Discussion Paper Series

RIEB

Kobe University

DP2014-36

An Axiomatic Approach to Measuring Degree of Stochastic Dominance

Takashi KAMIHIGASHI
John STACHURSKI

November 10, 2014

Research Institute for Economics and Business Administration
Kobe University
2-1 Rokkodai, Nada, Kobe 657-8501 JAPAN

An Axiomatic Approach to Measuring Degree of Stochastic Dominance[1]

Takashi Kamihigashi[a,b] and John Stachurski[c]

[a]Research Institute for Economics and Business Administration, Kobe University

[b]IPAG Business School, Paris, France

[c]Research School of Economics, Australian National University

October 23, 2014

ABSTRACT. In recent years, a range of measures of "partial" or "degree of" stochastic dominance have been introduced. These measures attempt to determine the extent to which one distribution is dominated by another. In order to systematically assess these proposed measures and their relationship to partial stochastic dominance, we adopt an axiomatic approach. We propose axioms for measuring degree of stochastic dominance and study the relationship between them. Among other findings, we show that one and only one measure satisfies all the axioms.

JEL Classifications: D81, G11
Keywords: Stochastic dominance, stochastic order

1. INTRODUCTION

Stochastic dominance is a standard method for comparing distributions in economics, finance, statistics and related fields. For example, it is a fundamental concept in the theory of choice under uncertainty (see, e.g., Machina (1982)). It plays an important role in the theory of finance, being closely related to existence of arbitrage opportunities and portfolio choice (e.g., Jarrow (1986), Hadar and Russell (1971)). In social welfare studies, policies or decision rules that shift outcomes up

[1]This paper has benefited from comments from Andy McLennan and Rabee Tourky, as well as financial support from the Japan Society for the Promotion of Science and Australian Research Council Discovery Grant DP120100321.

Email addresses: tkamihig@rieb.kobe-u.ac.jp, john.stachurski@anu.edu.au

in terms of stochastic dominance can be regarded as universally preferred under weak assumptions on individual choices. Stochastic dominance also serves as a standard ordering for considering parametric monotonicity in settings where distributions are compared.[2]

In recent years, researchers interested in comparing distributions have introduced a range of different notions of partial first order stochastic dominance. One well-known example is so-called "restricted stochastic dominance," which compares the order of cumulative distribution functions (cdfs) only up to some specified point in the domain (Atkinson (1987); Davidson and Duclos (2000, 2013)). Another is "almost stochastic dominance," which was developed in the finance literature (Leshno and Levy (2002); Levy (2009)). This idea has been applied widely and extended in various directions to cover additional use cases or simplify calculation (see, e.g., Lizyayev and Ruszczyński (2012), Denuit et al. (2013) and Tzeng et al. (2013)). A third kind of measure was proposed in Fields et al. (2002), who analyze degree of stochastic dominance in the context of income mobility analysis by comparing orders of cdfs at quantile markers. Still more measures are considered in Stoyanov et al. (2012).

The proliferation of distinct measures suggests the need for a systematic study of "degree of" or "partial" stochastic dominance. In this paper we provide an axiomatic treatment. Measures of degree of stochastic dominance are regarded as functions that evaluate the extent to which one probability distribution is dominated by another in the sense of stochastic dominance. We consider axioms that have some claim to be natural properties for such measures, and discuss their logical consistency. We then compare the various measures of degree of stochastic dominance found in the literature with our axioms. Among other findings, we show that exactly one measure satisfies all the axioms.

[2] Here and below, stochastic dominance always refers to first order stochastic dominance. The basic properties of stochastic dominance were initially provided by Lehmann (1955), and introduced to economics by Quirk and Saposnik (1962). As stated by Border (1992), "the concept has been independently discovered too many times for an exhaustive listing." Recent applications related to parametric monotonicity include Acemoglu and Jensen (2012) and Balbus et al. (2012). For textbook treatments see Stokey and Lucas (1989), Lindvall (2002), or Föllmer and Schied (2011).

Section 2 discusses preliminaries, formalizes the notion of a measure of degree of stochastic dominance and gives examples. Section 3 proposes axioms and studies their logical relationships and implications. Section 4 considers measures proposed in the literature in light of the axioms. Section 5 concludes.

2. Set Up

We begin with some preliminary definitions and notation, as well as a discussion of existing measures of degree of stochastic dominance.

2.1. **Preliminaries.** To discuss stochastic dominance, the space S over which distributions are defined must at least have some notion of order. A number of the measures of degree of stochastic dominance we consider require that S is an interval of $\mathbb{R}$, paired with its usual order $\leq$. Other measures are defined on more general spaces. To avoid an excessively taxonomical discussion and make available a number of equivalent views of stochastic dominance, we always assume that S is a complete, separable metric space (i.e., a Polish space) that contains at least two elements, and that S is ordered by a closed partial order $\preceq$.[3] This setting includes most applications of interest.

As usual, $h \colon S \to \mathbb{R}$ is called *increasing* if $x \preceq y$ implies $h(x) \leq h(y)$. A set $B \subset S$ is called an *increasing set* if its indicator function $\mathbb{1}_B$ is an increasing function (equivalently, $x \in B$ and $x \preceq y$ implies $y \in B$). Let $\mathcal{B}$ denote the Borel sets of S and let $\mathcal{O}$ be the open sets. We write $\mathcal{B}_I$ for the increasing Borel sets, $\mathcal{O}_I$ for the increasing open sets and $\mathcal{H}$ for the increasing Borel measurable functions from S to $[0,1]$. Let $\mathcal{P}$ be the probability measures on $(S, \mathcal{B})$. We say that $\mu \in \mathcal{P}$ is *stochastically dominated* by $\nu \in \mathcal{P}$ and write $\mu \preceq_{sd} \nu$ if $\int h \, d\mu \leq \int h \, d\nu$ for all $h \in \mathcal{H}$. Under our assumptions on $(S, \preceq_{sd})$, the relation $\preceq_{sd}$ is a partial order on the set $\mathcal{P}$ of probability measures (Kamae and Krengel, 1978, theorem 2).

Stochastic dominance can be characterized in several ways. The statement $\mu \preceq_{sd} \nu$ is equivalent to $\mu(I) \leq \nu(I)$ for all $I \in \mathcal{B}_I$. See theorem 1 of Kamae et al. (1977). We can also characterize the order by couplings. In particular, given $(\mu, \nu) \in \mathcal{P} \times \mathcal{P}$, a pair of S-valued random variables (X, Y) defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called a *coupling* of (μ, ν) if $\mu(B) = \mathbb{P}\{X \in B\}$ and $\nu(B) = \mathbb{P}\{Y \in B\}$

[3] A *closed partial order* on S is a binary relation $\preceq$ on $S \times S$ that is reflexive, transitive and antisymmetric, and such that its graph $\{(x,y) \in S \times S : x \preceq y\}$ is closed in the product topology.

for all $B \in \mathcal{B}$. The set of all couplings of (μ, ν) is denoted below by $\Gamma(\mu, \nu)$. With this notation we have the following well known theorem:

$$\mu \preceq_{sd} \nu \quad \iff \quad \exists\,(X, Y) \in \Gamma(\mu, \nu) \text{ s.t. } \mathbb{P}\{X \preceq Y\} = 1. \tag{1}$$

See Kamae et al. (1977) or chapter 4 in Lindvall (2002). Finally, if S is an interval in $\mathbb{R}$ then cdf F is stochastically dominated by cdf G if $G(x) \leq F(x)$ for all $x \in \mathbb{R}$. See, for example, section 6.D of Mas-Colell et al. (1995).

2.2. **Measures of Degree of Stochastic Dominance.** In this section we discuss various measures of degree of stochastic dominance that have been introduced in the literature. Here and below, a *measure of degree of stochastic dominance* is any function

$$\delta \colon \mathcal{P} \times \mathcal{P} \to [0, 1] \text{ such that } \delta(\mu, \nu) = 1 \text{ whenever } \mu \preceq_{sd} \nu. \tag{2}$$

The idea is that $\delta(\mu, \nu)$ measures the extent to which μ is dominated by ν, attaining its maximum value when domination is complete. At this stage the definition is intentionally weak, so that the main notions of degree of stochastic dominance from the existing literature can be cast in our framework after minor modifications.

One simple measure of degree of stochastic dominance is the quantile ratio measure proposed in Fields et al. (2002). Let $S = [a, b]$ and let $\preceq$ be the usual order $\leq$ on $\mathbb{R}$. Letting F and G be cdfs on $[a, b]$, define

$$q(F, G) := \frac{m}{n} \quad \text{where} \quad m := \sum_{i=1}^{n} \mathbb{1}\{G(x_i) \leq F(x_i)\}. \tag{3}$$

Here $\{x_i\}$ is a grid of n specified values in $[a, b]$, typically corresponding to some quantile points. Thus $q(F, G)$ measures the fraction of times that $G(x_i) \leq F(x_i)$ is observed over specified test points. Clearly q satisfies $q(F, G) = 1$ when $G \leq F$ pointwise (i.e., $F \preceq_{sd} G$), which corresponds to (2).

Another version of partial stochastic dominance is so-called *restricted* stochastic dominance (see, e.g. Atkinson (1987) or Davidson and Duclos (2013)). Let F and G be one dimensional cdfs on an interval $[a, b]$. Given $c \in [a, b]$, distribution F is said to be dominated by G in the restricted sense if $G(x) \leq F(x)$ for all $x \leq c$. We can turn this into a measure by considering the largest such c, defined by

$$c^*(F, G) := \sup\{c \in [a, b] \,:\, G(x) \leq F(x),\ \forall\, x \leq c\}. \tag{4}$$

After normalizing we get

$$r(F, G) := \frac{c^*(F, G) - a}{b - a}. \tag{5}$$

Evidently r is a measure of degree of stochastic dominance in the sense of (2).[4]

Another popular measurement for partial stochastic dominance is so-called *almost* stochastic dominance (Levy (1992); Leshno and Levy (2002)). Once again the context is $S = [a, b]$ with the usual order $\leq$. For cdfs F and G on $[a, b]$, the measure can be expressed as

$$\alpha(F, G) := \frac{\int (F(x) - G(x))_+ \, dx}{\int |F(x) - G(x)| \, dx} \tag{6}$$

where $v_+ := \max\{v, 0\}$ for any $v \in \mathbb{R}$. Intuitively, if F is almost dominated by G, then $G \leq F$ on most of its domain, and $(F(x) - G(x))_+ = |F(x) - G(x)|$ for most x. Hence $\alpha(F, G)$ is close to 1. In order to ensure that the measure is defined for all pairs F, G, we need to consider the case $F = G$. Since $F = G$ implies $F \preceq_{sd} G$, we adopt the convention that $\alpha(F, G) = 1$ for such pairs.[5]

More measures of degree of stochastic dominance were studied in the (Stoyanov et al., 2012, p. 21). Among other possibilities, they suggest measures that consider the difference $\int h d\mu - \int h d\nu$ as h varies over some class of real-valued increasing functions on S. One natural way to implement this idea is to define

$$\rho(\mu, \nu) := 1 - \sup_{h \in \mathcal{H}} \left\{ \int h d\mu - \int h d\nu \right\}. \tag{7}$$

It is easy to see that ρ is a measure of degree of stochastic dominance in the sense defined above.[6] As clarified below, when $S = \mathbb{R}$ a simple expression for (7) exists in terms of cdfs.

---

[4]One could also consider taking $r(F, G)$ to be $\lambda\{x \in S : G(x) \leq F(x)\}$, where λ is the uniform distribution or some other probability measure. A similar idea appears in Stoyanov et al. (2012). This is simpler to express than (5) and closely related to q in (3) but more difficult to compute in applications.

[5]Our definition of α in (6) differs from the standard definition in that $\alpha(F, G) = 1 - s$ where s is the standard definition. This is because our convention is that values close to one should indicate almost stochastic dominance.

[6]Stoyanov et al. (2012) also suggest replacing $\mathcal{H}$ with $\mathcal{L}$ in (7), where $\mathcal{L}$ is the set of all $h \in \mathcal{H}$ satisfying $|h(x) - h(y)| \leq d(x, y)$ for all $x, y \in S$. Here d is the metric on S. The measure γ satisfies (2) because the supremum on the right hand side is zero when $\mu \preceq_{sd} \nu$. Such a measure has the advantage that it respects weak convergence on $\mathcal{P}$. On the other hand it is difficult to calculate in applications and fails to satisfy all the axioms below. We do not discuss it further.

3. Axioms

In this section we consider properties that have some claim to be regarded as axiomatic for a well behaved measure of degree of stochastic dominance.

Axiom 3.1. If $(\mu, \nu) \in \mathcal{P} \times \mathcal{P}$ and $\delta(\mu, \nu) = 1$, then $\mu \preceq_{sd} \nu$.

Axiom 3.1 is a relatively weak and naturally desirable property. Our definition of a measure of degree of stochastic dominance already provides the converse implication (see (2)), so if δ satisfies axiom 3.1 then it takes the value 1 if and only if the two distributions in question are stochastically ordered.

Axiom 3.2. Let $\lambda \in [0, 1]$ and let (μ, ν) be any pair in $\mathcal{P} \times \mathcal{P}$ satisfying

$$\mu = \lambda\mu' + (1 - \lambda)\mu'' \quad \text{and} \quad \nu = \lambda\nu' + (1 - \lambda)\nu'' \tag{8}$$

for some $\mu', \mu'', \nu', \nu'' \in \mathcal{P}$. If, in addition, $\mu' \preceq_{sd} \nu'$, then $\delta(\mu, \nu) \geq \lambda$.

Axiom 3.2 is motivated by a desire for continuity, in the sense that $\delta(\mu, \nu)$ should be close to 1 when μ is "nearly" dominated by ν. To see the connection, consider a pair $\mu' \preceq_{sd} \nu'$, so that $\delta(\mu', \nu') = 1$. Suppose we now mix in two other lotteries μ'' and ν'', by taking $\lambda < 1$ in (8). This produces new lotteries μ and ν. If μ' and ν' are the main components of μ and ν in the sense that λ is close to 1, then μ is "almost dominated" by ν. Axiom 3.2 assures us that in this setting $\delta(\mu, \nu)$ will be close to one.

Axiom 3.3. For each $(\mu, \nu) \in \mathcal{P} \times \mathcal{P}$ and $\epsilon > 0$, there exists a $(X, Y) \in \Gamma(\mu, \nu)$ with

$$\mathbb{P}\{X \preceq Y\} \geq \delta(\mu, \nu) - \epsilon. \tag{9}$$

Axiom 3.3 is motivated by the desire that $\delta(\mu, \nu)$ should not be close to 1 unless (μ, ν) is "almost ordered." To see this, recall from (1) that if μ is dominated by ν then we can find a coupling (X, Y) such that $\mathbb{P}\{X \preceq Y\} = 1$. To retain consistency with this idea, $\delta(\mu, \nu)$ should not be large unless we can find a coupling (X, Y) such that $\mathbb{P}\{X \preceq Y\}$ is also large.

3.1. Other Candidates.

Another possible axiom for a measure of degree of stochastic dominance is $\delta(\mu, \nu) = 1$ and $\delta(\nu, \mu) = 1$ implies $\mu = \nu$. This is motivated by the fact that $\preceq_{sd}$ is antisymmetric. However, such an implication holds as soon as axiom 3.1 holds. Hence we omit it. Similarly, one might consider a transitivity requirement such as: if $\{\mu_n\}$, $\{\mu'_n\}$ and $\{\mu''_n\}$ are sequences in $\mathcal{P}$, then

$$\delta(\mu_n, \mu'_n) \to 1 \text{ and } \delta(\mu'_n, \mu''_n) \to 1 \implies \delta(\mu_n, \mu''_n) \to 1.$$

However it can be shown that transitivity is an implication of axioms 3.2 and 3.3, and hence we omit it from the list. Finally, it might be argued that

$$\delta(\mu, \nu) + \delta(\nu, \mu) = 1 \quad \text{whenever} \quad \mu \neq \nu \tag{10}$$

is a desirable property for a measure of degree of stochastic dominance. However, (10) is logically inconsistent with both axiom 3.2 and axiom 3.3. For example, if $\mu = \nu$ then both $\mu \preceq_{sd} \nu$ and $\nu \preceq_{sd} \mu$ are true. If we now perturb these measures very slightly to produce μ' and ν', continuity implies that both $\delta(\mu', \nu')$ and $\delta(\nu', \mu')$ will be close to one, so (10) fails.[7]

3.2. **Implications.** We now consider the logical relationships and implications of axioms 3.1–3.3. First, no pair of axioms is logically inconsistent. Theorem 4.1 below shows that all can hold at the same time. The main logical connection between the axioms is that axiom 3.3 implies axiom 3.1:

Proposition 3.1. *If δ is a measure of degree of stochastic dominance that satisfies axiom 3.3, then it satisfies axioms 3.1 as well.*

Axioms 3.2 and 3.3 are logically independent. Neither one implies the other, as is clear from the direction of the bounds they put on δ. Axiom 3.2 does not imply axiom 3.1, as can be seen from examples in section 4.1.

Next we show that the axioms we have listed are strong enough to give uniqueness. The following theorem records this fact, with δ and δ' being any two measures of degree of stochastic dominance.

Theorem 3.1. *If δ and δ' satisfy axioms 3.1–3.3, then $\delta = \delta'$.*

In view of proposition 3.1, we can alternatively state theorem 3.1 by saying that if δ and δ' satisfy axioms 3.2 and 3.3, then $\delta = \delta'$.

[7] More formally, suppose first that axiom 3.2 holds. If we take $\mu' = \nu' = \lambda$ in (8), where λ is some fixed element of $\mathcal{P}$, then axiom 3.2 implies that $\delta(\mu, \nu) \geq \lambda$ and $\delta(\nu, \mu) \geq \lambda$. For $\lambda > 1/2$ this is inconsistent with (10). Furthermore, if S contains elements that are not ordered in either direction, then the left-hand side of (10) can be made equal to zero.

4. Measures vs Axioms

We now reconsider the measures of stochastic dominance listed in the introduction in light of the axioms.

4.1. **Examples and Comparisons.** A first observation is that all measures of degree of stochastic dominance considered in section 2.2 satisfy axiom 3.1 except for the measure q defined in (3). The proofs are straightforward and hence omitted.

Regarding axiom 3.2, note first that q fails axiom 3.2 whenever the grid $\{x_i\}$ has at least three points. To see this let F', F'' and G'' be any cdfs on $[a,b]$ such that $F'' < G''$ on (a,b). Let $\lambda \in (0,1)$ and let

$$F = \lambda F' + (1-\lambda)F'', \quad G = \lambda F' + (1-\lambda)G''.$$

Since $F' \preceq_{sd} F'$, axiom 3.2 implies that $q(F,G) \geq \lambda$. On the other hand, $G(x_i) > F(x_i)$ on any interior point x_i. Hence m in (3) is at most 2, and $q(F,G) \leq 2/n$. Since λ can be arbitrarily close to 1 this contradicts axiom 3.2.

For this same pair F, G, the fact that $G > F$ on (a,b) implies that $c^*(F,G) = 0$ for c^* defined in (4), and hence $r(F,G) = 0$ for the restricted stochastic dominance measure defined in (5). Likewise, for the same pair, $\alpha(F,G) = 0$, where α is the almost stochastic dominance measure. Hence r and α also fail to satisfy axiom 3.2.

On the other hand, the stochastic dominance measure ρ defined in (7) satisfies axiom 3.2. We establish this fact below, in theorem 4.1.

Regarding axiom 3.3, the measures q, r and α all fail the axiom, while ρ satisfies it. To see this, consider the following pair of distributions on $S = [0,1]$. Given some small positive number ϵ, let μ put mass ϵ on 0 and $1-\epsilon$ on 1, and let ν put all mass on $1-\epsilon$. Let F be the cdf of μ and let X be a draw from μ. Let G and Y be the cdf of and a draw from ν respectively. Since Y is certainly $1-\epsilon$ we have $X \leq Y$ if and only if $X = 0$, and hence $\mathbb{P}\{X \leq Y\} = \epsilon$ for all couplings. On the other hand, $F(x) > G(x)$ iff $0 \leq x < 1-\epsilon$, and hence $r(F,G) = 1-\epsilon$. If $\epsilon < 1/2$ then $1-\epsilon > \epsilon$, and hence r fails axiom 3.3.

For this pair (μ,ν), the measures q and α also give values larger than ϵ when ϵ is small. The arguments are not difficult and the details are omitted. By contrast, we can see that $\sup_{h \in \mathcal{H}} \left\{ \int h d\mu - \int h d\nu \right\} \geq 1-\epsilon$ by choosing $h = \mathbb{1}_{(1-\epsilon,1]}$. It follows from the definition of ρ in (7) that $\rho(\mu,\nu) \leq \epsilon$.

4.2. **Properties of ρ.** The following result has been alluded to above and we state it here for the record:

Theorem 4.1. *The measure ρ defined in (7) is the unique measure of degree of stochastic dominance that satisfies axioms 3.1–3.3.*

In view of this result we invest in uncovering some additional properties of ρ. Our first observation is the following equivalences:

Proposition 4.1. *For any $(\mu, \nu) \in \mathcal{P} \times \mathcal{P}$, the value $\rho(\mu, \nu)$ can also be expressed as*

$$\rho(\mu, \nu) = 1 - \sup_{I \in \mathcal{B}_I} \{\mu(I) - \nu(I)\} = 1 - \sup_{I \in \mathcal{O}_I} \{\mu(I) - \nu(I)\}.$$

These equalities can help simplify calculations. For example, consider the case of $S = \mathbb{R}$, where many applications of stochastic dominance take place. Observe that, by taking complements, $\sup_{I \in \mathcal{O}_I} \{\mu(I) - \nu(I)\}$ is equal to the supremum of $\nu(D) - \mu(D)$ over all closed decreasing sets D. In the one-dimensional case, this class of sets is the intervals $(-\infty, x]$ over $x \in \mathbb{R}$. It follows that for one dimensional distributions with cdfs F and G, the measure ρ has the simple representation

$$\rho(F, G) = 1 - \sup_x \{G(x) - F(x)\}. \tag{11}$$

If F and G are absolutely continuous with densities f and g respectively, then the maximizer in (11) exists and the first order condition is $f(x) = g(x)$.

Finally, one desirable property for a measure of degree of stochastic dominance is that if a pair μ, ν almost satisfies $\mu \preceq_{sd} \nu$ in the sense of being close to a pair that is ordered, then $\delta(\mu, \nu)$ should be close to one. It turns out that the axioms enforce this. In the statement of the result we let $\|\mu - \nu\| := \sup_{B \in \mathcal{B}} |\mu(B) - \nu(B)|$.

Proposition 4.2. *For any two pairs of distributions (μ, ν) and (μ', ν') we have*

$$|\rho(\mu, \nu) - \rho(\mu', \nu')| \leq \|\mu - \mu'\| + \|\nu - \nu'\|.$$

5. Conclusion

In this paper we propose an axiomatic treatment for the problem of measuring degree of (or extent of) first order stochastic dominance. We put forward three axioms and discuss their rationale and relationship to one another. Among other things, we show that there exists exactly one measure of degree of stochastic dominance that satisfies all axioms, and characterize its properties. Further discussion of axioms and the development of similar ideas in the context of higher order stochastic dominance are avenues for future research.

6. PROOFS

This section collects all remaining proofs. We begin with a useful lemma.

Lemma 6.1. *For any* $(\mu, \nu) \in \mathcal{P} \times \mathcal{P}$ *we have*

$$\int h d\mu - \int h d\nu \leq 1 - \mathbb{P}\{X \preceq Y\}, \quad \forall (X, Y) \in \Gamma(\mu, \nu), \forall h \in \mathcal{H}. \tag{12}$$

Proof. Fix $h \in \mathcal{H}$ and any coupling (X, Y) of (μ, ν). Observe that

$$\begin{aligned}\int h d\mu - \int h d\nu &= \mathbb{E}\, h(X) - \mathbb{E}\, h(Y) \\ &= \mathbb{E}\,[h(X) - h(Y)] \mathbb{1}\{X \preceq Y\} + \mathbb{E}\,[h(X) - h(Y)] \mathbb{1}\{X \npreceq Y\}.\end{aligned}$$

Since $h \in \mathcal{H}$ we know that h is increasing. As a result, we have $h(X) \leq h(Y)$ on $\{X \preceq Y\}$. Using this inequality and nonnegativity of h we obtain

$$\int h d\mu - \int h d\nu \leq \mathbb{E}\,[h(X) - h(Y)] \mathbb{1}\{X \npreceq Y\} \leq \mathbb{E}\, h(X) \mathbb{1}\{X \npreceq Y\}.$$

Since $h \leq 1$ the last term is dominated by $\mathbb{P}\{X \npreceq Y\}$. Hence (12) holds. □

Proof of proposition 3.1. Suppose that δ satisfies axiom 3.3 and that $\delta(\mu, \nu) = 1$. From these hypotheses we have existence of a sequence of couplings $(X_n, Y_n) \in \Gamma(\mu, \nu)$ with $\mathbb{P}\{X_n \preceq Y_n\} \to 1$. The claim $\mu \preceq_{sd} \nu$ now follows from lemma 6.1. □

Proof of theorem 3.1. Let δ and δ' be two maps from $\mathcal{P} \times \mathcal{P} \to [0, 1]$ that both satisfy axioms 3.2 and 3.3. Fix $(\mu, \nu) \in \mathcal{P} \times \mathcal{P}$. Without loss of generality we assume that $\delta(\mu, \nu) \leq \delta'(\mu, \nu)$. We claim that the reverse inequality holds. Note that if $\delta'(\mu, \nu) = 0$ then the result is trivial, and hence we take $\delta'(\mu, \nu) > 0$ in what follows.

Let $\epsilon > 0$ be given. By axiom 3.3, there exists a coupling (X, Y) of (μ, ν) such that

$$\mathbb{P}\{X \preceq Y\} \geq \delta'(\mu, \nu) - \epsilon. \tag{13}$$

Define $\lambda := \mathbb{P}\{X \preceq Y\}$ and the measures

$$\mu'(B) = \frac{\mathbb{P}\{X \in B,\ X \preceq Y\}}{\lambda}, \quad \nu'(B) = \frac{\mathbb{P}\{Y \in B,\ X \preceq Y\}}{\lambda},$$
$$\mu''(B) = \frac{\mathbb{P}\{X \in B,\ X \npreceq Y\}}{1 - \lambda}, \quad \nu''(B) = \frac{\mathbb{P}\{Y \in B,\ X \npreceq Y\}}{1 - \lambda}.$$

It is straightforward to check that these are indeed probability measures. For example, the definition of λ gives $\mu'(S) = \mathbb{P}\{X \in S,\ X \preceq Y\}/\lambda = \mathbb{P}\{X \preceq Y\}/\lambda = 1$. It is also true that $\mu' \preceq_{sd} \nu'$. To see this, let $I \in \mathcal{B}_I$ and observe that, since I is increasing,

$$\mu'(I) = \frac{\mathbb{P}\{X \in I,\ X \preceq Y\}}{\lambda} \leq \frac{\mathbb{P}\{Y \in I,\ X \preceq Y\}}{\lambda} = \nu'(I).$$

Finally, expression (8) holds true. Indeed for μ we have $\mu(B) = \mathbb{P}\{X \in B\}$, which can be decomposed as

$$\mathbb{P}\{X \in B,\ X \preceq Y\} + \mathbb{P}\{X \in B,\ X \npreceq Y\} = \lambda\mu'(B) + (1-\lambda)\mu''(B).$$

Since δ satisfies axiom 3.2, it now follows that $\delta(\mu,\nu) \geq \lambda$. Using the definition of λ and (13) we have $\delta(\mu,\nu) \geq \delta'(\mu,\nu) - \epsilon$. As ϵ was arbitrary we now have $\delta(\mu,\nu) \geq \delta'(\mu,\nu)$, completing the proof of equality. □

During the next proof we require some additional notation. For our partial order $\preceq$ on S, let $\mathbb{G}$ be the graph. That is $\mathbb{G} := \{(x,y) \in S \times S : x \preceq y\}$. Let $\mathcal{B} \otimes \mathcal{B}$ be the product σ-algebra on $S \times S$. Let π_i be the i-th coordinate projection, so that $\pi_1(x,y) = x$ and $\pi_2(x,y) = y$ for any $(x,y) \in S \times S$. As usual, given $Q \subset S \times S$, we let $\pi_1(Q)$ be all $x \in S$ such that $(x,y) \in Q$ for some $y \in S$, and similarly for π_2. Let $\mathcal{C}$ be the closed sets in S and let $\mathcal{C}_D$ be the decreasing sets in $\mathcal{C}$.

Proof of theorem 4.1. It is convenient if we first establish that ρ satisfies axiom 3.3 and then return to axiom 3.2. Fix $\mu,\nu \in \mathcal{P}$. We start by establishing the following inequalities:

$$(14) \qquad 1 - \sup_{(X,Y)\in\Gamma(\mu,\nu)} \mathbb{P}\{X \preceq Y\} \leq \sup_{I \in \mathcal{O}_I}\{\mu(I) - \nu(I)\} \leq \sup_{I \in \mathcal{B}_I}\{\mu(I) - \nu(I)\}.$$

The second inequality is obvious and stated only for convenience. The first inequality is subtle, and also tighter than necessary for the current result, but useful in the sequel. It relies on Strassen's theorem (Strassen (1965)). In the present context, Strassen's theorem tells us that, for any $\epsilon \geq 0$ and any closed set $K \subset S \times S$, there exists a probability measure λ on $(S \times S, \mathcal{B} \otimes \mathcal{B})$ with marginals μ and ν such that $\lambda(K) \geq 1 - \epsilon$ if and only if

$$\nu(F) \leq \mu(\pi_1(K \cap (S \times F))) + \epsilon, \qquad \forall\, F \in \mathcal{C}.$$

For any $F \subset S$, let F^d be the smallest decreasing set containing F. In other words, F^d is all $x \in S$ such that $x \preceq y$ for some $y \in F$. Note that if F is closed then, since $\preceq$ is a closed partial order, so is the set F^d (i.e., $F \in \mathcal{C} \implies F^d \in \mathcal{C}_D$). Let $\epsilon := \sup_{D \in \mathcal{C}_D} \{\nu(D) - \mu(D)\}$. Evidently

$$\epsilon \geq \sup_{F \in \mathcal{C}} \{\nu(F^d) - \mu(F^d)\} \geq \sup_{F \in \mathcal{C}} \{\nu(F) - \mu(F^d)\}.$$

Noting that F^d can be expressed as $\pi_1(\mathbb{G} \cap (S \times F))$, it follows that, for any $F \in \mathcal{C}$,

$$\nu(F) \leq \mu(\pi_1(\mathbb{G} \cap (S \times F))) + \epsilon.$$

Since $\preceq$ is a closed partial order, the set $\mathbb{G}$ is closed, and Strassen's theorem applies. From this theorem we obtain a probability measure λ on the product space $S \times S$ such that (a) $\lambda(\mathbb{G}) \geq 1 - \epsilon$ and (b) λ has marginals μ and ν.

Because complements of increasing sets are decreasing and vice versa, we have

$$\sup_{I \in \mathcal{O}_I} \{\mu(I) - \nu(I)\} = \sup_{D \in \mathcal{C}_D} \{\nu(D) - \mu(D)\} = \epsilon \geq 1 - \lambda(\mathbb{G}). \tag{15}$$

Now consider the probability space $(\Omega, \mathcal{F}, \mathbb{P}) = (S \times S, \mathcal{B} \otimes \mathcal{B}, \lambda)$, and let $X = \pi_1$ and $Y = \pi_2$. We then have

$$\lambda(\mathbb{G}) = \lambda\{(x, y) \in S \times S : x \preceq y\} = \mathbb{P}\{X \preceq Y\}.$$

Combining this equality with (15), we have shown the existence of a $(X, Y) \in \Gamma(\mu, \nu)$ with $\sup_{I \in \mathcal{O}_I} \{\mu(I) - \nu(I)\} \geq 1 - \mathbb{P}\{X \preceq Y\}$. Inequality (14) now follows.

Returning to axiom 3.3, from (14) we easily obtain

$$\sup_{h \in \mathcal{H}} \left\{ \int h d\mu - \int h d\nu \right\} \geq 1 - \sup_{(X,Y) \in \Gamma(\mu,\nu)} \mathbb{P}\{X \preceq Y\}, \tag{16}$$

or, equivalently, $\rho(\mu, \nu) \leq \sup_{(X,Y) \in \Gamma(\mu,\nu)} \mathbb{P}\{X \preceq Y\}$. Hence axiom 3.3 holds.

The proof that ρ satisfies axiom 3.2 runs as follows. Fix $\mu, \nu \in \mathcal{P}$ and let the decompositions in (8) be given, with $\lambda \in [0, 1]$ and $\mu' \preceq_{sd} \nu'$. Pick any $h \in \mathcal{H}$. We have

$$\int h \, d\mu - \int h \, d\nu = \lambda \int h \, d\mu' + (1 - \lambda) \int h \, d\mu'' - \lambda \int h \, d\nu' - (1 - \lambda) \int h \, d\nu''.$$

Using $0 \leq h \leq 1$ we have $(1 - \lambda) \int h \, d\mu'' - (1 - \lambda) \int h \, d\nu'' \leq 1 - \lambda$, and hence

$$\int h \, d\mu - \int h \, d\nu \leq \lambda[\mu' - \nu'] + 1 - \lambda.$$

Since $\mu' \preceq_{sd} \nu'$ and $h \in \mathcal{H}$ we obtain $\int h \, d\mu - \int h \, d\nu \leq 1 - \lambda$. Taking the supremum and rearranging gives $\rho(\mu, \nu) = 1 - \sup_{h \in \mathcal{H}} \{\int h \, d\mu - \int h \, d\nu\} \geq \lambda$. In other words, axiom 3.2 holds. $\square$

Proof of proposition 4.1. After taking suprema, lemma 6.1 tells us that

$$\sup_{h \in \mathcal{H}} \left\{ \int h d\mu - \int h d\nu \right\} \leq 1 - \sup_{(X,Y) \in \Gamma(\mu,\nu)} \mathbb{P}\{X \preceq Y\}.$$

Combining this with (14), we see that all the terms in (14) are in fact equal, and take the common value $\sup_{h \in \mathcal{H}} \left\{ \int h d\mu - \int h d\nu \right\}$. This confirms the claim in proposition 4.1. □

Proof of proposition 4.2. Let μ, ν, μ' and ν' be as in the statement of the proposition. From proposition 4.1 we have

$$\begin{aligned} \left|\rho(\mu',\nu') - \rho(\mu,\nu)\right| &= \left| 1 - \sup_{I \in \mathcal{B}_I} \{\mu'(I) - \nu'(I)\} - 1 + \sup_{I \in \mathcal{B}_I} \{\mu(I) - \nu(I)\} \right| \\ &= \left| \sup_{I \in \mathcal{B}_I} \{\mu'(I) - \nu'(I)\} - \sup_{I \in \mathcal{B}_I} \{\mu(I) - \nu(I)\} \right| \\ &\leq \sup_{I \in \mathcal{B}_I} \left| \mu'(I) - \nu'(I) - \{\mu(I) - \nu(I)\} \right| \\ &\leq \sup_{B \in \mathcal{B}} \left| \mu'(B) - \mu(B) \right| + \sup_{B \in \mathcal{B}} \left| \nu'(B) - \nu(B) \right|. \end{aligned}$$

This is equivalent to the statement of proposition 4.2. □

References

Acemoglu, Daron and Martin Kaae Jensen (2012) "Robust comparative statics in large dynamic economies,"Technical report, National Bureau of Economic Research.

Atkinson, Anthony B (1987) "On the measurement of poverty," *Econometrica*, Vol. 55, pp. 749–764.

Balbus, Lukasz, Paweł Dziewulski, Kevin Reffett, and Lukasz Wozny (2012) "Equilibria in Large Games with Strategic Complementarities," *Available at SSRN 2020979*.

Border, Kim C (1992) "Revealed preference, stochastic dominance, and the expected utility hypothesis," *Journal of Economic Theory*, Vol. 56, pp. 20–42.

Davidson, Russell and Jean-Yves Duclos (2000) "Statistical inference for stochastic dominance and for the measurement of poverty and inequality," *Econometrica*, Vol. 68, pp. 1435–1464.

——— (2013) "Testing for restricted stochastic dominance," *Econometric Reviews*, Vol. 32, pp. 84–125.

Denuit, Michel, Rachel Huang, Larry Tzeng et al. (2013) "Bivariate Almost Stochastic Dominance,"Technical report, UCL.

Fields, Gary S, Jesse B Leary, and Efe A Ok (2002) "Stochastic dominance in mobility analysis," *Economics Letters*, Vol. 75, pp. 333–339.

Föllmer, H. and A. Schied (2011) *Stochastic Finance: An Introduction in Discrete Time*, De Gruyter Textbook Series: De Gruyter.

Hadar, Josef and William R Russell (1971) "Stochastic dominance and diversification," *Journal of Economic Theory*, Vol. 3, pp. 288–305.

Jarrow, Robert (1986) "The relationship between arbitrage and first order stochastic dominance," *The Journal of Finance*, Vol. 41, pp. 915–921.

Kamae, Teturo and Ulrich Krengel (1978) "Stochastic partial ordering," *The Annals of Probability*, Vol. 6, pp. 1044–1049.

Kamae, Teturo, Ulrich Krengel, and George L O'Brien (1977) "Stochastic inequalities on partially ordered spaces," *The Annals of Probability*, pp. 899–912.

Lehmann, Erich L (1955) "Ordered families of distributions," *The Annals of Mathematical Statistics*, pp. 399–419.

Leshno, Moshe and Haim Levy (2002) "Preferred by all and preferred by most decision makers: Almost stochastic dominance," *Management Science*, Vol. 48, pp. 1074–1085.

Levy, Haim (1992) "Stochastic dominance and expected utility: survey and analysis," *Management Science*, Vol. 38, pp. 555–593.

Levy, Moshe (2009) "Almost stochastic dominance and stocks for the long run," *European Journal of Operational Research*, Vol. 194, pp. 250–257.

Lindvall, Torgny (2002) *Lectures on the coupling method*: Courier Dover Publications.

Lizyayev, Andrey and Andrzej Ruszczyński (2012) "Tractable almost stochastic dominance," *European Journal of Operational Research*, Vol. 218, pp. 448–455.

Machina, Mark J (1982) "'Expected utility' analysis without the independence axiom," *Econometrica*, pp. 277–323.

Mas-Colell, A., M.D. Whinston, and J.R. Green (1995) *Microeconomic Theory*: Oxford University Press.

Quirk, James P and Rubin Saposnik (1962) "Admissibility and measurable utility functions," *The Review of Economic Studies*, pp. 140–146.

Stokey, Nancy and Robert E Lucas (1989) *Recursive Methods in Economic Dynamics (with EC Prescott)*: Harvard University Press.

Stoyanov, Stoyan V, Svetlozar T Rachev, and Frank J Fabozzi (2012) "Metrization of stochastic dominance rules," *International Journal of Theoretical and Applied Finance*, Vol. 15.

Strassen, Volker (1965) "The existence of probability measures with given marginals," *The Annals of Mathematical Statistics*, pp. 423–439.

Tzeng, Larry Y, Rachel J Huang, and Pai-Ta Shih (2013) "Revisiting almost second-degree stochastic dominance," *Management Science*, Vol. 59, pp. 1250–1254.