to test them statistically.

The idea of using Kolmogorov-Smirnov-type statistics to test for univariate stochastic dominance of X over Y comes from (McFadden, 1989) who stated the null hypothesis $H_0 : F_X(z) \leq F_Y(z) \quad \forall z \in [0, 1]$ against alternative $H_1 : F_X(z) > F_Y(z)$ for some $z \in [0, 1]$.

As dominance condition $F_X(z) \leq F_Y(z) \quad \forall z \in [0, 1]$ is equivalent to $\sup_{z \in [0, 1]} (F_X - F_Y)(z) \leq 0$ it is natural to search for an empirical or *plug-in* analogue of this inequality.

Using same size samples $(x_1, \dots, x_n)$ and $(y_1, \dots, y_n)$ to construct empirical cdfs F_X^n and F_Y^n we can define $D_n(z) = \sqrt{n}(F_X^n(z) - F_Y^n(z))$ and then a two sample test statistic for H_0 is:

$$D_n^* = \max_{w \in [0, 1]} D_n(w) \quad (2.5)$$

The only information we need then to carry on testing XSD_1Y in the univariate case is the asymptotic distribution of test statistic D_n^* , which (McFadden, 1989) shows to be of Smirnov type (assuming continuity for F_X and F_Y). The *pivotality* of Kolmogorov-Smirnov statistics

such as 2.5, that is, its independence on the underlying distributions, is the key factor here, as those distributions are generally unknown.

As proven by Simpson 1951 this invariance property for the asymptotic distribution of Kolmogorov-Smirnov statistics is lost in the two dimensional case. On the other hand Schmid 1958 shows how discontinuity accounts for loss of pivotality even for the univariate case. To work around these difficulties we'll have to use empirical processes theory and bootstrap techniques in Sections 4.2 and 4.3.

For second order univariate stochastic dominance McFadden 1989 proposes the test statistic:

$$S_n^* = \max_{w \in [0,1]} S_n(w) \quad (2.6)$$

where $S_n(w) = \sqrt{n} \int_0^w (F_Y^n(z) - F_X^n(z)) dz$, but didn't completely characterize its asymptotic distribution.

Our goal in Sections 4.2 and 4.3 is to develop bivariate first and second order stochastic dominance tests for distributions with compact support¹ in some sense extending Barret and Donald 2003 univariate tests, but using empirical bootstrap theory as in Section 3.7 of van der Vaart W. and Wellner 2000.

3 Modularity Classes and Bivariate Stochastic Dominance

Given a class $\Phi = \{\phi\}$ of continuous functions $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ we say that there is stochastic dominance of random vector (X_1, Y_1) over (X_2, Y_2) if for all $\phi \in \Phi$ we have $E(\phi(X_1, Y_1)) \geq E(\phi(X_2, Y_2))$.

If we assume that (X_1, Y_1) and (X_2, Y_2) have a cdfs F_1 and F_2 respectively, with common compact support U , we can put dominance in terms of integrals:

$$\int_U \phi(s, t) dF_1 \geq \int_U \phi(s, t) dF_2 \quad \forall \phi \in \Phi \quad (3.1)$$

As in the univariate case, to obtain sufficient conditions for dominance in terms of inequalities involving only cdfs and it's integrals, we have to restrict our attention to some special classes of test functions.

Following Atkinson and Bourguignon 1982 and Perez 2018 we define first order modularity classes².

Definition 3.1. Let $\Omega \subseteq \mathbb{R}^2$ be an open connected set and let $\phi : \mathbb{R}^2 \rightarrow \mathbb{R} \in \mathcal{C}^2(\Omega)$ be an increasing function on each of its arguments. We say that:

1. $\phi \in \Phi^+$ if $\forall (x, y) \in \Omega$ we have $\frac{\partial^2 \phi}{\partial x \partial y}(x, y) \geq 0$.
2. $\phi \in \Phi^-$ if $\forall (x, y) \in \Omega$ we have $\frac{\partial^2 \phi}{\partial x \partial y}(x, y) \leq 0$.

¹Compact support for underlying distributions is not essential in first and second order dominance. For higher dominance orders, however, integral operators as in 2.4 are in general unbounded and need to be dealt with using additional tools such as weighted Kolmogorov-Smirnov statistics (Horváth, Kokoszka and Zitikis, 2006).

²Modularity conditions can be stated more generally without differentiability assumptions and conditions 3.1 shown to be equivalent for the case of $\mathcal{C}^2$ functions. See Topkis 1998 or Perez 2015.

In the first case we say that ϕ is *supermodular* while in the second case we call it *submodular*.

In the context of microeconomics, if we think of ϕ as a utility function, the condition of supermodularity accounts for *strategic complementarity* of input goods (Topkis, 1998). This is what makes bivariate stochastic dominance a useful tool in welfare economics as shown in Section 5 of this paper.

For any cdf F with support $\bar{U} = [0, 1] \times [0, 1]$, we compute its marginal distributions $F^X(x) = F(x, 1)$ and $F^Y(y) = F(1, y)$. Following the notation of Atkinson and Bourguignon 1982 we then define the associated function $K(x, y) = -(F(x, y) - F^X(x) - F^Y(y))$.

Given random vectors (X_1, Y_1) and (X_2, Y_2) with cdfs F_1 and F_2 and $(s, t) \in \bar{U}$ we denote $\Delta F(s, t) = F_1(s, t) - F_2(s, t)$, $\Delta F^X(s) = F_1^X(s) - F_2^X(s)$, $\Delta F^Y(t) = F_1^Y(t) - F_2^Y(t)$ and $\Delta K(s, t) = K_1(s, t) - K_2(s, t)$.

Hadar and Russell 1974 provided sufficient conditions for stochastic dominance of vector (X_1, Y_1) over (X_2, Y_2) over modularity classes Φ^+ and Φ^- assuming absolutely continuous distributions. Perez 2018 proved that those conditions are sufficient for the general case where underlying distributions are only assumed to have compact support.

This conditions are analogue to univariate sufficient conditions of Hadar and Russell 1969 in the sense that they involve just cdfs F_1 and F_2 and its integrals. For first order bivariate stochastic dominance we can state them as in Perez 2015:

$$\Delta F(s, t) \leq 0 \quad \forall (s, t) \in \bar{U} \Rightarrow F_1 SD_1 F_2 \text{ over } \Phi^- \quad (3.2)$$

$$\Delta K(s, t) \leq 0, \Delta F^X(s), \Delta F^Y(t) \leq 0 \quad \forall (s, t) \in \bar{U} \Rightarrow F_1 SD_1 F_2 \text{ over } \Phi^+ \quad (3.3)$$

Note that the specific marginal conditions are only relevant in the supermodular case, as they are trivial in the submodular case.

To obtain sufficient conditions for bivariate second order dominance we first have to set our higher order modularity classes. We follow Perez 2018 and define:

Definition 3.2. Let $\Omega \subseteq \mathbb{R}^2$ be an open connected set and let $\phi : \mathbb{R}^2 \rightarrow \mathbb{R} \in \mathcal{C}^2(\Omega)$ be an increasing function on each of its arguments. We say that:

1. $\phi \in \Phi^{--}$ if $\phi \in \Phi^-$ and $\forall (x, y) \in \Omega$ we have:

$$\frac{\partial^2 \phi}{\partial x^2}(x, y), \frac{\partial^2 \phi}{\partial y^2}(x, y) \leq 0 \quad (3.4)$$

$$\frac{\partial^3 \phi}{\partial x^2 \partial y}(x, y), \frac{\partial^3 \phi}{\partial x \partial y^2}(x, y) \geq 0 \quad (3.5)$$

$$\frac{\partial^4 \phi}{\partial x^2 \partial y^2}(x, y) \leq 0 \quad (3.6)$$

2. $\phi \in \Phi^{++}$ if $\phi \in \Phi^+$ and $\forall (x, y) \in \Omega$ we have:

$$\frac{\partial^2 \phi}{\partial x^2}(x, y), \frac{\partial^2 \phi}{\partial y^2}(x, y) \leq 0 \quad (3.7)$$

$$\frac{\partial^3 \phi}{\partial x^2 \partial y}(x, y), \frac{\partial^3 \phi}{\partial x \partial y^2}(x, y) \leq 0 \quad (3.8)$$

$$\frac{\partial^4 \phi}{\partial x^2 \partial y^2}(x, y) \geq 0 \quad (3.9)$$

To establish again sufficient conditions involving a single inequality we have to define integral functionals:

$$H(x, y, F) = \int_0^x \int_0^y F(s, t) dt ds \quad (3.10)$$

$$L(x, y, F) = \int_0^x \int_0^y K(s, t) dt ds \quad (3.11)$$

Those functionals are defined for any cdf F with support $\bar{U}$ and in particular using subindexes 1 and 2 we denote its analogue for F_1 and F_2 .

We are also going to need marginal functionals defined as:

$$H^X(x, F) = \int_0^x F^X(s) ds \quad (3.12)$$

$$H^Y(y, F) = \int_0^y F^Y(t) dt \quad (3.13)$$

We set difference operators as before: $\Delta H(s, t) = H_1(s, t) - H_2(s, t)$, $\Delta H^X(s, t) = H_1^X(s, t) - H_2^X(s, t)$, $\Delta H^Y(s, t) = H_1^Y(s, t) - H_2^Y(s, t)$ and $\Delta L(s, t) = L_1(s, t) - L_2(s, t)$. The following conditions were proven to be sufficient by Hadar and Russell 1974 for the absolutely continuous case and by Perez 2018 for the general compact support case:

$$\Delta H \leq 0, \Delta H^X \leq 0, \Delta H^Y \leq 0 \quad \forall (s, t) \in \bar{U} \Rightarrow F_1 SD_2 F_2 \text{ over } \Phi^{--} \quad (3.14)$$

$$\Delta L \leq 0, \Delta H^X \leq 0, \Delta H^Y \leq 0 \quad \forall (s, t) \in \bar{U} \Rightarrow F_1 SD_2 F_2 \text{ over } \Phi^{++} \quad (3.15)$$

Note that marginal conditions are the same in both cases. Our goal in the next sections is to transform this sufficient conditions for dominance into statistical tests, using appropriate empirical versions of the functionals and Kolmogorov-Smirnov-type test statistics.

4 Testing for Bivariate Stochastic Dominance

4.1 General Setting

To test for first order stochastic dominance we first extend the definition of empirical functionals from Davidson and Duclos 2000 to the bivariate case. Sampling from the random vector (X, Y) we obtain bivariate random samples $\{(X_1, Y_1), \dots, (X_n, Y_n)\}$ and we can construct the associated empirical measure $P_n(B) = \frac{1}{n} \sum_{i=1}^n \delta_{(X_i, Y_i)}(B)$ for each Borel set $B \subseteq \mathbb{R}^2$.

As vectors (X_i, Y_i) are random, this is a random measure³.

The empirical distribution function is obtained as:

$$\hat{F}_n(s, t) = P_n((-\infty, s] \times (-\infty, t]) = \frac{1}{n} \sum_{i=1}^n 1_{\{X_i \leq s\}}(s) 1_{\{Y_i \leq t\}}(t) \quad (4.1)$$

³On a technical note, we are going to assume that the sample space is a separable metric space and that continuum hypothesis holds, to ensure that the range of the Brownian bridge is separable (see (Dudley, 2003), Sec.13.1).

Marginal empirical cdfs are thus:

$$\hat{F}_n^X(s) = P_n((-\infty, s] \times \mathbb{R}) = \frac{1}{n} \sum_{i=1}^n 1_{\{X_i \leq s\}}(s) \quad (4.2)$$

$$\hat{F}_n^Y(t) = P_n(\mathbb{R} \times (-\infty, t]) = \frac{1}{n} \sum_{i=1}^n 1_{\{Y_i \leq t\}}(t) \quad (4.3)$$

We also have an empirical version of $K(s, t)$ given by:

$$\hat{K}(s, t) = - \left(\hat{F}(s, t) - \hat{F}^X(s) - \hat{F}^Y(t) \right) \quad (4.4)$$

Now assume we have samples $\{(X_1^1, Y_1^1), (X_2^1, Y_2^1), \dots, (X_m^1, Y_m^1)\}$ from distribution F_1 (with measure P) and $\{(X_1^2, Y_1^2), (X_2^2, Y_2^2), \dots, (X_m^2, Y_m^2)\}$ from distribution F_2 (measure Q), with associated empirical measures P_m and Q_n respectively. The general two sample statistic is then $\nu_{m,n} = \left(\frac{mn}{m+n} \right)^{1/2} (P_m - Q_n)$.

Following the same logic as in McFadden 1989 we are going to apply $\nu_{m,n}$ to sets of the form $[0, s] \times [0, t]$, and then take supremum to obtain a uniform condition of Kolmogorov-Smirnov type.

We then have the test statistic for first order stochastic dominance in the submodular bivariate case:

$$\lambda_{m,n} = \left(\frac{mn}{m+n} \right)^{1/2} \sup_{(s,t) \in [0,1] \times [0,1]} (\hat{F}_{1m} - \hat{F}_{2n})(s, t) \quad (4.5)$$

where $\hat{F}_{1m}$ and $\hat{F}_{2n}$ are the cdfs obtained from empirical measures P_m and Q_n .

For the supermodular class we introduce the analogue test statistic⁴:

$$\kappa_{m,n} = \left(\frac{mn}{m+n} \right)^{1/2} \sup_{[0,1] \times [0,1]} (\hat{K}_{1m} - \hat{K}_{2n})(s, t) \quad (4.6)$$

To obtain limit distributions for test statistics 4.5 and 4.6 we are going to use bootstrap techniques as in (van der Vaart W. and Wellner, 2000), Sec.3.7. To simplify notation let's put $V_i = (X_i^1, Y_i^1)$, the i -th element in the sample from F_1 , and $W_j = (X_j^2, Y_j^2)$, the j -th element in the sample from F_2 .

We start with the *pooled sample* of size $N = n + m$:

$$\{Z_{N1}, Z_{N2}, \dots, Z_{NN}\} = \{V_1, \dots, V_m, W_1, \dots, W_n\} \quad (4.7)$$

From the pooled sample we can define the *pooled empirical distribution*:

$$H_N = \frac{1}{N} \sum_{i=1}^N \delta_{Z_{Ni}} = \alpha_N P_m + (1 - \alpha_N) Q_n \quad (4.8)$$

where $\alpha_N = m/n$.

⁴Note that we are not centering our attention on the marginal conditions for dominance as they can be tested with the already known univariate test by Barret and Donald 2003.

Now if we re-sample from 4.7 with replacement we obtain bootstrap samples of size N and for each of this samples we have bootstrapped empirical measures:

$$\hat{P}_{m,N}^B = \frac{1}{m} \sum_{i=1}^m \delta_{\hat{Z}_{Ni}} \quad \hat{Q}_{n,N}^B = \frac{1}{n} \sum_{i=1}^n \delta_{\hat{Z}_{N,m+i}} \quad (4.9)$$

We are going to use a couple of results from (van der Vaart W. and Wellner, 2000) but we first need to introduce some notation. Given a signed measure Q and a class of measurable functions $\mathcal{F}$ we think of Q acting on functions by integration, i.e. $Qf = \int f dQ$. We denote $\|Q\|_{\mathcal{F}} = \sup\{|Qf| : f \in \mathcal{F}\}$. We denote weak convergence of a sequence of probability measures by $\rightsquigarrow$.

For the definition of Donsker classes and the Brownian bridge we refer to van der Vaart W. and Wellner 2000 Sec.2.1, and for outer integral P^*f to Sec.1.2 in the same reference.

In the proofs of our main results in Sections 4.2 and 4.3, we'll use the following results from Sec.3.7. of van der Vaart W. and Wellner 2000:

Theorem 4.1. *Let $\mathcal{F}$ be a class of measurable functions which is both P -Donsker and Q -Donsker, and such that $\|P\|_{\mathcal{F}} < \infty$, $\|Q\|_{\mathcal{F}} < \infty$. If $m, n \rightarrow \infty$ such that $m/N \rightarrow \alpha \in (0, 1)$, then $\sqrt{m}(\hat{P}_{m,N}^B - H_N) \rightsquigarrow G_H$ in probability given $V_1, V_2, \dots, W_1, W_2, \dots$*

Here G_H is a tight Brownian bridge corresponding to measure $H = \alpha P + (1 - \alpha)Q$.

Proof. It is Theorem 3.7.6 from van der Vaart W. and Wellner 2000.⁵ □

Theorem 4.2. *If in addition $\mathcal{F}$ possesses an envelope function F with $P^*F^2 < \infty$ and $Q^*F^2 < \infty$, then $\sqrt{m}(\hat{P}_{m,N}^B - H_N) \rightsquigarrow G_H$ given almost any sequence $V_1, V_2, \dots, W_1, W_2, \dots$*

Proof. It is Theorem 3.7.6 from van der Vaart W. and Wellner 2000. □

4.2 Nonparametric Tests for First Order Bivariate Stochastic Dominance

In what follows we are going to consider the class of indicator functions:

$$\mathcal{F} = \{1_{[0,s] \times [0,t]} : s, t \in [0, 1]\}$$

and the space $l^\infty(\mathcal{F})$ of all real bounded functions over $\mathcal{F}$. A typical element from $l^\infty(\mathcal{F})$ is for example a finite signed measure, acting over the elements of $\mathcal{F}$ by integration.

It will be useful to assign to each function $f \in \mathcal{F}$ a couple of related functions also in $\mathcal{F}$, given by:

$$f = 1_{[0,s] \times [0,t]} \longrightarrow \begin{cases} f^x = 1_{[0,s] \times [0,1]} \\ f^y = 1_{[0,1] \times [0,t]} \end{cases}$$

Let's consider now the following operators, $T_1, T_2 : l^\infty(\mathcal{F}) \rightarrow \mathbb{R}$, given by:

$$T_1(P) = \sup_{f \in \mathcal{F}} Pf \quad (4.10)$$

⁵Theorem 3.7.6 and Theorem 3.7.7 are for general sample spaces, we are going to use them for the special case of random vectors in 2 dimensions.

$$T_2(P) = \sup_{f \in \mathcal{F}} (Pf^x + Pf^y - Pf) \quad (4.11)$$

Note that if P and Q are probability measures on $[0, 1] \times [0, 1]$, then they have associated cdfs $F_1(s, t)$ and $F_2(s, t)$ (and functions $K_1(s, t), K_2(s, t)$ given as in Sec.3) and then:

$$T_1(P - Q) = \sup_{(s,t) \in [0,1] \times [0,1]} (F_1 - F_2)(s, t) \quad (4.12)$$

$$T_2(P - Q) = \sup_{(s,t) \in [0,1] \times [0,1]} (K_1 - K_2)(s, t) \quad (4.13)$$

Theorem 4.3 (First Order Bivariate Stochastic Dominance Test). *Let's consider test statistics $\lambda_{m,n}^B$ and $\kappa_{m,n}^B$ given by using two sample bootstrap empirical distributions (given by 4.9) in formulas 4.5 and 4.6. Then if $m, n \rightarrow \infty$ such that $m/N \rightarrow \alpha \in (0, 1)$:*

1. $\lambda_{m,n}^B \rightsquigarrow T_1(G_H)$
2. $\kappa_{m,n}^B \rightsquigarrow T_2(G_H)$

given almost any sequence $V_1, V_2, \dots, W_1, W_2, \dots$, where $H = \alpha P + (1 - \alpha)Q$ and G_H is a version of the tight Brownian bridge corresponding to measure H .

Proof. Our class of functions $\mathcal{F}$ is a Donsker class because it is a subclass of rectangles in $\mathbb{R}^2$ which is known to be Donsker (Dudley, 2003; van der Vaart W. and Wellner, 2000). On the other hand, constant function 1 is an envelope satisfying conditions in Theorem 4.2. Consequently we have:

$$\sqrt{\frac{mn}{m+n}}(\hat{P}_{m,N} - \hat{Q}_{n,N}) \rightsquigarrow G_H$$

Now we have $\lambda_{m,n}^B = T_1(\sqrt{\frac{mn}{m+n}}(\hat{P}_{m,N} - \hat{Q}_{n,N}))$ and $\kappa_{m,n}^B = T_2(\sqrt{\frac{mn}{m+n}}(\hat{P}_{m,N} - \hat{Q}_{n,N}))$. Note that empirical measures (and its bootstrap versions) are random elements taking values in the metric space $(l^\infty(\mathcal{F}), |||_{\mathcal{F}})$.

Under our assumptions we know that the range of the Brownian bridge is separable and then, as weak convergence is preserved by continuous mappings (van der Vaart W. and Wellner 2000, Theorem 1.11.1) our result will follow if we show that operators T_1 and T_2 are continuous. Now if P and Q are elements from $l^\infty(\mathcal{F})$:

$$|T_1(P) - T_1(Q)| = \left| \sup_{f \in \mathcal{F}} Pf - \sup_{f \in \mathcal{F}} Qf \right| \leq \sup_{f \in \mathcal{F}} |(P - Q)f| = \|P - Q\|_{\mathcal{F}}$$

this shows that T_1 is continuous and proves the first part.

Now let P and Q be elements from $l^\infty(\mathcal{F})$, then:

$$\begin{aligned} |T_2(P) - T_2(Q)| &= \left| \sup_{f \in \mathcal{F}} (Pf^x + Pf^y - Pf) - \sup_{f \in \mathcal{F}} (Qf^x + Qf^y - Qf) \right| \\ &\leq \sup_{f \in \mathcal{F}} \{ |(Pf^x + Pf^y - Pf) - (Qf^x + Qf^y - Qf)| \} \leq \\ &\leq \sup_{f \in \mathcal{F}} \{ |(P - Q)f^x| + |(P - Q)f^y| + |(Q - P)f| \} \end{aligned}$$

It is clear that $\sup_{f \in \mathcal{F}} \{|(P - Q)f^x|\} \leq \|P - Q\|_{\mathcal{F}}$, as we are covering a smaller class of functions. The same applies to the term with $\{|(P - Q)f^y|\}$ and then we have:

$$|T_2(P) - T_2(Q)| \leq 3 \|P - Q\|_{\mathcal{F}}$$

which shows T_2 is continuous and this finishes our proof. $\square$

As a consequence of Theorem 4.3, we have a sequence of critical values given by the quantiles of bootstrap distribution, that is:

$$\hat{c}_{mn}^- = \inf \{t : P_{\hat{Z}}(\lambda_{mn}^B > t) \leq \beta\} \quad (4.14)$$

for the submodular case and:

$$\hat{c}_{mn}^+ = \inf \{t : P_{\hat{Z}}(\kappa_{mn}^B > t) \leq \beta\} \quad (4.15)$$

for the supermodular case.

The consistency of the test then follows in the sense that:

$$\hat{c}_{mn}^- \longrightarrow c_H^- = \inf \{t : P(T_1(G_H) > t) \leq \beta\} \quad (4.16)$$

$$\hat{c}_{mn}^+ \longrightarrow c_H^+ = \inf \{t : P(T_2(G_H) > t) \leq \beta\} \quad (4.17)$$

In the supermodular case for example, we have a one-tail test, rejecting hypothesis PSD_1Q (equivalent to $K_1 \leq K_2$ uniformly) if $\kappa_{m,n} > \hat{c}_{mn}^+$. Given the pooled sample $V_1, \dots, V_m, W_1, \dots, W_n$, there are N^N bootstrap samples, each with the same probability then N^N possible values for κ_{mn}^B (counting with multiplicity).

Note that rejecting PSD_1Q does not imply accepting QSD_1P , simply because the fact that $\sup_{s,t} (F_1(s, t) - F_2(s, t)) \not\leq 0$ does not imply $\sup_{s,t} (F_2(s, t) - F_1(s, t)) \leq 0$. That is why we don't have a two tailed test, and instead we have to test hypothesis PSD_1Q and QSD_1P separately. This is similar for the tests in Barret and Donald 2003.

Our test rejects the hypothesis if the value of κ_{mn} is on the unpper β fraction of the values κ_{mn}^B , resulting then in a probability of type I error less than β . The sequence of tests then results asymptotically of level β .

4.3 Testing Second Order Bivariate Stochastic Dominance

To test second order dominance we first need an empirical version of the integral operator H :

$$\hat{H}(x, y) = H(x, y, \hat{F}_N) = \int_0^x \int_0^y \hat{F}_N(s, t) ds dt \quad (4.18)$$

If we note that:

$$\begin{aligned} L(x, y) &= \int_0^x \int_0^y K(s, t) dt = \int_0^x \int_0^y (F^X(s) + F^Y(t) - F(s, t)) ds dt = \\ &= -H(x, y) + yF^X(x) + xF^Y(y) \end{aligned} \quad (4.19)$$

we can simply use $\hat{H}$, $\hat{F}^X$ and $\hat{F}^Y$ to obtain $\hat{L}$.

Using the same two samples as in Section 4.2 we define for the submodular case:

$$\mu_{m,n} = \left(\frac{mn}{m+n} \right)^{1/2} \sup_{[0,1] \times [0,1]} (\hat{H}_{1m} - \hat{H}_{2n})(s, t) \quad (4.20)$$

and similarly for the supermodular case:

$$\gamma_{m,n} = \left(\frac{mn}{m+n} \right)^{1/2} \sup_{[0,1] \times [0,1]} (\hat{L}_{1m} - \hat{L}_{2n})(s, t) \quad (4.21)$$

Note that this test statistics are multidimensional generalizations of those in (Barret and Donald, 2003).

Working with the same definitions and references as in Sec.4.2 we'll define two operators, $T_3, T_4 : l^\infty(\mathcal{F}) \rightarrow \mathbb{R}$.

To define T_3 , we start by assigning to each element $P \in l^\infty(\mathcal{F})$ a function $F_P : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ given by:

$$F_P(s, t) = P1_{[0,s] \times [0,t]} \quad (4.22)$$

As $P \in l^\infty(\mathcal{F})$ the function in 4.22 is bounded, and so it turns to be Lebesgue integrable⁶ over compact set $[0, 1] \times [0, 1]$.

We can thus define:

$$T_3(P) = \sup_{(x,y) \in [0,1] \times [0,1]} \int_0^x \int_0^y F_P(s, t) ds dt \quad (4.23)$$

Note that if P probability measure, then F_P turns out to be its associated cdf and we have:

$$T_3(P - Q) = \sup_{(x,y) \in [0,1] \times [0,1]} \int_0^x \int_0^y (F_P - F_Q)(s, t) ds dt \quad (4.24)$$

For the supermodular case let's start by defining for each $P \in l^\infty(\mathcal{F})$ the function $K_P : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ by:

$$K_P(s, t) = P1_{[0,s] \times [0,1]} + P1_{[0,1] \times [0,t]} - P1_{[0,s] \times [0,t]} \quad (4.25)$$

and as above we can define:

$$T_4(P) = \sup_{(x,y) \in [0,1] \times [0,1]} \int_0^x \int_0^y K_P(s, t) ds dt \quad (4.26)$$

so that for the case of probability measures:

$$T_4(P - Q) = \sup_{(x,y) \in [0,1] \times [0,1]} \int_0^x \int_0^y (K_P - K_Q)(s, t) ds dt \quad (4.27)$$

We are going to use two sample bootstrap to find critical values for test statistics. The following theorem is analogous to the one we proved for first order bivariate dominance.

⁶Should it be the case that function 4.22 was not measurable (perhaps P is some strange element in $l^\infty(\mathcal{F})$), we would use the exterior integral. For the case of finite signed measures it has to be measurable, because it is pretty much the same as a distribution function.

Theorem 4.4 (Second Order Bivariate Stochastic Dominance Test). : *Let's consider test statistics $\mu_{m,n}^B$ and $\gamma_{m,n}^B$ obtained using the empirical two sample bootstrap distributions (given by 4.9) in formulas 4.20 and 4.21. Then if $m, n \rightarrow \infty$ such that $m/N \rightarrow \alpha \in (0, 1)$:*

$$1. \mu_{m,n}^B \rightsquigarrow T_3(G_H)$$

$$2. \gamma_{m,n}^B \rightsquigarrow T_4(G_H)$$

given almost any sequence $V_1, V_2, \dots, W_1, W_2, \dots$, where $H = \alpha P + (1 - \alpha)Q$ and G_H is a version of the tight Brownian bridge corresponding to measure H .

Proof. We are in the same conditions as in Theorem 4.3, so it will suffice to show that T_3 and T_4 are continuous, since $\mu_{m,n}^B = T_3(\sqrt{\frac{mn}{m+n}}(\hat{P}_{m,N} - \hat{Q}_{n,N}))$ and $\gamma_{m,n}^B = T_4(\sqrt{\frac{mn}{m+n}}(\hat{P}_{m,N} - \hat{Q}_{n,N}))$. Now if P and Q are elements in $l^\infty(\mathcal{F})$ we have:

$$\begin{aligned} |T_3(P) - T_3(Q)| &= \\ &= \left| \sup_{(x,y) \in [0,1] \times [0,1]} \int_0^x \int_0^y F_P(s,t) ds dt - \sup_{(x,y) \in [0,1] \times [0,1]} \int_0^x \int_0^y F_Q(s,t) ds dt \right| \leq \\ &\leq \sup_{(x,y) \in [0,1] \times [0,1]} \left| \int_0^x \int_0^y F_P(s,t) ds dt - \int_0^x \int_0^y F_Q(s,t) ds dt \right| \leq \\ &\leq \sup_{(x,y) \in [0,1] \times [0,1]} \left| \int_0^x \int_0^y (F_P - F_Q)(s,t) ds dt \right| \leq \\ &\leq \sup_{(x,y) \in [0,1] \times [0,1]} \int_0^x \int_0^y |(F_P - F_Q)(s,t)| ds dt \leq \\ &\leq \sup_{(x,y) \in [0,1] \times [0,1]} \int_0^x \int_0^y \|P - Q\|_{\mathcal{F}} ds dt \leq \|P - Q\|_{\mathcal{F}} \quad (4.28) \end{aligned}$$

which implies T_3 is continuous and this proves the first part.

The same technique can be used to prove the second part, using triangular inequality as we did to prove the second part of Theorem 4.3. $\square$

Once we established the result, we can follow the same reasoning as in Sec. 4.2 to obtain data dependent critical values for the sequence of tests, which turns out to be consistent in the same sense as before.

5 Applications to Labor Market Econometrics: Income and Human Capital in Argentina (1996-2006)

5.1 Argentina and the 2001 Crisis

In the decade going from 1996 to 2006 Argentina's labor market suffered extreme changes. From 1996-1998, the last growing years in the 90s to recession at the turn of the century, to full-out depression with the collapse of the Argentinian peso's currency board in 2002 and recovery starting in 2004. This extreme macroeconomic events obviously had huge impact in

the labor market, with unemployment reaching peaks of 20% in 2002 and income inequality bursting into levels not seen since the hyperinflationary years at the end of the 80s.

The effect of these changes on aggregate economic welfare have been analyzed focusing mainly on income distribution. The work of Gasparini, Marchionni and Sosa Escudero 2001 traces the determinants of income inequality and Perez 2005 examines the effects of both income inequality and income polarization on the level of social conflict.

Appropriate multidimensional analysis for this period, within the statistical approach, has been lacking and in filling this gap we have an excellent opportunity to put into practice the stochastic dominance tests we developed in the previous sections.

5.2 Dimensions of Economic Inequality and Data Sources

Following the approach of Atkinson and Bourguignon 1982 and Perez 2015 we adopt bivariate Bergson-Samuelson functionals as a measure of economic welfare. Taking as input the bivariate distribution function $F(x, y)$ with compact support $\bar{U}$ for the economy as a whole, we have aggregate welfare:

$$W(F) = \int_{\bar{U}} \phi(x, y) dF(x, y) \quad (5.1)$$

where $\phi(x, y)$ is a utility function over $\mathbb{R}_+^2$. The key point here is that a higher aggregate welfare for a bivariate distribution F over a distribution G is represented by $W(F) \geq W(G)$ and that's exactly stochastic dominance over the class of utility functions $\{\phi\}$. This makes stochastic dominance a powerful tool in welfare analysis as it is directly translated into welfare comparisons. That's the reason for Atkinson and Bourguignon 1982 to choose stochastic dominance in their study multidimensional economic inequality, and we adopt the same approach.

There has been discussion on the topic of selection of both the number and the nature of the involved dimensions (see Perez 2015, Ch.2). From the computational point of view, although our tests are theoretically easy to extend to n -variables, the curse of dimensionality imposes severe restrictions, as the necessary sample size grows exponentially with each additional dimension.

We set to 2 dimensions, choosing income and human capital as variables. The choice of income is pretty much uncontroversial, as its impact on economic welfare is universally accepted. The selection of human capital as the second variable seems reasonable on several counts: (1) human capital is the key factor determining income distribution, as proven by the econometric analysis on the returns to schooling (Griliches, 1977); (2) other proposed variables as some health measure (Atkinson and Bourguignon, 1982) or employment status (Marks et al., 2000) are highly correlated with human capital and then their effect on welfare can to some extent be incorporated into the former's.

Having decided to use income and human capital joint distributions as input for our economic welfare functions we have to specify our data sources and explicitly define the variables. Our main data source is the household survey database from the National Institute of Statistics and Census in Argentina⁷. This is Argentina's most comprehensive and used source for micro

⁷Available at: https://www.indec.gob.ar/indec/web/Institucional-Indec-bases_de_datos_

data both in academic studies and in applied policy programs. It samples from urban households along the whole country, periodically surveying a wide range of socioeconomic variables. Among those variables we focus on household income and education to construct our variables of interest.

To construct our income variable we follow Perez 2005 letting $y_i = \log(Y_i)$, where $Y_i = \frac{TFI}{\sqrt{\sigma}}$. Here *TFI* stands for *Total Family Income* and it measures the aggregate household income, while σ is the household size. Although in poverty analysis it is usual to implement income normalization using the sum of *adult equivalent factor* for the household members, in the study of inequality the *square root rule* is prevalent (see Duclos, Esteban and Ray, 2004; OECD, 2008). To account for inflationary effects, we normalize using the consumer price index.

For human capital we adopt the approach of ? who measured the distribution and aggregate amount of human capital in Japan. According to that, we transform years of formal education (s_i) and age (a_i) for the i -th individual in the sample, into human capital (h_i) by means of a *nonlinear mincerian function*, that combines a second degree polynomial in experience and an exponential of education:

$$\log(h_i) = \gamma_1(a_i - s_i - 6) + \gamma_2(a_i - s_i - 6)^2 + \frac{\alpha}{1 - \psi} s_i^{1-\psi} \quad (5.2)$$

(5.2)

As our household survey data base doesn't survey cohort data, we can't easily model and estimate returns to schooling for Argentina directly. Thus, for the parameters α and ψ we take the values from Bils and Klenow 2000 who apply meta-regression analysis to the estimated returns to schooling for a set of 56 countries obtained by Psacharopoulos 1994. The values are $\gamma_1 = 0.0512$, $\gamma_2 = -0.00071$, $\alpha = 0.32$ and $\psi = 0.28$. For the conversion of attained education level (data surveyed) into years of education (the needed variable) we follow the convention in Margot 2001.

5.3 Test Implementation and Computations

To implement the tests developed in Section 4, we have first to decide to use either the submodular or supermodular classes. It is reasonable to assume that income and human capital are strategic complements in terms of utility, that is, for a bivariate utility function $\phi(x, y)$ with income and human capital as variables, we expect $\frac{\partial^2 \phi}{\partial x \partial y} \geq 0$. This is sensible, studies showing that marginal utility for income increases with human capital, as a more educated consumer has a wider range of goods and services available (Clark and Oswald, 1995).

We have then to use the tests developed for the supermodular case, which is slightly more complex than the submodular case because we have also to test for the marginal conditions. We use the test statistics developed in Section 4 for first and second order stochastic dominance in the supermodular case, recalling them here for the reader's convenience:

$$\kappa_{m,n} = \left(\frac{mn}{m+n} \right)^{1/2} \sup_{[0,1] \times [0,1]} (\hat{K}_{1m} - \hat{K}_{2n})(s, t) \quad (5.3)$$

$$\gamma_{m,n} = \left(\frac{mn}{m+n} \right)^{1/2} \sup_{[0,1] \times [0,1]} (\hat{L}_{1m} - \hat{L}_{2n})(s, t) \quad (5.4)$$

Test computations are straightforward, according to the following steps:

1. Given samples for different distributions F and G we first compute the overall test statistic ($\kappa_{m,n}$ or $\gamma_{m,n}$) for the two samples. Here, the different samples come from different periods, for example, when we are comparing welfare in the first semester of 1998 (distribution F) and in the second semester of 2003 (distribution G).
2. We then construct the pooled sample and proceed to resample with replacement, generating a pair of bootstrap distributions of the original sizes. We compute the test statistic for every such pair of bootstrap samples and then construct the bootstrap distribution for the test statistic.
3. Having set a significance level α we compare the critical value obtained from the bootstrap distribution with the value of the test statistic computed for the original sample, rejecting the hypothesis $H_0 : FSD_j G$ when the later is greater than the critical value.

We compute 100 bootstrap iterations with $\alpha = 0.05$, which is acceptable in terms of consistency (Efron and Tibshirani, 1994) given our big sample size of more than 25.000 observations for each variable. On the other hand, as our sample sizes impacts nonlinearly on the number of required computations to get the critical values, we can't afford a lot more of resamples. We implement the tests with Matlab routines, whose code is available at <https://github.com/luciano-alejo-perez/SD-code>.

5.4 Test Results and Welfare Rankings

Having two samples for each year, corresponding roughly to first and second semester, from 1996 to 2006⁸, we proceed to make head to head welfare comparisons between distributions. For each pair of distributions we have to test both $H_0 : FSD_j G$ and $\tilde{H}_0 : GSD_j F$ (here $j = 1$ means 1st order dominance and $j = 2$ means 2nd order).

Head to head comparisons are presented in the form of a dominance matrix (Table 2) where a value of 1 means the row dominates over the column, a value of -1 means the columns dominates the row, and a value of 0 means there is no definite dominance in either way. From the dominance matrices we can in turn develop welfare rankings (Table 1). As 1st order dominance implies 2nd order dominance, the ranking for 2nd order will not revert any previous welfare judgment but can result more decisive. This is usual for stochastic dominance, even for the univariate case.

To avoid an excessive use of space for the matrices (as each is would be of dimensions 22x22) we combine the results for both 1st and 2nd order dominance tests in only one matrix. Thus, we insert 2 values in each place, the first being the test result for 1st order dominance, the second (separated by a comma) the result for 2nd order dominance. Also, the nested nature of

⁸Except for the second semester of 2003 for which the data available for household income and size was lacking.

stochastic dominance allows us to only test for 2nd order dominance when there is not definite dominance at first order. That's why in Table 2 only the (off-diagonal) places with a 0 for 1st order comparisons have a second number representing the 2nd order comparison.

In analyzing the results shown in Table 2 we notice a remarkable blockwise structure of the dominance matrix. The distributions from first semester 2002 to first semester 2006 are dominated by any previous distribution. This shows the extreme impact of the collapse of the Argentine Peso currency board in 2002, which caused a jump in the threshold of inequality, a phenomenon already highlighted by Perez 2005 for income distribution.

Distribution (YY-Semester)	SD_1	Ranking	Distribution (YY-Semester)	SD_2	Ranking
1999-2		1	1999-2		1
1996-2		2	1996-2		2
2006-2		3	2006-2		3
1996-1		4	1999-1		4
1998-1		4	1998-1		5
1999-1		4	1996-1		6
2000-1		4	2000-1		6
2001-1		4	2001-1		6
2001-2		4	2001-2		6
1998-1		10	1998-2		10
1997-1		11	1997-1		11
1997-2		12	1997-2		12
2000-2		12	2000-2		12
2005-1		14	2005-1		14
2004-1		15	2004-2		15
2004-2		15	2005-2		15
2005-2		15	2006-1		17
2006-1		15	2004-1		18
2003-1		19	2003-1		19
2002-1		20	2002-1		20
2002-2		20	2002-2		20

Table 1: Welfare Ranking. 1st (left) and 2nd (right) Order Stochastic Dominance.

Dist	96-1	96-2	97-1	97-2	98-1	98-2	99-1	99-2	00-1	00-2	01-1	01-2	02-1	02-2	03-1	04-1	04-2	05-1	05-2	06-1	06-2
96-1	0	-1	1	1	0,-1	1	0,-1	-1	0,0	1	0,0	0,0	1	1	1	1	1	1	1	1	-1
96-2	1	0	1	1	1	1	1	-1	1	1	1	1	1	1	1	1	1	1	1	1	1
97-1	-1	-1	0	1	-1	-1	-1	-1	-1	-1	1	-1	-1	1	1	1	1	1	1	1	1
97-2	-1	-1	-1	0	-1	-1	-1	-1	-1	0,0	-1	-1	1	1	1	1	1	1	1	1	1
98-1	0,1	-1	1	1	0	1	0,-1	-1	0,1	1	0,1	0,1	1	1	1	1	1	1	1	1	1
98-2	-1	-1	1	1	-1	0	-1	-1	-1	1	-1	-1	1	1	1	1	1	1	1	1	1
99-1	0,1	-1	1	1	0,1	1	0	-1	0,1	1	0,1	0,1	1	1	1	1	1	1	1	1	1
99-2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
00-1	0,0	1	1	1	0,-1	1	0,-1	-1	0	1	0,0	0,0	1	1	1	1	1	1	1	1	1
00-2	-1	-1	-1	0,0	-1	-1	-1	-1	-1	0	-1	-1	1	1	1	1	1	1	1	1	1
01-1	0,0	1	1	1	0,-1	1	0,-1	-1	0,0	1	0	0,0	1	1	1	1	1	1	1	1	1
01-2	0,0	1	1	1	0,-1	1	0,-1	-1	0,0	1	0,0	0	1	1	1	1	1	1	1	1	1
02-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	0	0,0	-1	-1	-1	-1	-1	-1	-1
02-2	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	0,0	0	-1	-1	-1	-1	-1	-1	-1
03-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	1	1	0	-1	-1	-1	-1	-1	-1
04-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1	0	0,-1	-1	0,-1	0,-1	0,-1
04-2	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1	0,1	0	-1	0,0	0,0	0,0
05-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	0	1	1	1
05-2	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1	0,1	0,0	-1	0	0,1	0,1
06-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1	0,1	0,0	-1	0,-1	0	0
06-2	1	-1	1	1	1	1	1	-1	1	1	1	1	1	1	1	1	1	1	1	1	0

Table 2: Head to Head Bivariate Stochastic Dominance (1st and 2nd Order). Argentina 1996-2006

The matrix in Table 2 also shows that bivariate stochastic dominance is for the most part able to rank distributions decisively. This is an empirical property of stochastic dominance that was already recognized for income distribution analysis (Sen, 2006) and it is another reason to choose stochastic dominance in practice, especially in the assessment of active distributional policy.

In this respect, it has to be noticed that in going from univariate stochastic dominance in income distribution to bivariate dominance on the joint distribution of income and human capital, the comparisons can, and in fact do in some cases, become less decisive. That's because bivariate dominance is not a simple aggregation of marginal comparisons. The key aspect of bivariate dominance is that it incorporates correlation effects between dimensions. This means that for a change in joint distribution to have a positive impact on social welfare, it is not only necessary that it implies an improvement (equalization) on income or human capital (or both) distributions, but also an increase in the correlation between human capital and income.

That's why, although income distribution improved steadily after the 2002 crisis, bivariate stochastic dominance doesn't necessarily show such a monotone improvement. The key reason behind this different results lies probably in the fact that in the first years of economic recovery, the reduction of unemployment was focused primarily in the construction and public sectors, of typical low productivity.

6 Test Properties

6.1 Bootstrap Power Assessment

We've seen that our tests are consistent in the sense that our data dependent critical values converge to the quantiles of the underlying Brownian bridge distribution. The analysis of statistical power of our test is more difficult, as both the null hypothesis ($H_0 : FSD_j G$), and the alternative hypothesis ($H_1 : F \text{ does not dominate } G$) are composite. We can't then find directly a probability of accepting H_1 as the underlying law is unknown.

An usual technique in the assessment of the statistical power of a test is the use of *contiguous hypothesis*. These are alternative hypothesis whose laws Q_n are asymptotically absolutely continuous (in the sense of Radon-Nykodim) respect to the undelaying law for H_0 , lets call it P . A sufficient condition for this is that $d_h(P, Q_n) \rightarrow 0$ under the Heilinger metric (van der Vaart W. and Wellner, 2000). In practice, however, these asymptotic results are extremely hard to evaluate explicitly.⁹

We the recall an idea in Efron and Tibshirani 1994: *"Bootstrap tests are useful in situations where the alternative hypothesis is not well-specified"*. The idea is then, following Good 2005, to make a local power assessment in the following way: starting with a basic known distribution P we modify it in such a way that we know that the hypothesis $H_0 : PSD_j Q$ does not hold. If we don't alter the distribution too much, we'll have a law that's similar enough, although different, allowing us to evaluate the statistical power of our test.

⁹van der Vaart W. and Wellner 2000, p.408, obtain the power of a one sample *KS* test. This is not useful for us, since we don't have a fixed law under the null hypothesis, but even in this case the authors remark that results cannot be evaluated explicitly.

To assure that the null hypothesis does not hold we have to use *increasing correlation transforms*, as developed by Epstein and Tanny 1980. As higher correlation (keeping marginals intact) implies bivariate stochastic dominance, changing P into Q this way, we can ensure that $H_0 : PSD_j Q$ does not hold.

We proceed as follows: (i) we take a sample from a known law P ; (ii) we modify it, obtaining a distribution Q , such that $H_0 : PSD_j Q$ does not hold; (iii) we then sample from each, obtaining P_n and Q_n ; (iv) we implement our test; (v) repeat several times. The power of our test will be the proportion of times that the test rejects H_0 .

Starting with a sample of size 40.000 from a bivariate normal distribution in Matlab we apply a series of correlation increasing transformations, always keeping the differences in no more than 2% of the points, as to ensure the locality of the power assessment.

Taking samples of size 20.000 from each distribution and implementing our tests with a significance level of 0.05, the results on 100 trials were of 70% rejections for our 1st order dominance test and a little higher for the 2nd order test. This results are very near to the ideal of 80% usually mentioned in the experiment design literature (Suresh and Chandrashekara, 2012).

Note that both Bickel 1969 and Barret and Donald 2003 prove that the probability of rejecting the null if false converges asymptotically to 1 for their tests. But in both cases, the continuity of the underlying distributions is assumed, and turns out to be essential for the proof.

6.2 Advantages and Comparison to Other Tests

There are two clear advantages of our approach in testing multivariate economic inequality. First, stochastic dominance has a direct interpretation in terms of aggregate social welfare, by means of Bergson-Samuelson functionals. Also, our tests are straightforward to implement using almost any software package, as bootstrap methods are now widely used, and the method we developed to obtain critical values dependent on data has no complications.

For these reasons, we consider our tests to be practical and useful tools to both the empirical econometrician, the applied economist and the policy maker.

As for how are tests compare to other approaches to stochastic dominance, such as Crawford 2005 we have to point out that this test, based on Anderson 1996 approach, develops a Pearson goodness of fit test, working with grouped data. Working with grouped data simplifies the calculations, but at the same time implies a loss of consistency (Davidson and Duclos, 2000). Working with point-to-point comparisons of distribution functions as we do in our tests, ensures consistency. Besides the matter of statistical consistency, it has to be remarked that tests based on grouped, unlike our test, don't have a direct economic interpretation in terms of welfare, which is a very desirable property in applications.

We finally remark that Crawford test, as much of the literature on stochastic dominance, rely on the continuity of the underlying laws, an assumption our approach can dispense.

7 Concluding Remarks, Extensions and Limitations

The main contribution of this paper is the development of fully consistent tests of bivariate stochastic dominance at first and second order. As proven in the previous sections, this tests

have several desirable properties and advantages. First, as the tests are implemented with bootstrap methods, they turn out to be very straightforward to apply, not requiring neither complex asymptotic calculations nor specific distribution tabulations. This makes our tests very useful tools for the applied literature, both in the economic inequality and risk theory fields.

Second, as we work with the entire distribution function instead of grouped data or concentration indexes, our tests are fully consistent, with high power, and can be implemented in practice with minimal preparations for the data.

Third, but related to the last observation, working directly with the bivariate distribution function allows for a direct interpretation in the field of welfare economics, without further assumptions or axioms needed. In this respect, the tests presented in this paper can be regarded as the first fully consistent statistical implementation of the multivariate stochastic dominance conditions given by Atkinson and Bourguignon 1982 and Hadar and Russell 1974.

This remarks obviously don't mean there is not room for improvement. There are two obvious ways in which we could want to generalize our results. On the one hand, for higher orders of dominance in the bivariate case, and on the other hand, for the case with more than two variables. There are not much difficulties in doing so, at least formally. We can as Perez 2015 define a sequence of operators $T_j : l^\infty(\mathcal{F}) \rightarrow \mathbb{R}$ and proceed as before to show that those operators are continuous.

For stochastic dominance in higher dimensions we can also generalize modularity classes in a direct way, just by imposing conditions on the sign of higher order partial derivatives. If we can get around the complexity of the great number of integrals that the use of the integration by parts formula would imply even in the three dimensional case, we could obtain sufficient conditions generalizing those in Hadar and Russell 1974, Atkinson and Bourguignon 1982 and Perez 2018.

But the key point here is not of formal but if practical kind. Both in higher orders of dominance and in higher dimensions we are dealing with an increasing number of multivariate integrals. The problem arises then as a manifestation of the well-known curse of dimensionality: as we increase the number of integrals involved in our tests, the size of the data sets needed for efficient testing grows exponentially. As we are dealing with bootstrap re-sampling, this in turn implies that the amount of computations involved rapidly grows.

So in most practical applications at the time we have to content ourselves with first and second order bivariate stochastic dominance. As shown in the previous sections, this suffices for interesting applications in the field of welfare economics, such as the interaction between human capital and income and its impact on two-dimensional economic inequality.

References

- Anderson, G. 1996. Nonparametric tests of stochastic dominance in income distributions, *Econometrica* **64**(5): 1183–1193.
- Atkinson, A. B. 1970. On the measurement of inequality, *Journal of Economic Theory* **2**: 244–263.

- Atkinson, A. B. and Bourguignon, F. 1982. The comparison of multi-dimensioned distributions of economic status, *The Review of Economic Studies* **49**(2): 183–201.
- Barret, G. and Donald, S. 2003. Consistent tests for stochastic dominance, *Econometrica* **73**(1): 71–104.
- Barret, G. and Donald, S. 2004. Consistent nonparametric tests for lorenz dominance, *Econometric Society, Australasian Meeting 2004*.
- Bickel, P. 1969. A distribution free version of the smirnov two sample test in the p-variate case, *Ann. Math. Statist.* **40**(1): 1–23.
- Billingsley, P. 1999. *Convergence of probability measures*, Wiley Series in Probability and Statistics, second edn, John Wiley & Sons Inc., New York.
- Bils, M. and Klenow, P. 2000. Does schooling cause growth?, *The American Economic Review* **90**: 1160–1183.
- Brufman, J. and Perez, L. 2011. Distribución muestral de índices de desigualdad económica: Enfoque no paramétrico, *XXVI Jornadas Nacionales de Docentes de Matemática de Facultades de Ciencias Económicas y Afines*, Facultad de Ciencias Económicas, Universidad Nacional de Misiones.
- Cabaña, A. and Cabaña, E. 1997. Transformed empirical processes and modified kolmogorov-smirnov statistics for multivariate distributions, *The Annals of Statistics* **25**(6): 2388–2409.
- Clark, A. and Oswald, A. 1995. Satisfaction and comparison income, *Journal of Public Economics* **61**(3): 359–381.
- Crawford, I. 2005. A nonparametric test of stochastic dominance in multivariate distributions, *Discussion Papers in Economics, University of Surrey* **DP 12/05**.
- 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.
- Duclos, J.-Y., Esteban, J. and Ray, D. 2004. Polarization: Concepts, measurement, estimation, *Econometrica* **72**(6): 1737–1772.
- Dudley, R. 1999. *Uniform Central Limit Theorems*, Vol. 63 of *Cambridge Studies in Advanced Mathematics*, Cambridge University Press, Cambridge.
- Dudley, R. 2003. *Real Analysis and Probability*, Vol. 74 of *Cambridge Studies in Advanced Mathematics*, Cambridge University Press, Cambridge.
- Efron, B. and Tibshirani, R. 1994. *An Introduction to the Bootstrap*, Chapman & Hall/CRC Monographs on Statistics & Applied Probability, Taylor & Francis.
- Epstein, L. and Tanny, S. 1980. Increasing generalized correlation: A definition and some economic consequences, *The Canadian Journal of Economics / Revue Canadienne d'Econometique* **13**(1): 16–34.

- Friedman, J. and Rafsky, L. 1979. Multivariate generalization of the wald-wolfowitz and smirnov two sample tests, *The Annals of Statistics* **7**(4): 697–717.
- Gasparini, L., Marchionni, M. and Sosa Escudero, W. 2001. *Distribución del Ingreso en Argentina: Perspectivas y Efectos sobre el Bienestar*, Fundación Arcor.
- Good, P. 2005. *Permutation, Parametric and Bootstrap Tests of Hypotheses*, Springer Series in Statistics, Springer, New York.
- Griliches, Z. 1977. Estimating the returns to schooling: Some econometric problems, *Econometrica* **45**(1): 1–22.
- Hadar, J. and Russell, W. R. 1969. Rules for ordering uncertain prospects, *Amer. Econ. Rev.* **59**: 25–34.
- Hadar, J. and Russell, W. R. 1974. Stochastic dominance in choice under uncertainty, in D. L. M. M. S. Balch and Y. Wu (eds), *Essays on Economic Behavior under Uncertainty*, North Holland, Amsterdam.
- Horváth, L., Kokoszka, P. and Zitikis, R. 2006. Testing for stochastic dominance using the weighted mcfadden type statistic, *Journal of Econometrics* **133**(1): 191–205.
- Margot, D. 2001. Rendimientos a la educación en argentina: Un análisis de cohortes, *Documento de Trabajo, Facultad de Ciencias Económicas, UNLP* **33**.
- Marks, G. et al. 2000. *The Measurement of Socioeconomic Background for the Reporting of Nationally Comparable Outcomes of Schooling, Draft Report*, National Education Performance Monitoring Taskforce, Australian Council for Educational Research & Sociology Program Research School of Social Sciences Australian National University.
- McFadden, D. 1989. Testing for stochastic dominance, in T. B. Fomby and T. K. Seo (eds), *Studies in the Economics of Uncertainty. In Honor of Josef Hadar*, Springer, New York.
- Miyazawa, K. 2011. Measuring human capital in japan, *RIETI Discussion Paper Series* **11-E-037**.
- Migheli, M. 2018. Breadwinner's Education and Household's Income in India, *International Journal of Statistics & Economics* **19**(2): 61–83.
- OECD 2008. *Growing Unequal? Income Distribution and Poverty in OECD Countries*, OECD, Paris.
- Perez, L. A. 2005. Polarización del ingreso en la argentina. 1999-2004. <http://finanzas.x10host.com/Perez2005.pdf>.
- Perez, L. A. 2015. *Tests de Dominancia Estocástica en Base a Estadísticos de Kolmogorov-Smirnov Multivariados, con Aplicaciones a la Desigualdad Económica Multidimensional*, PhD thesis, University of Buenos Aires, Facultad de Ciencias Económicas, Argentina. http://bibliotecadigital.econ.uba.ar/download/tesis/1501-1250_PerezL.pdf.

- Perez, L. A. 2018. Modularity classes and boundary effects in multivariate stochastic dominance, *Cuadernos del Cimbago* **2**(20): 55–77.
- Psacharopoulos, G. 1994. Returns to investment in education: A global update, *World Development* **22**(9): 1325–1343.
- Schmid, P. 1958. On the kolmogorov and smirnov limit theorems for discontinuous distribution functions, *Ann. Math. Stat.* **29**(24): 1011–1027.
- Sen, A. 2006. *Inequality Reexamined*, Claredon Press, Oxford.
- Sen, P. 1973. An almost sure invariance principle for kolmogorov smirnov statistics, *The Annals of Probability* **1**(3): 488–496.
- Simpson, P. 1951. A note on the estimation of a bivariate distribution function, *Ann. Math. Statist.* **22**(3): 476–478.
- Suresh, K. and Chandrashekara, S. 2012. Sample size estimation and power analysis for clinical research studies, *J Hum Reprod Sci* **5**(1): 7–13.
- Topkis, D. 1998. *Supermodularity and Complementarity*, Princeton University Press, Princeton.
- Tsui, K. 1995. Multidimensional generalizations of the relative and absolute inequality indices: The atkinson-kolm-sen approach, *J. Ec. Theo* **67**(1): 251–265.
- van der Vaart W. and Wellner, J. 2000. *Weak Convergence and Empirical Processes. With Applications to Statistics*, Springer Series in Statistics, Springer, New York.
- Whang, Y.-J. 2019. *Econometric Analysis of Stochastic Dominance Concepts, Methods, Tools, and Applications*, Cambridge University Press, Cambridge.