A new characterization of second-order stochastic dominance

Yuanying Guan* Muqiao Huang† Ruodu Wang‡

September 30, 2024

Abstract

We provide a new characterization of second-order stochastic dominance, also known as increasing concave order. The result has an intuitive interpretation that adding a risk with negative expected value in adverse scenarios makes the resulting position generally less desirable for risk-averse agents. A similar characterization is also found for convex order and increasing convex order. The proof techniques for the main result are based on properties of Expected Shortfall, a family of risk measures that is popular in banking and insurance regulation. Applications in risk management and insurance are discussed.

Keywords: Expected Shortfall, stochastic dominance, convex order, dependence, Strassen's theorem

1 Introduction

Second-order stochastic dominance (SSD), also known as increasing concave order, is one of the most fundamental tools in stochastic comparison and decision making under risk (e.g., Hadar and Russell (1969) and Rothschild and Stiglitz (1970)). For general treatments, we refer to the monographs Müller and Stoyan (2002) and Shaked and Shanthikumar (2007).

*Department of Mathematical Sciences and Department of Finance & Real Estate, DePaul University, USA. E-mail: yguan8@depaul.edu.

†Department of Statistics and Actuarial Science, University of Waterloo, Canada. E-mail: m5huang@uwaterloo.ca.

‡Department of Statistics and Actuarial Science, University of Waterloo, Canada. E-mail: wang@uwaterloo.ca.

This short paper provides a new characterization of SSD. Let L^1 be the set of integrable random variables in an atomless probability space $(\Omega, \mathcal{F}, \mathbb{P})$, which we fix throughout. We first give the standard definitions for some stochastic orders. For $X, Y \in L^1$, we say that X dominates Y

(a) in SSD, denoted by $X \geq_{\mathrm{ssd}} Y$, if $\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$ for all increasing concave functions u;

(b) in increasing convex order, denoted by $X \geq_{\mathrm{icx}} Y$, if $\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$ for all increasing convex functions u;

(c) in convex order, denoted by $X \geq_{\mathrm{cx}} Y$, if $\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$ for all convex functions u.

Throughout the paper, "increasing" is in the non-strict sense.

In decision theory, X and Y in comparison are usually interpreted as random payoffs or wealths. Instead, by interpreting X and Y as losses, SSD can be converted into increasing convex order, since $X \leq_{\mathrm{ssd}} Y$ is equivalent to $-X \geq_{\mathrm{icx}} -Y$. Increasing convex order is also known as stop-loss order in actuarial science; see e.g., Dhaene et al. (2002). We write $X \stackrel{\mathrm{d}}{=} Y$ if X and Y are identically distributed, which is precisely the symmetric part of each relation above.

Classic results on the representation of SSD are obtained by the celebrated work of Strassen (1965) and Rothschild and Stiglitz (1970). This representation result can be summarized as follows: For $X, Y \in L^1$, $X \geq_{\mathrm{ssd}} Y$ holds if and only if $Y \stackrel{\mathrm{d}}{=} W + Z$ for some $W, Z \in L^1$ such that $W \stackrel{\mathrm{d}}{=} X$ and

$$\mathbb{E}[Z|W] \leq 0. \tag{1}$$

See Theorem 4.A.5 of Shaked and Shanthikumar (2007) for this result. Condition (1) means that $(W, W+Z)$ forms a supermartingale.

The main result of this paper is to provide a different representation of SSD, where the corresponding condition for the additive payoff is weaker than (1).

Theorem 1. *For any $X, Y \in L^1$, $X \geq_{\mathrm{ssd}} Y$ holds if and only if $Y \stackrel{\mathrm{d}}{=} W + Z$ for some $W, Z \in L^1$ such that $W \stackrel{\mathrm{d}}{=} X$ and*

$$\mathbb{E}[Z|W \leq x] \leq 0 \text{ for all relevant values of } x. \tag{2}$$

In (2), relevant values of x are those that satisfy $\mathbb{P}(W \leq x) > 0$.

Note that (2) implies $\mathbb{E}[Z] \leq 0$ by taking $x \to \infty$. Condition (2) is clearly weaker than (1), because the latter can be equivalently written as $\mathbb{E}[Z|W = x] \leq$

0 for all almost every x in the range of W (here, the conditional expectations are chosen as a regular version). If Z is a function of W, then (1) is very restrictive, as it means $Z \leq 0$, whereas (2) can hold for a wide range of models that does not require $Z \leq 0$. Moreover, (2) is much easier to check in practice, since the event $\{W \leq x\}$ has a positive probability for every relevant x, whereas the event $\{W = x\}$ has zero probability for all x when W is continuously distributed. Two examples comparing (1) and (2) are presented in Section 4. In Section 5, we discuss applications of the new condition to risk management and insurance, including stochastic improvers, marketable insurance contracts, and stop-loss premium calculation.

The main interpretation of Theorem 1 is that, for a risk-averse decision maker with random wealth W, adding a risk Z with negative expectation in adverse scenarios, that is, when W is small, makes the resulting position $W+Z$ generally less desirable than W; see Section 6 for more discussions.

To prove Theorem 1, the main step is to justify $W + Z \leq_{\rm ssd} W$, which we summarize in the following proposition.

Proposition 2. *For any $W, Z \in L^1$ satisfying (2), $W + Z \leq_{\rm ssd} W$ holds.*

Up to the best of our knowledge, both Theorem 1 and Proposition 2 are new to the literature.

Proposition 2 is closely related to the main result of Brown (2017), where the author showed that for W taking values in $[0, 1]$ and Z taking values in $[-1, 1]$, if $\mathbb{E}[Z] = 0$ and

$$\mathbb{E}[Z|W \geq x] \geq 0 \text{ for all relevant values of } x, \tag{3}$$

then $W + Z \geq_{\rm cx} X$. This result also follows from Corollary 3.3 of Li et al. (2016). Despite the close connection, there are several additional merits of our results and the proof approach. First, our result works for both SSD and convex order, whereas the condition (3) of Brown (2017) does not generalize to SSD. Indeed, Brown (2017, Corollary 1) claimed that (3) together with $\mathbb{E}[Z] \leq 0$ yields $X + Z \leq_{\rm ssd} X$, but $\mathbb{E}[Z] < 0$ is not possible if (3) holds; thus SSD is not covered except for the case of convex order. Similarly, results in Li et al. (2016) rely on the notion of expectation dependence (Wright (1987); see Section 6), but our condition (2) is different from expectation dependence unless $\mathbb{E}[Z] = 0$. Second, our proof techniques are completely different from those of Brown (2017). Our proof is much shorter, and it is based on risk measures, in particular, the Expected Shortfall (ES, also known as CVaR or TVaR), one of the most important risk measures in finance and insurance (McNeil et al.

(2015)). Thus, the proof is more accessible to scholars in risk management. Third, our result is formulated on L^1 without any restriction on the range of the random variables, and the proof argument is unified for all random variables without using discrete approximation or taking limits.

2 Proof of the main result

Let us first define the risk measure ES used in our proof. ES at level $p \in [0, 1)$ is defined by

$$\mathrm{ES}_p(X) = \frac{1}{1-p}\int_p^1 Q_X(t)\mathrm{d}t, \quad X \in L^1,$$

where

$$Q_X(t) = \inf\{x \in \mathbb{R} : \mathbb{P}(X \le x) > t\}$$

is the right t-quantile of X at $t \in (0, 1)$. For $X \in L^1$, denote by ϕ_X the function on $[0, 1]$ given by $\phi_X(p) = (1-p)\mathrm{ES}_p(X)$ on $[0, 1)$ and $\phi_X(1) = 0$. We first state a few simple facts on ES and the quantile function in the following lemma.

Lemma 3. *For $X, Y \in L^1$, the following statements hold:*

(i) $X \ge_{\mathrm{icx}} Y$ *if and only if* $\mathrm{ES}_p(X) \ge \mathrm{ES}_p(Y)$ *for all* $p \in (0, 1)$*;*

(ii) for $p \in (0, 1)$, $\mathrm{ES}_p(X) \ge \mathbb{E}[X|B]$ *for any* $B \in \mathcal{F}$ *with* $\mathbb{P}(B) = 1 - p$*;*

(iii) for $p \in (0, 1)$*, if* $\mathbb{P}(X < Q_X(p)) = p$*, then* $\mathrm{ES}_p(X) = \mathbb{E}[X|X \ge Q_X(p)]$*;*

(iv) for $p \in (0, 1)$*, if* $\mathbb{P}(X < Q_X(q)) < \mathbb{P}(X < Q_X(p))$ *for all* $q \in (0, p)$*, then* $\mathbb{P}(X < Q_X(p)) = p$*;*

(v) the function ϕ_X *is continuous and concave on* $[0, 1]$*, and its derivative is* $-Q_X(p)$ *at almost every* $p \in (0, 1)$.

Proof of Lemma 3. These properties are all well-known or straightforward to check. For (i), see Theorem 4.A.3 of Shaked and Shanthikumar (2007). For (ii), see Lemma 3.1 of Embrechts and Wang (2015). For (iii), see Lemma A.7 of Wang and Zitikis (2021). Statement (iv) follows from Lemma 1 of Guan et al. (2024). Statement (v) follows directly from the definition of ES. □

Proof of Proposition 2. Note that $W + Z \le_{\mathrm{ssd}} W$ is equivalent to $-W - Z \ge_{\mathrm{icx}} -W$, which we show below. Write $X = -W$. By (i) of Lemma 3, it suffices to show $\mathrm{ES}_p(X - Z) \geq \mathrm{ES}_p(X)$ for all $p \in (0,1)$. Denote by $P_X = \{p \in (0,1) : \mathbb{P}(X < Q_p(X)) = p\}$. For $p \in P_X$, we have

$$\begin{aligned}
\mathrm{ES}_p(X - Z) &\geq \mathbb{E}[X - Z | X \geq Q_X(p)] && \text{[by (ii)]} \\
&= \mathrm{ES}_p(X) - \mathbb{E}[Z | X \geq Q_X(p)] && \text{[by (iii)]} \\
&\geq \mathrm{ES}_p(X). && \text{[by (2)]}
\end{aligned}$$

The argument is complete here if X is continuously distributed, as in that case $P_X = (0,1)$. We continue with the case that the distribution of X may have atoms.

Suppose that an interval $(a,b) \subseteq (0,1)$ does not intersect P_X (such intervals may not exist). For any $p \in (a,b)$, let $p^* = \inf\{q \in (0,1) : Q_X(q) = Q_X(p)\}$. If $p^* = 0$, then $q \mapsto Q_X(q)$ is constant on $(0,p)$, and since $p \in (a,b)$ is arbitrary, we have that $q \mapsto Q_X(q)$ is constant on $(0,b)$. Next suppose $p^* > 0$. Since $q \mapsto Q_X(q)$ is right-continuous, we have $Q_X(p^*) = Q_X(p)$. The definition of p^* implies $Q_X(q) < Q_X(p^*)$ for any $q \in (0, p^*)$. By (iv), $\mathbb{P}(X < Q_X(p^*)) = p^*$ and hence $p^* \in P_X$. This yields $p^* \leq a$. Therefore, $q \mapsto Q_X(q)$ is constant on the interval (a,b) in both cases.

By (v), ϕ_X is linear on any interval outside P_X and ϕ_{X-Z} is concave, and both are continuous. We claim that these properties and $\phi_{X-Z} \geq \phi_X$ on P_X imply that $\phi_{X-Z} \geq \phi_X$ holds on $(0,1)$; see Figure 1 for an illustration. This would be sufficient for $X - Z \geq_{\mathrm{icx}} X$, and below we show this claim.

Suppose that there exists $p \in (0,1)$ such that $\phi_{X-Z}(p) < \phi_X(p)$. Since both ϕ_{X-Z} and ϕ_X are continuous, there exists a neighbourhood of p on which $\phi_{X-Z} < \phi_X$. Let $a = \inf\{q \in (0,1) : \phi_{X-Z} < \phi_X \text{ on } (q,p]\}$ and $b = \sup\{q \in (0,1) : \phi_{X-Z} < \phi_X \text{ on } [p,q)\}$. If $a > 0$, then by continuity we have $\phi_{X-Z}(a) = \phi_X(a)$ and $\phi_{X-Z}(b) = \phi_X(b)$, noting that $\phi_{X-Z}(1) = \phi_X(1) = 0$. If $a = 0$, then $\phi_{X-Z}(a) \geq \phi_X(a)$ because

$$\phi_{X-Z}(0) = \mathbb{E}[X] - \mathbb{E}[Z] \geq \mathbb{E}[X] = \phi_X(0),$$

where $\mathbb{E}[Z] \leq 0$ is guaranteed by (2). Therefore, in either case,

$$\phi_{X-Z}(a) \geq \phi_X(a) \quad \text{and} \quad \phi_{X-Z}(b) \geq \phi_X(b). \tag{4}$$

Note that (a,b) does not intersect P_X since $\phi_{X-Z} \geq \phi_X$ on P_X. Since ϕ_{X-Z} is concave and ϕ_X is linear on (a,b), (4) implies $\phi_{X-Z} \geq \phi_X$ on (a,b); see the areas with dashed lines in Figure 1. This yields a contradiction. □

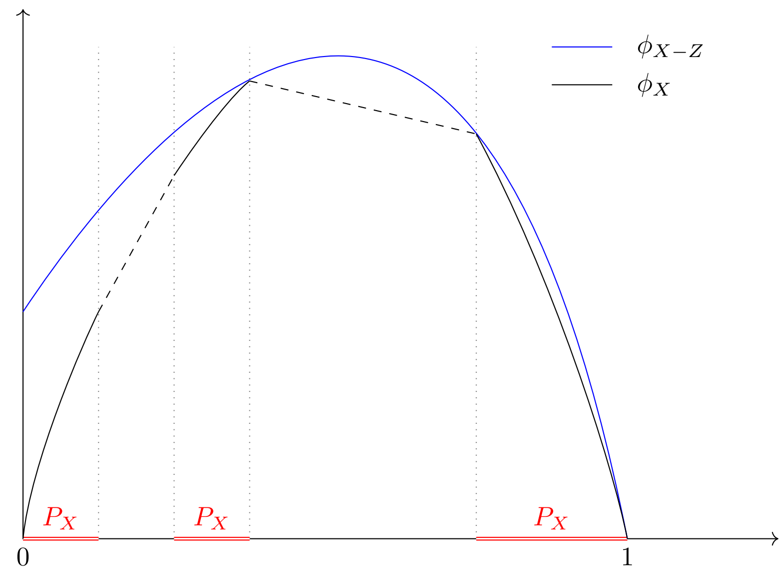

Figure 1: An illustration of $\phi_{X-Z} \geq \phi_X$: after the inequality is shown to hold on P_X, it also holds outside P_X due to concavity of ϕ_{X-Z} and piece-wise linearity of ϕ_X (dashed lines).

Proof of Theorem 1. The "if" statement follows from Proposition 2 via $X \stackrel{\mathrm{d}}{=} W \geq_{\mathrm{ssd}} W + Z \stackrel{\mathrm{d}}{=} Y$. The "only if" statement follows from the fact that (1) is stronger than (2), and thus via the representation mentioned in the introduction (Theorem 4.A.5 of Shaked and Shanthikumar (2007)). $\square$

3 Convex order and increasing convex order

We present a few immediate corollaries of Theorem 1 on increasing convex order and convex order, commonly used in risk management and actuarial science.

Corollary 4. *For any $X, Y \in L^1$, $X \leq_{\mathrm{icx}} Y$ holds if and only if $Y \stackrel{\mathrm{d}}{=} W + Z$*

for some $W, Z \in L^1$ *such that* $W \stackrel{\mathrm{d}}{=} X$ *and*

$$\mathbb{E}[Z|W \geq x] \geq 0 \text{ for all relevant values of } x.$$

Proof. This corollary follows by applying Theorem 1 to the relation $-X \geq_{\mathrm{ssd}} -Y$, and let $-Y \stackrel{\mathrm{d}}{=} -W - Z$ with $\mathbb{E}[-Z| - W \leq -x] \leq 0$ for all relevant values of x. □

Corollary 5. *For any* $X, Y \in L^1$*, the following are equivalent.*

(i) $X \leq_{\mathrm{cx}} Y$;

(ii) $X \geq_{\mathrm{ssd}} Y$ *and* $\mathbb{E}[X] = \mathbb{E}[Y]$;

(iii) $Y \stackrel{\mathrm{d}}{=} W + Z$ *for some* $W, Z \in L^1$ *such that* $W \stackrel{\mathrm{d}}{=} X$ *and* $\mathbb{E}[Z|W] = 0$.

(iv) $Y \stackrel{\mathrm{d}}{=} W + Z$ *for some* $W, Z \in L^1$ *such that* $W \stackrel{\mathrm{d}}{=} X$ *and*

$$\mathbb{E}[Z] = 0 \text{ and } \mathbb{E}[Z|W \leq x] \leq 0 \text{ for all relevant values of } x;$$

(v) $Y \stackrel{\mathrm{d}}{=} W + Z$ *for some* $W, Z \in L^1$ *such that* $W \stackrel{\mathrm{d}}{=} X$ *and*

$$\mathbb{E}[Z] = 0 \text{ and } \mathbb{E}[Z|W \geq x] \geq 0 \text{ for all relevant values of } x.$$

Proof. The equivalence between (i), (ii) and (iii) is well known; see e.g., Shaked and Shanthikumar (2007, Theorems 3.A.4 and 4.A.35). The equivalence between (ii) and (iv) follows from Theorem 1. The equivalence between (iv) and (v) follows by noting that for Z with $\mathbb{E}[Z] = 0$,

$$\mathbb{E}[Z|W \leq x] \leq 0 \ \forall x \iff \mathbb{E}[Z|W > x] \geq 0 \ \forall x \iff \mathbb{E}[Z|W \geq x] \geq 0 \ \forall x,$$

where the last equivalence is argued by $\lim_{x \uparrow y}\{W > x\} = \{W \geq y\}$ and $\lim_{x \downarrow y}\{W \geq x\} = \{W > y\}$. □

The implication (v) ⇒ (i) in Corollary 5 is also obtained by Li et al. (2016) and Brown (2017). To compare (iv) and (v) with the classic characterization (iii), if Z is a function of W, $\mathbb{E}[Z|W] = 0$ does not hold except for the trivial case $Z = 0$, whereas $\mathbb{E}[Z|W \leq x] \leq 0$ and $\mathbb{E}[Z|W \geq x] \geq 0$ in (iv) and (v) can hold for many models of Z, thus allowing for flexibility in applications.

4 Two illustrative examples

We now illustrate our results with two simple examples, one with Gaussian distributions and one with Bernoulli distributions. The purpose here is to compare the classic condition (1) with our condition (2), and to see how much more flexibility (2) offers. Note that both (1) and (2) are sufficient conditions for $W + Z \leq_{\text{ssd}} W$, but they may not be necessary. In both examples below, (2) allows a larger range of parameters than (1).

Example 1 (Gaussian distributions)**.** It is well known that for two normal random variables $X \sim \mathrm{N}(\mu_X, \sigma_X^2)$ and $Y \sim \mathrm{N}(\mu_Y, \sigma_Y^2)$, $X \geq_{\text{ssd}} Y$ holds if and only if

$$\mu_X \geq \mu_Y \quad \text{and} \quad \sigma_X^2 \leq \sigma_Y^2 \tag{5}$$

see e.g., Example 4.A.46 of Shaked and Shanthikumar (2007). Suppose that (W, Z) is jointly Gaussian with mean vector (μ_W, μ_Z) and covariance matrix

$$\begin{pmatrix} \sigma_W^2 & \rho\sigma_W\sigma_Z \\ \rho\sigma_W\sigma_Z & \sigma_Z^2 \end{pmatrix},$$

where ρ is the correlation coefficient. Since SSD is invariant up to location-scale transforms, we assume $\mu_W = 0$ and $\sigma_W = 1$ without loss of generality. We analyze values of the parameters $\mu_Z \in \mathbb{R}$, $\sigma_Z > 0$ and $\rho \in [-1, 1]$ obtained from $W + Z \leq_{\text{ssd}} W$, (1), and (2), respectively.

(a) Since $\mathbb{E}[W + Z] = \mu_Z$ and $\mathrm{var}(W + Z) = 1 + \sigma_Z^2 + 2\rho\sigma_Z$, we have from (5) that $W + Z \leq_{\text{ssd}} W$ if and only if $\mu_Z \leq 0$ and $\rho \geq -\sigma_Z/2$.

(b) Using the conditional distribution of the bivariate Gaussian distribution, we have $\mathbb{E}[Z|W] = \mu_Z + \rho\sigma_Z W$. Condition (1) is $\mu_Z + \rho\sigma_Z W \leq 0$. Since the support of W is the real line, this means $\mu_Z \leq 0$ and $\rho = 0$.

(c) Condition (2) is $\mu_Z + \rho\sigma_Z \mathbb{E}[W|W \leq x] \leq 0$ for all $x \in \mathbb{R}$. If $\rho \geq 0$, this holds true if and only if $\mu_Z \leq 0$, because $\sup_{x\in\mathbb{R}} \mathbb{E}[W|W \leq x] = \mathbb{E}[W] = 0$. If $\rho < 0$, this does not hold for any μ_Z and σ_Z, since $\mathbb{E}[W|W \leq x]$ is unbounded from below.

In this example, condition (2) is more flexible than (1), although it is stronger than the equivalent condition for SSD. Table 1 summarizes these observations.

Example 2 (Bernoulli distributions)**.** Consider a Bernoulli random variable W with parameter $1/2$ and a random variable Z distributed as $W - c$ for a

sufficient condition (1)	$\mu_Z \le 0$ and $\rho = 0$
our sufficient condition (2)	$\mu_Z \le 0$ and $\rho \ge 0$
equivalent condition for SSD	$\mu_Z \le 0$ and $\rho \ge -\sigma_Z/2$

Table 1: Conditions for $W + Z \le_{\rm ssd} W$ in Example 1 (Gaussian)

sufficient condition (1)	$c \ge 1/2$ and $1 - 2c \le \rho \le 2c - 1$
our sufficient condition (2)	$c \ge 1/2$ and $1 - 2c \le \rho$
equivalent condition for SSD	$c \ge 1/2$ and $1 - 2c \le \rho$

Table 2: Conditions for $W + Z \le_{\rm ssd} W$ in Example 2 (Bernoulli)

constant $c \in \mathbb{R}$. Let ρ be the correlation coefficient of (W, Z). Note that ρ fully determines the joint distribution of (W, Z), where the only degree of freedom is $\mathbb{P}(W = Z + c = 1) = (1 + \rho)/4$. We analyze values of the parameters $c \in \mathbb{R}$ and $\rho \in [-1, 1]$ obtained from $W + Z \le_{\rm ssd} W$, (1), and (2), respectively.

(a) First, $\mathbb{P}(W + Z = 2 - c) = \mathbb{P}(W + Z = -c) = (1 + \rho)/4$ and $\mathbb{P}(W + Z = 1 - c) = (1 - \rho)/2$. For $W + Z \le_{\rm ssd} W$ to hold, it is necessary and sufficient that $\mathbb{E}[(W + Z) \wedge t] \le \mathbb{E}[W \wedge t]$ for all $t \in \mathbb{R}$, where $a \wedge b = \min\{a, b\}$. Since $W + Z$ takes values only at three points $-c, 1 - c, 2 - c$, it suffices to check these three points. Checking the point $t = 2 - c$ yields $c \ge 1/2$. Checking the point $t = 1 - c$ yields $\rho \ge 1 - 2c$. Checking the point $t = -c$ yields $c \ge 0$. Therefore, the equivalent condition is $c \ge 1/2$ and $\rho \ge 1 - 2c$.

(b) We can directly compute $\mathbb{E}[Z|W] = (1-\rho)/2-c+\rho W$. Therefore, condition (1) means $(1 - \rho)/2 - c \le 0$ if $\rho \ge 0$ and $(1 - \rho)/2 - c + \rho \le 0$ if $\rho < 0$. Putting the two cases together, it is $1 - 2c \le \rho \le 2c - 1$ (which implies $c \ge 1/2$).

(c) Note that $\mathbb{E}[W|W \le x]$ for relevant x takes value 0 or $1/2$. Hence, condition (2) is $(1 - \rho)/2 - c \le 0$ and $(1 - \rho)/2 - c + \rho/2 \le 0$, and thus $c \ge 1/2$ and $\rho \ge 1 - 2c$.

In this example, condition (2) is more flexible than (1), and it turns out to be necessary and sufficient. Table 2 summarizes these observations.

In both examples, the condition via (2) and the equivalent SSD condition are both quite intuitive: the mean of Z is less than 0 (this is necessary for the

SSD relation), and the correlation of (W, Z) cannot be too small, as negative correlation reduces the aggregate risk. The condition via (1) does not have this interpretation.

For a bi-atomic distribution of Z on two arbitrary points (more general than Z in Example 2), we can check that condition (2) is equivalent to $Z+W \leq_{\mathrm{ssd}} W$ when the spread of Z is less than or equal to 1, and otherwise it is stronger than necessary. We omit these calculations.

5 Risk management and insurance applications

In this section, we provide three applications of our main results: stochastic improvers, insurance marketability, and stop-loss premium calculation.

5.1 Risk reducers and stochastic improvers

As a risk management tool, the concept of risk reducers is introduced by Cheung et al. (2014) and further studied by He et al. (2016). A *risk reducer* for $X \in L^1$ (Cheung et al. (2014)) is an additive payoff $Z \in L^1$ such that $X+Z \leq_{\mathrm{cx}} X+\mathbb{E}[Z]$. Intuitively, risk reducers are random payoffs Z that make the combined payoff $X+Z$ less risky than the original wealth X adjusted by the mean of Z.

Inspired by this, we define a *stochastic improver* for $X \in L^1$, that is, an additive payoff $Z \in L^1$ such that $X+Z \geq_{\mathrm{ssd}} X$. Recall that risk-averse expected utility agents are modelled by increasing concave utility functions. The intuition of a stochastic improver is that every risk-averse expected utility agent would prefer $X+Z$ over X, and thus the additive payoff Z improves the utility of the agent with random wealth X.

Note that $X+Z \geq_{\mathrm{ssd}} X+\mathbb{E}[Z]$ is equivalent to $X+Z \leq_{\mathrm{cx}} X+\mathbb{E}[Z]$. Hence, for Z with $\mathbb{E}[Z]=0$, it is a risk reducer if and only if it is a stochastic improver. However, for random variables with non-zero mean, these two concepts are generally incompatible. If $\mathbb{E}[Z] \geq 0$, then a risk reducer is necessarily a stochastic improver, because

$$X+Z \geq_{\mathrm{ssd}} X+\mathbb{E}[Z] \geq_{\mathrm{ssd}} X.$$

However, the converse is not true; a stochastic improver need not be a risk reducer. For instance, for any nonnegative $X \in L^2$ with positive variance, we have $X+X \geq_{\mathrm{ssd}} X$ but $\mathrm{var}(X+X) > \mathrm{var}(X+\mathbb{E}[X])$; therefore X is a stochastic improver for itself but not a risk reducer; in fact, $X+X \geq_{\mathrm{cx}} X+\mathbb{E}[X]$ holds (Theorem 3.A.17 of Shaked and Shanthikumar (2007)), the opposite of being a risk reducer. Moreover, a risk reducer Z is not a stochastic improver if $\mathbb{E}[Z]<0$.

Denote by $\mathcal{S}_X$ the set of all stochastic improvers for $X \in L^1$, and let

$$\mathcal{N}_X = \{Z \in L^1 : \mathbb{E}[Z|X+Z \le x] \ge 0 \text{ for all relevant values of } x\}.$$

The definition of SSD implies that the set $\mathcal{S}_X$ is convex. The following result connects the above two sets by using our main result.

Proposition 6. *For $X \in L^1$, $\mathcal{N}_X \subseteq \mathcal{S}_X$.*

Proof. It suffices to verify that for $Z \in \mathcal{N}_X$, $X + Z \ge_{\text{ssd}} X$. Let $W = X + Z$. Using Proposition 2, the condition $\mathbb{E}[-Z|W \le x] \le 0$ for all relevant x is sufficient for $W \ge_{\text{ssd}} W - Z$, which is $X + Z \ge_{\text{ssd}} X$. □

One may wonder whether the converse statement to Proposition 6 also holds, that is, $\mathcal{S}_X = \mathcal{N}_X$. The quick answer is negative. As we see in Table 1, for two standard Gaussian random variables W and Z, $W - Z \le_{\text{ssd}} W$ if and only if the correlation coefficient between W and Z is smaller than or equal to $1/2$. By writing $X = W - Z$, the above condition is equivalent to $Z \in \mathcal{S}_X$. On the other hand, $Z \in \mathcal{N}_X$ if and only if the correlation coefficient between W and Z is nonpositive. Therefore, $Z \in \mathcal{S}_X$ but $Z \notin \mathcal{N}_X$.

There is a special setting in which $\mathcal{S}_X$ and $\mathcal{N}_X$ coincide. A random vector (X, Y) is *comonotonic* if there exist increasing functions f and g such that $X = f(X + Y)$ and $Y = g(X + Y)$ almost surely. He et al. (2016) studied risk reducers Z when $(X, X+Z)$ is comonotonic. Below we obtain a characterization of stochastic improvers under the same assumption of comonotonicity.

Proposition 7. *Let $X, Z \in L^1$ be such that $(X, X + Z)$ is comonotonic. Then $Z \in \mathcal{N}_X$ if and only if $Z \in \mathcal{S}_X$.*

Proof. The "only-if" statement follows from Proposition 6. We show the "if" statement below. Suppose $Z \notin \mathcal{N}_X$. By definition of $\mathcal{N}_X$, there exists $x \in \mathbb{R}$ such that $\mathbb{E}[Z|X + Z \le x] < 0$ and $\mathbb{P}(X + Z \le x) > 0$. Let $A = \{X + Z \le x\}$ and $p = 1 - \mathbb{P}(A)$. Since $\mathbb{E}[Z] \ge 0$ as required by $Z \in \mathcal{S}_X$, we have $\mathbb{P}(A) \in (0, 1)$. It follows that

$$\mathbb{E}[X + Z|A] = \mathbb{E}[X|A] + \mathbb{E}[Z|A] < \mathbb{E}[X|A]. \tag{6}$$

Let $W = -X$. Since $(X, X + Z)$ is comonotonic, $(W, \mathbb{1}_A)$ is also comonotonic. Therefore, for almost every $\omega \in A$ and $\omega' \in A^c$, we have $W(\omega) \ge W(\omega')$. Such A is called a p-tail event of W by Wang and Zitikis (2021); intuitively, it is a set on which W takes larger values than on its complement. By definition, A is also a p-tail event of $-X-Z=W-Z$. Lemma A.7 of Wang and Zitikis (2021) gives $\mathbb{E}[W|A]=\mathrm{ES}_p(W)$ and $\mathbb{E}[W-Z|A]=\mathrm{ES}_p(W-Z)$. Putting the above observations with (6), we have $\mathrm{ES}_p(W-Z)>\mathrm{ES}_p(W)$, and hence by Lemma 3 part (i), $W-Z\le_{\mathrm{icx}} W$ cannot hold. This means that $X+Z\ge_{\mathrm{ssd}} X$ does not hold, and $Z\notin\mathcal{S}_X$. Thus, $Z\in\mathcal{S}_X$ implies $Z\in\mathcal{N}_X$. $\square$

Applying Proposition 7 to $Z-\mathbb{E}[Z]$ with $(X,X+Z)$ comonotonic, we obtain that the following conditions are equivalent:

(a) $Z-\mathbb{E}[Z]$ is a stochastic improver for X;

(b) Z is a risk reducer for X;

(c) $\mathbb{E}[Z|X+Z\le x]\ge\mathbb{E}[Z]$ for all relevant x.

This result can be compared with He et al. (2016, Theorem 3.2), which states that for Z that is $\sigma(X)$-measurable with $(X,X+Z)$ comonotonic, (b) is equivalent to

(d) $\mathbb{E}[Z|X\le x]\ge\mathbb{E}[Z]$ for all relevant x.

The condition (d) is called negative expectation dependence of Z on X (Wright (1987)). Generally, the two conditions (c) and (d) are not equivalent even if $(X,X+Z)$ is comonotonic; for instance, if X is a constant, then (d) always holds true but (c) never holds true unless Z is also a constant. Nevertheless, when Z is $\sigma(X)$-measurable, (d) implies (c) because the set of events $\{\{X\le x\}:x\in\mathbb{R}\}$ contains $\{\{X+Z\le y\}:y\in\mathbb{R}\}$ in this case. Therefore, Proposition 7 implies Theorem 3.2 of He et al. (2016) as a special case when Z is $\sigma(X)$-measurable and condition (d) holds.

An example of a stochastic improver satisfying the conditions in Proposition 7 is the purchase of a protective put in a Black–Scholes financial market. Below, a random vector (X,Y) is *counter-monotonic* if $(X,-Y)$ is comonotonic.

Example 3 (Protective put)**.** Consider a continuous-time financial market model with 0 interest rate on a time interval $[0,T]$. For simplicity, we will assume a Black–Scholes market with a stock price process $(X_t)_{t\in[0,T]}$ that has a nonpositive return rate and constant volatility. The assumption of nonpositive return rate is unusual, but it is needed for the analysis below. For details on the Black–Scholes market model used in this example, see Shreve (2004). This market is complete, and thus any payoff can be priced with a risk-neutral probability measure Q. For $t\in[0,T]$, let $\mathcal{F}_t=\sigma(X_s:s\le t)$ and P_t be the time-t price of a put option that gives the payoff $(K-X_T)_+$ at maturity T, where $K > 0$ represents the strike price. The strategy of holding the stock and purchasing the put option is called a protective put. The Black–Scholes formula gives that P_t is a decreasing function of X_t for each $t \in [0, T]$. Let $Z_t = P_t - P_0$, that is, the time-t value of purchasing the put option at time 0. Girsanov's theorem gives the explicit formula of $Y_t := \mathbb{E}[\mathrm{d}Q/\mathrm{d}\mathbb{P}|\mathcal{F}_t]$, which guarantees that Y_t is an increasing function of X_t under the assumption of nonpositive return rate. Hence,

$$P_0 = \mathbb{E}^Q[P_t] = \mathbb{E}\left[\mathbb{E}\left[\frac{\mathrm{d}Q}{\mathrm{d}\mathbb{P}}|\mathcal{F}_t\right] P_t\right] \leq \mathbb{E}\left[\frac{\mathrm{d}Q}{\mathrm{d}\mathbb{P}}\right] \mathbb{E}[P_t] = \mathbb{E}[P_t],$$

where the inequality is due to the counter-monotonicity of (P_t, Y_t). Moreover, $X_t + Z_t$ is an increasing function of X_t for $t \in [0, T]$, which can be seen from e.g., the put-call parity. Using counter-monotononicity of $(P_t, X_t + Z_t)$, which implies negative expectation dependence of P_t on $X_t + Z_t$, we have, for any relevant $x \in \mathbb{R}$,

$$\mathbb{E}[Z_t|X_t + Z_t \leq x] = \mathbb{E}[P_t - P_0|X_t + Z_t \leq x] \geq \mathbb{E}[P_t] - P_0 \geq 0.$$

Therefore, the conditions in Proposition 7 hold for (X_t, Z_t), and hence Z_t is a stochastic improver for X_t. In conclusion, for any risk-averse expected utility agent, entering a protective put at time 0 improves the expected utility of the future payoff at each time spot up to the maturity. Recall that this conclusion only holds if the asset has a nonpositive return rate, making it undesirable for most investors. The assumption of nonpositive return is replaced by that of a nonnegative return if the investor has a short position of the stock. In that case, a call option is a stochastic improver, following the above argument.

5.2 Widely marketable insurance contracts

Next, we consider an insurance market. Let L^1_+ be the set of nonnegative random variables in L^1, and elements in L^1_+ represent insurable losses in this section. Cheung et al. (2014) introduced the concept of universal markability. An *indemnity schedule* is a function $I : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying $0 \leq I(x) \leq x$ for each $x \geq 0$. An indemnity schedule I is *universally marketable* if for any $X \in L^1_+$, $w \in \mathbb{R}$, and increasing concave utility function u, a solution P^* to the equation

$$\mathbb{E}[u(w - X + I(X) - P)] = \mathbb{E}[u(w - X)], \quad P \in \mathbb{R}, \tag{7}$$

satisfies $P^* \geq \mathbb{E}[I(X)]$. Intuitively, it means that every risk-averse expected utility agent with insurable loss X would accept to purchase the insurance contract with payoff $I(X)$ at some price higher than or equal to the mean of $I(X)$; the insurance price being no less than the mean of the payoff is a natural requirement for the insurance provider to participate (see Arrow (1963)); later we will discuss a few examples where this is violated. Cheung et al. (2014, Theorem 3) showed that an indemnity schedule is universally marketable if and only if it is 1-Lipschitz.

Below, we offer a different angle. In an insurance market, the indemnity schedule I and the loss X are not separately considered; for instance, the indemnity schedule should be different for property insurance and for health insurance. Therefore, instead of looking for I that is marketable for all X, it is natural to look for I that is marketable for a specific X. Moreover, the insurance company may be concerned about a different premium principle than the expected value (see e.g., Denneberg (1990) and Wang et al. (1997)). We let $P_0 \geq 0$ represent the minimum acceptable price for the insurer for the indemnity I; in the previous setting it is $P_0 = \mathbb{E}[I(X)]$.

To incorporate the above two features, we say that an indemnity schedule I is *widely marketable for* $(X, P_0) \in L^1_+ \times \mathbb{R}_+$ if for any $w \in \mathbb{R}$ and increasing concave utility function u, a solution P^* to (7) satisfies $P^* \geq P_0$. Here, we choose the word "widely" to reflect that this requirement is less general than universal marketability (which holds for all X), but it is still quite broad, as it applies to all risk-averse expected utility agents. Our main results allow us to study this property for flexible choices of P_0, which is summarized in the next proposition.

Proposition 8. *Let I be an indemnity schedule, $X \in L^1_+$, and $P_0 \in \mathbb{R}_+$. If $\mathbb{E}[I(X)|X - I(X) \geq x] \geq P_0$ for all relevant x, then I is widely marketable for (X, P_0).*

Proof. Since the utility function u is increasing and concave, a solution $P^* \geq P_0$ to (7) exists for every u if $-X + I(X) - P_0 \geq_{\text{ssd}} -X$. By Proposition 2 with $W = -X + I(X) - P_0$ and $Z = P_0 - I(X)$, a sufficient condition for the desired SSD relation is

$$\mathbb{E}[P_0 - I(X)| - X + I(X) - P_0 \leq x] \leq 0 \quad \text{for all relevant values of } x.$$

By rearranging terms, the above condition is equivalent to $\mathbb{E}[I(X)|X - I(X) \geq x] \geq P_0$ for all relevant x. ∎

Due to the specification of X and P_0, the indemnity I in Proposition 8 need not be continuous (see the example below), thus a wider class of indenmity schedules than those in Theorem 3 of Cheung et al. (2014) can be included.

Example 4 (Fixed idenmity plan)**.** Let the indenmity schedule I be given by $I(x) = \mathbb{1}_{\{x\geq 1\}}$. That is, an pre-determined payment of 1 is paid if the loss reaches or exceeds 1. This kind of contract is called a fixed indemnity plan in health insurance (e.g., a fixed amount is paid upon hospital admission). Clearly, I is not continuous, but it satisfies all conditions to be an indemnity schedule. Suppose that X is exponentially distributed with mean 1. We can compute, for $x \in [0, 1]$,

$$\begin{aligned}\mathbb{E}[I(X)|X - I(X) \geq x] &= \mathbb{P}(X \geq 1|X - I(X) \geq x)\\ &= \frac{\mathbb{P}(X \geq 1 \text{ and } X - I(X) \geq x)}{\mathbb{P}(X - I(X) \geq x)}\\ &= \frac{\mathbb{P}(X \geq 1 + x)}{\mathbb{P}(X \in [x, 1) \cup [1 + x, \infty))}\\ &= \frac{e^{-(1+x)}}{e^{-(1+x)} + e^{-x} - e^{-1}} = \frac{1}{1 + e - e^{x}} \geq e^{-1}.\end{aligned}$$

Moreover, for $x > 1$, $\mathbb{E}[I(X)|X - I(X) \geq x] = 1$. Therefore, with any $P_0 \in [0, e^{-1}]$, I is widely marketable for (X, P_0). In particular, I is widely marketable for $(X, \mathbb{E}[I(X)])$, by noting $\mathbb{E}[I(X)] = e^{-1}$.

If $P_0 > \mathbb{E}[I(X)]$, then the condition $\mathbb{E}[I(X)|X - I(X) \geq x] \geq P_0$ in Proposition 8 cannot hold for all relevant x. Therefore, if the insurance company charges more than the expected value of the insurance payment, then its contract cannot be attractive to all risk-averse expected utility agents (although it may still be attractive to a subset of such agents). This is because risk-averse agents include risk-neutral ones, who do not want to pay anything more than the expected insurance payment and are not the typical insurance buyers. This may be seen as a limintation of the applicability of Proposition 8. The same limitation applies to the formulation of Cheung et al. (2014), which relies on the stronger condition $P_0 = \mathbb{E}[I(X)]$. On the positive side, the additional flexibility of P_0 provided by Proposition 8 allows an insurance company to quantitatively understand how to attract all risk-averse expected utility agents for a particular product, if they wish to, by lowering their premium to the level that the condition in Proposition 8 holds. This is relevant in the contexts of many different insurance products, commercial promotions, and government subsidized insurance. In each of these contexts, the insurance company may have an incentive to let the premium go below the expected insurance payment.

5.3 Stop-loss premium calculation

The SSD relation is also closely related to the stop-loss premium in insurance. Let $X \in L^1_+$ be an insurable loss. For a deductible level $d \geq 0$, the insurance contract that pays $(X-d)_+$ is called a stop-loss insurance contract, which is a popular form of insurance coverage; Arrow (1963) showed that the stop-loss contract is the optimal form for a risk-averse insured and a risk-neutral insurer under general conditions. The stop-loss premium of X with deductible d is then defined as $\mathbb{E}[(X-d)_+]$, widely studied in actuarial science; see e.g., Denuit and Vermandele (1998) and Dhaene et al. (2002). Our results imply a simple relation on the stop-loss premiums of two random losses.

Proposition 9. *If $X \in L^1_+$ and $Z \in L^1$ satisfy $\mathbb{E}[Z|X \geq x] \geq 0$ for all relevant values of x, then for any deductible $d \geq 0$, $X+Z$ has a larger stop-loss premium than X.*

Proposition 9 is a straightforward consequence of Corollary 4 and the well-known fact that the partial order over L^1_+ induced by stop-loss premiums at all deductible levels is equivalent to increasing convex order; see e.g., Dhaene et al. (2002).

6 Further discussions

In this section we discuss some issues related to our results. We first present an extension of the representation in Theorem 1.

In the formulation of Theorem 1, in addition to the additive payoff Z, we relied on an extra random variable $W \stackrel{\mathrm{d}}{=} X$ satisfying $Y \stackrel{\mathrm{d}}{=} W + Z$ instead of directly using $W = Y - Z$. This is needed in the classic representation (called the Strassen theorem); see Theorem 4.A.5 of Shaked and Shanthikumar (2007). The technical reason for such a construction is that the existence of W satisfying certain distributional requirements depends on the choice of Y and the underlying probability space. For instance, an independent noise to Y does not exist if the σ-algebra of Y is equal to $\mathcal{F}$.

In a recent paper, Nutz et al. (2022) established a refinement of the Strassen theorem by allowing $W = Y - Z$ in an arbitrary atomless probability space. This refinement, based on the theory of martingale optimal transport, is highly non-trivial. Using this result, we obtain a representation without involving an additional random variable W. Our results in the previous sections are presented in their current forms for consistency with the classic theorem of Strassen in its most familiar form. In what follows, $X \geq_{\rm st} Y$ means $\mathbb{P}(X > x) \geq \mathbb{P}(Y > x)$ for all $x \in \mathbb{R}$.

Theorem 10. *For any $X, Y \in L^1$, $X \leq_{\rm cx} Y$ holds if and only if $X \stackrel{\rm d}{=} Y - Z$ for some $Z \in L^1$ such that $\mathbb{E}[Z] = 0$ and*

$$\mathbb{E}[Z|Y - Z \leq x] \leq 0 \text{ for all relevant values of } x. \tag{8}$$

Moreover, $X \geq_{\rm ssd} Y$ holds if and only if $X \geq_{\rm st} Y - Z$ for some $Z \in L^1$ such that (8) holds.

Proof. The second equivalence is a direct consequence of the first one by decomposing $\geq_{\rm ssd}$ into $\geq_{\rm st}$ and $\leq_{\rm cx}$; see Theorem 4.A.6 of Shaked and Shanthikumar (2007). Below we only show the first equivalence. The "if" statement follows from Corollary 5 with $W = Y - Z$. For the "only if" statement, by Theorem 3.1 of Nutz et al. (2022), there exists $W \in L^1$ such that $W \stackrel{\rm d}{=} X$ and $\mathbb{E}[Y|W] = W$. Let $Z = Y - W$. It follows that $\mathbb{E}[Y - W|W] = 0$ and hence $\mathbb{E}[Z|Y - Z] = 0$. Therefore, Z satisfies both $X \stackrel{\rm d}{=} Y - Z$ and (8). □

We make another remark regarding the our results and comonotonicity. Assume that X and Y in Theorem 1 are comonotonic. By choosing $Z = Y - X$ and $W = X$, condition (2) becomes

$$\mathbb{E}[X|X \leq x] \geq \mathbb{E}[Y|X \leq x] \text{ for all relevant values of } x.$$

If X is continuously distributed, then, using comonotonicity, this condition is

$$\int_0^p Q_X(t)\mathrm{d}t \leq \int_0^p Q_Y(t)\mathrm{d}t \quad \text{for all } p \in (0, 1),$$

which is a well-known equivalent condition for $X \geq_{\rm ssd} Y$ (see Lemma 3).

We conclude the paper by discussing the interpretation of our representation results in relation to dependence concepts. As usual, W is interpreted as the random wealth of a decision maker. In case $\mathbb{E}[Z] = 0$, the condition (2) is positive expectation dependence of Z on W (Wright (1987)); see Li et al. (2016) for its generalizations to higher order. Therefore, the implication from (2) to $W + Z \geq_{\rm cx} W$, which is Proposition 2 with $\mathbb{E}[Z] = 0$, yields the intuitive interpretation that adding a positively expectation dependent pertubation increases the risk. Generally speaking, adding a positively dependent (in some vague sense) noise is risky; see Dhaene et al. (2002) and Puccetti and Wang (2015) for summaries on the intimate links between dependence concepts and stochastic orders. However, if $\mathbb{E}[Z] \neq 0$, then (2) no longer has an interpretation of positive dependence, as $\mathrm{cov}(W, Z)$ may be negative (see Example 2). Indeed, (2) is strictly weaker than the combination of positive expectation dependence and $\mathbb{E}[Z] \leq 0$. The resulting relation $W + Z \leq_{\mathrm{ssd}} W$ now means that an additive payoff with negative expected value in adverse scenarios of the random wealth W (i.e., on events of the form $\{W \leq x\}$), even if positively dependent, makes the overall position less desirable for any decision makers who respect SSD, that is, those who prefer more wealth to less and are risk averse.

Acknowledgements

The authors thank an Editor, two anonymous referees, David P. Brown, and Felix Liebrich for helpful discussions and comments. Ruodu Wang acknowledges financial support from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2024-03728) and Canada Research Chairs (CRC-2022-00141).

References

Arrow, K. J. (1963). Uncertainty and the welfare economics of medical care. *American Economic Review*, **53**(5), 941–973.

Brown, D. P. (2017). New characterizations of increasing risk. *Journal of Mathematical Economics*, **69**, 7–11.

Cheung, K. C., Dhaene, J., Lo, A. and Tang, Q. (2014). Reducing risk by merging counter-monotonic risks. *Insurance: Mathematics and Economics*, **54**, 58–65.

Denneberg, D. (1990). Premium calculation: Why standard deviation should be replaced by absolute deviation. *ASTIN Bulletin*, **20**, 181–190.

Denuit, M. and Vermandele, C. (1998). Optimal reinsurance and stop-loss order. *Insurance: Mathematics and Economics*, **22**(3), 229–233.

Dhaene, J., Denuit, M., Goovaerts, M. J., Kaas, R. and Vynche, D. (2002). The concept of comonotonicity in actuarial science and finance: Theory. *Insurance: Mathematics and Economics*, **31**(1), 3–33.

Embrechts, P. and Wang, R. (2015). Seven proofs for the subadditivity of Expected Shortfall. *Dependence Modeling*, **3**, 126–140.

Guan, Y., Jiao, Z. and Wang, R. (2024). A reverse ES (CVaR) optimization formula. *North American Actuarial Journal*, **28**(3), 611–625.

Hadar, J. and Russell, W. R. (1969). Rules for ordering uncertain prospects. *American Economic Review*, **59**(1), 25–34.

He, J., Tang, Q. and Zhang, H. (2016). Risk reducers in convex order. *Insurance: Mathematics and Economics*, **70**, 80–88.

Li, J., Liu, D. and Wang, J. (2016). Risk aversion with two risks: A theoretical extension. *Journal of Mathematical Economics*, **63**, 100–105.

McNeil, A. J., Frey, R. and Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques and Tools*. Revised Edition. Princeton, NJ: Princeton University Press.

Müller, A. and Stoyan, D. (2002). *Comparison Methods for Stochastic Models and Risks*. Wiley, England.

Nutz, M., Wang, R., Zhang, Z. (2022). Martingale transports and Monge maps. *arXiv*: 2209.14432.

Puccetti, G. and Wang R. (2015). Extremal dependence concepts. *Statistical Science*, **30**(4), 485–517.

Rothschild, M. and Stiglitz, J. E. (1970). Increasing risk: I. A definition. *Journal of Economic Theory*, **2**(3), 225–243.

Shaked, M. and Shanthikumar, J. G. (2007). *Stochastic Orders*. Springer Series in Statistics.

Shreve, S. E. (2004). *Stochastic Calculus for Finance II: Continuous-time Models*. Springer, New York.

Strassen, V. (1965). The existence of probability measures with given marginals. *Annals of Mathematical Statistics*, **36**(2), 423–439.

Wang, R. and Zitikis, R. (2021). An axiomatic foundation for the Expected Shortfall. *Management Science*, **67**, 1413–1429.

Wang, S., Young, V. R. and Panjer, H. H. (1997). Axiomatic characterization of insurance prices. *Insurance: Mathematics and Economics*, **21**(2), 173–183.

Wright, R. (1987). Expectation dependence of random variables, with an application in portfolio theory. *Theory and Decision*, **22**(2), 111–124.