Probability in the Engineering and Informational Sciences (2023), **37**:4 888–906
doi:10.1017/S0269964822000250

CAMBRIDGE
UNIVERSITY PRESS

RESEARCH ARTICLE

Almost first-order stochastic dominance by distorted expectations

Jianping Yang[1], Tian Zhou[1] and Weiwei Zhuang[2]

[1]Department of Mathematical Sciences, School of Science, Zhejiang Sci-Tech University, Hangzhou, Zhejiang 310018, China. E-mails: yangjp@zstu.edu.cn; ztian980117@163.com
[2]International Institute of Finance, School of Management, University of Science and Technology of China, Hefei, Anhui 230026, China. E-mail: weizh@ustc.edu.cn.

Keywords: Almost first-order stochastic dominance, Decision analysis, Distortion function, Order statistics, ROC curve

2000 Mathematics Subject Classifications: 60E15, 91B16

Abstract
Almost stochastic dominance has been receiving a great amount of attention in the financial and economic literatures. In this paper, we characterize the properties of almost first-order stochastic dominance (AFSD) via distorted expectations and investigate the conditions under which AFSD is preserved under a distortion transform. The main results are also applied to establish stochastic comparisons of order statistics and receiver operating characteristic curves via AFSD.

1. Introduction

Since the pioneering contribution of Hanoch and Levy [14], stochastic dominance (SD) theory has become an important tool for comparing risks, which provides a systematic and efficient standard paradigm for analyzing people's decision-making behavior under uncertainty. SD theory is consistent with expected utility theory, but it does not require a specific utility function. It uses the whole probability distribution rather than some common numerical characteristics such as mean and standard derivation. SD theory has been well-developed, and there are hundreds of papers on SD and its applications (see, e.g., [13,28]). The most popular notions of SD in practical applications are the *first-order stochastic dominance* (FSD) and the *second-order stochastic dominance* (SSD). They both possess some simple and tractable properties such as equivalent criteria based on integral conditions and probability transfer [19,24].

Although SD theory is attractive, some related literatures have shown that its practical application scope is very limited [1]. The dominance relation between risky prospects F and G fails to hold even in the situation that the vast majority of investors would prefer F over G (F is almost entirely below G). To provide a standard for comparing this kind of risks and reveal a preference for "most" decision makers, Leshno and Levy [17] extended the SD theory to *almost stochastic dominance* (ASD) by eliminating pathological preferences. Bali *et al.* [3] illustrated that ASD can unambiguously explain some decision-making behaviors of investors, such as a higher stock to bond ratio for the long-term investment. Up to now, ASD has played an important role in a number of fields, especially in insurance and finance, and drawn many important applications [10,18,21,31].

In recent literatures, some papers connected distorted distributions with stochastic comparisons (see, e.g., [5,16,24,30]). A function H from $[0, 1]$ to $[0, 1]$ is called a distortion function if H is increasing, and satisfies $H(0) = 0$ and $H(1) = 1$. Let X be a random variable on an atomless probability space

$(\Omega, \mathscr{F}, P)$, with cumulative distribution function (cdf) F. The distorted distribution function of F is defined by $H \circ F$, which can be regarded as the cdf of some random variable X^H if H is right continuous. Generally, the distorted function H is interpreted as a subjective weight of the original cdf F, and reflects the risk attitude of decision makers [35]. Intuitively, a concave H attaches more weight to smaller expenses, conforming to the idea of risk aversion. Yaari [35] firstly applied the distorted distribution function in dual theory of choice under risk, and some researchers proved that some SD rules via expected utility theory can be characterized through the distortion transform [20,24]. For more applications of distortions in insurance and actuarial science, see, for example, Wang [32] and Denuit *et al.* [5].

In this paper, we devote to further develop of ASD theory and its applications in the fields of reliability theory and biostatistics. Throughout, a distortion function H is always assumed to be right continuous. The risk of X is valued as its distorted expectation $\mathbb{E}_H(X)$, defined by

$$\mathbb{E}_H(X) = \int_{-\infty}^{\infty} x \, \mathrm{d}H \circ F(x) = \int_0^1 F^{-1}(u) \, \mathrm{d}H(u), \tag{1.1}$$

where $F^{-1}(u) = \inf\{x : F(x) \geq u\}$ for $u \in [0, 1]$. Let $\mathscr{L} = \{X : X \text{ is a random variable on } (\Omega, \mathscr{F}, P)\}$. Some researchers have investigated the equivalent characterizations of classical SD rules based on distortion expectation. For any random variable $X_i \in \mathscr{L}$, $i = 1, 2$, Levy and Wiener [20] proved that X_1 is greater than X_2 in FSD, if and only if, for all distortion functions H,

$$\mathbb{E}_H(X_1) \geq \mathbb{E}_H(X_2). \tag{1.2}$$

Levy and Wiener [20] also proposed that X_1 is greater than X_2 in SSD if and only if (1.2) holds for all convex distortion functions H. Wang and Young [34] showed that X_1 is greater than X_2 in increasing convex order if and only if (1.2) holds for all concave distortion functions H. These results are very meaningful since many good risk measures can be expressed as the expectations of the distorted distributions. Although FSD, SSD and increasing convex order can all be characterized via distorted expectations, it is not true in general for ASD because SD rules are defined by integrated distributions, whereas distorted expectations are related to integrated quantiles [23]. We will explore the equivalent characterizations of AFSD rules based on distorted expectations, and find that, for any $0 < \varepsilon < 1/2$, X_1 is greater than X_2 in ε-AFSD is equivalent to the condition (1.2) for all distortion functions H in a certain class of distortion functions. Furthermore, we use this main result to derive some other properties of AFSD under distortion transform.

The paper is organized as follows. In Section 2, we recall the definition of AFSD and illustrate that AFSD does not possess invariance under increasing concave or convex transforms. The main characterization result of AFSD via distorted expectation is given in Section 3.1. The other properties of AFSD under distortion transform are listed in Section 3.2. The main results in Section 3 are applied to establish stochastic comparisons of order statistics and ROC curves via AFSD in Section 4.

2. Almost stochastic dominance and distorted expectation

2.1. Almost stochastic dominance

We begin by introducing some notations. Let X_i denote the random asset with cdf F_i, $i = 1, 2$, and let $\mathscr{U}$ be the set of all differentiable utility functions on $\mathbb{R}$. For $0 < \varepsilon < 1/2$, define

$$\mathscr{U}_1^{\varepsilon} = \left\{ U \mid U \in \mathscr{U}, U'(x) \leq \inf_{x \in \mathbb{R}} \{U'(x)\} \left[\frac{1}{\varepsilon} - 1 \right], \text{ for all } x \in \mathbb{R} \right\}.$$

Throughout, denote $x_+ = \max\{x, 0\}$ for any $x \in \mathbb{R}$, and define $\|\varphi\| = \int_{-\infty}^{\infty} |\varphi(x)| \, \mathrm{d}x$ for any function $\varphi : \mathbb{R} \to \mathbb{R}$. We recall the definition of AFSD.

https://doi.org/10.1017/S0269964822000250 Published online by Cambridge University Press

Table 1. *The distribution of X_1.*

The values of X_1	0	1	2
Probabilities	1/16	3/16	3/4

Table 2. *The distribution of X_2.*

The values of X_2	1	2
Probabilities	9/16	7/16

Definition 2.1 [17]**.** *We say that X_2 is dominated by X_1 in ε-AFSD, denoted by $X_1 \geq_{\varepsilon\text{-}AFSD} X_2$ or $F_1 \geq_{\varepsilon\text{-}AFSD} F_2$, if and only if*

(i) $\mathbb{E}[U(X_1)] \geq \mathbb{E}[U(X_2)]$ *for all* $U \in \mathscr{U}_1^{\varepsilon}$, *or*
(ii) $\|(F_1 - F_2)_+\| \leq \varepsilon \| F_1 - F_2 \|$.

FSD can be regarded as ε-AFSD with $\varepsilon = 0$. Levy and Wiener [20] studied FSD and SSD through distorted expectations and investigated the classes of distortion functions that preserve FSD and SSD. In this paper, we will mainly study the properties of AFSD under distortion transforms. It should be mentioned that AFSD does not possess invariance under increasing convex or concave transforms. Let us see the following two examples.

Example 2.1. Assume that X_1 and X_2 are two random variables with respective probability mass functions F_1 and F_2 described by Tables 1 and 2.

Hence,

$$F_1(x) - F_2(x) = \begin{cases} 1/16, & 0 \leq x < 1, \\ -5/16, & 1 \leq x < 2, \\ 0, & x \geq 2. \end{cases}$$

Thus, X_1 dominates X_2 in terms of ε-AFSD, if and only if,

$$\varepsilon \geq \frac{\|(F_1 - F_2)_+\|}{\|F_1 - F_2\|} = \frac{1}{6},$$

that is, $1/6 \leq \varepsilon < 1/2$. Choose the distortion function $H(u) = \sqrt{u}$, which is increasing and concave on $[0, 1]$. Then,

$$H \circ F_1(x) - H \circ F_2(x) = \begin{cases} 1/4, & 0 \leq x < 1, \\ -1/4, & 1 \leq x < 2, \\ 0, & x \geq 2. \end{cases}$$

Hence, $\|(H \circ F_1 - H \circ F_2)_+\| = 1/4$ and $\|H \circ F_1 - H \circ F_2\| = 1/2$. Thus, for any $0 < \varepsilon < 1/2$, it holds that

$$\varepsilon < \frac{\|(H \circ F_1 - H \circ F_2)_+\|}{\|H \circ F_1 - H \circ F_2\|} = \frac{1}{2}.$$

That is, X_1^H does not dominate X_2^H in ε-AFSD.

Example 2.2. Assume that X_1 and X_2 are two random variables with respective probability mass functions F_1 and F_2 described by Tables 3 and 4.

https://doi.org/10.1017/S0269964822000250 Published online by Cambridge University Press

Table 3. *The distribution of X_1.*

The values of X_1	1	2
Probabilities	3/4	1/4

Table 4. *The distribution of X_2.*

The values of X_2	0	1	2
Probabilities	1/2	1/6	1/3

Then,

$$F_1(x) - F_2(x) = \begin{cases} -1/2, & 0 \le x < 1, \\ 1/12, & 1 \le x < 2, \\ 0, & x \ge 2. \end{cases}$$

Hence, $\|(F_1 - F_2)_+\| = 1/12$ and $\|F_1 - F_2\| = 7/12$. Thus, X_2 is dominated by X_1 in ε-AFSD, if and only if,

$$\varepsilon \ge \frac{\|(F_1 - F_2)_+\|}{\|F_1 - F_2\|} = \frac{1}{7},$$

that is, $1/7 \le \varepsilon < 1/2$. Let $H(u) = u^2$, which is increasing and convex on $[0, 1]$. Then,

$$H \circ F_1(x) - H \circ F_2(x) = \begin{cases} -1/4, & 0 \le x < 1, \\ 17/144, & 1 \le x < 2, \\ 0, & x \ge 2. \end{cases}$$

Hence, $\|(H \circ F_1 - H \circ_2)_+\| = 17/144$ and $\|H \circ F_1 - H \circ F_2\| = 53/144$. Then, for $\varepsilon = 16/53$,

$$\varepsilon < \frac{\|(H \circ F_1 - H \circ F_2)_+\|}{\|H \circ F_1 - H \circ F_2\|} = \frac{17}{53}.$$

That is, X_1^H does not dominate X_2^H in 16/53-AFSD.

2.2. Distorted expectation

Distorted expectation is very important in the statistical, financial and economic literatures. Suppose that X is a random variable with cdf $F(x)$ and H is a right continuous distortion function that leads to a distorted probability measure $H \circ F$. It can be shown that the Choquet integral of a random variable X, with respect to a distortion function H, is equivalent to the expectation of the variable X under the distorted distribution $H \circ F$:

$$\begin{aligned} E_H(X) &= \int_{-\infty}^{\infty} x \, \mathrm{d}H \circ F(x) \\ &= -\int_{-\infty}^{0} H \circ F(x) \, \mathrm{d}x + \int_{0}^{\infty} (1 - H \circ F(x)) \, \mathrm{d}x. \end{aligned}$$

In finance and insurance, the Choquet integral with the distortion function has been proposed to measure risks [33]. For any nonnegative random variable X such as loss, depending on the chosen

https://doi.org/10.1017/S0269964822000250 Published online by Cambridge University Press

distortion function, different distorted risk measures are obtained. When $H(u) = 1-(1-u)^n$, $u \in [0,1]$, for a positive integer n, we obtain

$$\mathbb{E}_H(X) = \int_0^\infty [1 - F(x)]^n \, \mathrm{d}x,$$

which is the generalized Gini index introduced by Donaldson and Weymark [6]. When

$$H_1(u) = \begin{cases} 0, & \text{if } 0 \le u < \alpha \\ 1, & \text{if } \alpha \le u \le 1 \end{cases},$$

for any $0 \le \alpha \le 1$, we obtain

$$\mathbb{E}_{H_1}(X) = \int_0^\infty [1 - H \circ F(x)] \, \mathrm{d}x = F^{-1}(\alpha)$$

which is the expression of the popular risk measure VaR_α (Value-at-Risk at confidence level α). Let $\widetilde{H}(u) = 1-H(1-u)$, $u \in [0,1]$, denote the dual distortion function of H. In insurance, $\mathbb{E}_{\widetilde{H}}(X)$ has been proposed to measure the risk and compute risk premiums of X, see Wang [32,33] and Denuit *et al.* [5].

Let $\mathscr{H}$ be the class of all distortion functions. For $H \in \mathscr{H}$, define

$$H'_+(1) = \lim_{u \to 1^-} \frac{H(u) - 1}{u - 1}$$

and

$$H'_+(u) = \lim_{\triangle u \to 0^+} \frac{H(u + \triangle u) - H(u)}{\triangle u}, \quad u \in [0,1).$$

For $0 < \varepsilon < 1/2$, set

$$\mathscr{H}_1^\varepsilon = \left\{ H \mid H \in \mathscr{H}, H'_+(u) \le \inf_{u \in [0,1]} \{H'_+(u)\} \left[\frac{1}{\varepsilon} - 1\right], \text{ for all } u \in [0,1] \right\}. \tag{2.1}$$

It is clear that $H_1 \in \mathscr{H}_1^\varepsilon$, and some piece-wise linear distortion functions without differentiability provided by Balbás *et al.* [2] are also included in $\mathscr{H}_1^\varepsilon$, for example:

$$H_2(u) = \begin{cases} \frac{1}{3}u, & \text{if } 0 \le u < \frac{1}{3} \\ \frac{4}{3}u - \frac{1}{3}, & \text{if } \frac{1}{3} \le u \le 1 \end{cases}.$$

For any $H \in \mathscr{H}_1^\varepsilon$, the distorted expectation $E_H(\cdot)$ as risk measure, has some good properties such as law invariance, translation invariance, positive homogeneity and monotonicity. However, subadditivity does not necessarily satisfied because the distortion function $H \in \mathscr{H}_1^\varepsilon$ is not convex. The famous Kusuoka representation shows that under an atomless probability space, any law invariant coherent risk measure is a supremum of distorted expectations with convex distortions [8], Section 4.6] and one necessarily has a distorted expectation under additional assumption of being comonotone additive. In addition, subadditivity is not necessarily satisfied. The reader can see the next example.

Example 2.3. Suppose that identically distributed Bernoulli random variables X_1 and X_2 with the discrete joint distribution given in Table 5.

https://doi.org/10.1017/S0269964822000250 Published online by Cambridge University Press

Table 5. *Joint distribution of X_1 and X_2.*

The values of (X_1, X_2)	(0, 0)	(0, 1)	(1, 0)	(1, 1)
Probabilities	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$

Assume that distortion function H is as follows:

$$H(u) = \begin{cases} (u + \frac{3}{8})^2 - \frac{9}{64}, & \text{if } 0 \le u < \frac{5}{8}, \\ \frac{3}{8}u + \frac{5}{8}, & \text{if } \frac{5}{8} \le u \le 1. \end{cases}$$

Clearly, $H \in \mathscr{H}_1^{\varepsilon}$ where $0 < \varepsilon \le \frac{3}{19}$. The distorted expectation are

$$\mathbb{E}_H(X_1 + X_2) = \frac{15}{32},$$

and

$$\mathbb{E}_H(X_1) = \mathbb{E}_H(X_2) = \frac{9}{64}.$$

Thus, we obtain

$$\mathbb{E}_H(X_1 + X_2) > \mathbb{E}_H(X_1) + \mathbb{E}_H(X_2).$$

It is well known that the distorted risk measure with convex distortions preserves some popular stochastic dominance. In this paper, we will investigate the isotonicity of the distorted expectation with distortion function $H \in \mathscr{H}_1^{\varepsilon}$ under some stochastic dominances to develop the dual theory of choice under risk.

3. Main results

In this section, we characterize AFSD via distorted expectation, and then investigate the properties of AFSD under distortion transforms.

3.1. Characterization of AFSD via distorted expectations

In this subsection, we first show that the distorted expectation is isotonic under AFSD by using the following Lemma 3.1.

Lemma 3.1. *For any* $0 < \varepsilon < 1/2$, $X_1 \ge_{\varepsilon\text{-}AFSD} X_2$ *if and only if*

$$\int_0^1 [F_2^{-1}(u) - F_1^{-1}(u)]_+ \,\mathrm{d}u \le \varepsilon \int_0^1 |F_1^{-1}(u) - F_2^{-1}(u)| \,\mathrm{d}u. \tag{3.1}$$

Proof. Note that

$$\|F_1 - F_2\| = \int_0^1 |F_1^{-1}(u) - F_2^{-1}(u)| \,\mathrm{d}u$$

and

$$\|(F_1 - F_2)_+\| = \int_0^1 [F_2^{-1}(u) - F_1^{-1}(u)]_+ \,\mathrm{d}u.$$

The desired result now follows directly. □

https://doi.org/10.1017/S0269964822000250 Published online by Cambridge University Press

Theorem 3.2. *For $0 < \varepsilon < 1/2$, the following two statements are equivalent:*

(i) $X_1 \geq_{\varepsilon\text{-}AFSD} X_2$;
(ii) $\mathbb{E}_H[X_1] \geq \mathbb{E}_H[X_2]$ *for all* $H \in \mathscr{H}_1^{\varepsilon}$.

Proof. (i) ⇒ (ii): From Lemma 3.1, we have, for any $0 < \varepsilon < 1/2$, $X_1 \geq_{\varepsilon\text{-AFSD}} X_2$ if and only if

$$\int_0^1 [F_2^{-1}(u) - F_1^{-1}(u)]_+ \,\mathrm{d}u \leq \frac{\varepsilon}{1-\varepsilon} \int_0^1 [F_1^{-1}(u) - F_2^{-1}(u)]_+ \,\mathrm{d}u. \tag{3.2}$$

Thus, for any $H \in \mathscr{H}_1^{\varepsilon}$,

$$\begin{aligned}
&\mathbb{E}_H(X_1) - \mathbb{E}_H(X_2)\\
&= \int_0^1 F_1^{-1}(u)\,\mathrm{d}H(u) - \int_0^1 F_2^{-1}(u)\,\mathrm{d}H(u)\\
&= \int_0^1 [F_1^{-1}(u) - F_2^{-1}(u)]_+ \,\mathrm{d}H(u) - \int_0^1 [F_2^{-1}(u) - F_1^{-1}(u)]_+ \,\mathrm{d}H(u)\\
&= \int_0^1 [F_1^{-1}(u) - F_2^{-1}(u)]_+ H'_+(u)\,\mathrm{d}u - \int_0^1 [F_2^{-1}(u) - F_1^{-1}(u)]_+ H'_+(u)\,\mathrm{d}u\\
&\geq \int_0^1 [F_1^{-1}(u) - F_2^{-1}(u)]_+ H'_+(u)\,\mathrm{d}u - \inf_{u\in[0,1]}\{H'_+(u)\}\frac{1-\varepsilon}{\varepsilon}\int_0^1 (F_2^{-1}(u) - F_1^{-1}(u))_+ \,\mathrm{d}u\\
&\geq \inf_{u\in[0,1]}\{H'_+(u)\}\left\{\int_0^1 [F_1^{-1}(u) - F_2^{-1}(u)]_+ \,\mathrm{d}u - \frac{1-\varepsilon}{\varepsilon}\int_0^1 [F_2^{-1}(u) - F_1^{-1}(u)]_+ \,\mathrm{d}u\right\}\\
&\geq 0.
\end{aligned}$$

That is, $\mathbb{E}_H(X_1) \geq \mathbb{E}_H(X_2)$.

(ii) ⇒ (i): Define a distortion function H with right derivative

$$H'_+(u) = \begin{cases} (1-\varepsilon)/\varepsilon, & \text{for } F_2^{-1}(u) > F_1^{-1}(u),\\ 1, & \text{for } F_2^{-1}(u) \leq F_1^{-1}(u). \end{cases}$$

It is easy to verify that $H \in \mathscr{H}_1^{\varepsilon}$. Thus,

$$\begin{aligned}
\mathbb{E}_H(X_1) - \mathbb{E}_H(X_2) &= \int_0^1 [F_1^{-1}(u) - F_2^{-1}(u)]_+ H'_+(u)\,\mathrm{d}u - \int_0^1 [F_2^{-1}(u) - F_1^{-1}(u)]_+ H'_+(u)\,\mathrm{d}u\\
&= \int_0^1 [F_1^{-1}(u) - F_2^{-1}(u)]_+ \,\mathrm{d}u - \frac{1-\varepsilon}{\varepsilon}\int_0^1 [F_2^{-1}(u) - F_1^{-1}(u)]_+ \,\mathrm{d}u.
\end{aligned}$$

Hence, for all $H \in \mathscr{H}_1^{\varepsilon}$, $\mathbb{E}_H(X_1) \geq \mathbb{E}_H(X_2)$ implies (3.2). That is, $X_1 \geq_{\varepsilon\text{-AFSD}} X_2$. This completes the proof of the theorem. □

3.2. Properties under distortion transform

Levy and Wiener [20] studied the closure of SD under a distortion transform on the space of distribution functions. They found that all distortion functions preserve FSD, whereas only concave distortion functions preserve SSD. In this subsection, we will prove that under the proper conditions, the ε-AFSD is also preserved under a distortion transform.

Proposition 3.3. *For any $0 < \varepsilon, \varepsilon_1 < 1/2$ such that $0 < \varepsilon/\varepsilon_1 < 1/2$, we have*

$$X_1 \geq_{\varepsilon\text{-}AFSD} X_2 \implies X_1^H \geq_{(\varepsilon/\varepsilon_1)\text{-}AFSD} X_2^H, \quad \text{for all } H \in \mathscr{H}_1^{\varepsilon_1}.$$

https://doi.org/10.1017/S0269964822000250 Published online by Cambridge University Press

Proof. From Theorem 3.2, it follows that $X_1^H \geq_{(\varepsilon/\varepsilon_1)\text{-AFSD}} X_2^H$ if and only if $\mathbb{E}_h[X_1^H] \geq \mathbb{E}_h[X_2^H]$ for all $h \in \mathscr{H}_1^{\varepsilon/\varepsilon_1}$, that is,

$$\int_0^1 F_1^{-1}(H^{-1}(u))\,\mathrm{d}h(u) \geq \int_0^1 F_2^{-1}(H^{-1}(u))\,\mathrm{d}h(u) \tag{3.3}$$

holds. Set $H^{-1}(u) = v$, then (3.3) reduces to

$$\int_0^1 F_1^{-1}(v)\,\mathrm{d}h \circ H(v) \geq \int_0^1 F_2^{-1}(v)\,\mathrm{d}h \circ H(v). \tag{3.4}$$

Hence, to prove that (3.4) holds for all $h \in \mathscr{H}_1^{\varepsilon/\varepsilon_1}$, it suffices to verify $h \circ H \in \mathscr{H}_1^{\varepsilon}$. Since $h \in \mathscr{H}_1^{\varepsilon/\varepsilon_1}$ and $H \in \mathscr{H}_1^{\varepsilon_1}$, for $v \in [0,1]$, we have

$$\begin{aligned}
(h \circ H)'_+(v) &= h'_+(H(v)) \cdot H'_+(v) \\
&\leq \inf\{h'_+(H(v))\}\left[\frac{\varepsilon_1}{\varepsilon} - 1\right] \cdot \inf\{H'_+(v)\}\left[\frac{1}{\varepsilon_1} - 1\right] \\
&\leq \inf\{h'_+(H(v)) \cdot H'_+(v)\} \cdot \frac{(\varepsilon_1 - \varepsilon)(1 - \varepsilon_1)}{\varepsilon\varepsilon_1} \\
&\leq \inf\{(h \circ H)'_+(v)\} \cdot \frac{\varepsilon_1 - \varepsilon}{\varepsilon\varepsilon_1} \\
&\leq \inf\{(h \circ H)'_+(v)\} \cdot \frac{1 - \varepsilon}{\varepsilon}.
\end{aligned}$$

This implies that $h \circ H \in \mathscr{H}_1^{\varepsilon}$. This completes the proof. □

Corollary 3.4. *Let $\varepsilon_1, \varepsilon_2 \in (0, 1/2)$ such that $0 < \varepsilon_1/\varepsilon_2 < 1/2$. If H_1, H_2 be two distortion functions with $H_2 \circ H_1^{-1} \in \mathscr{H}_1^{\varepsilon_1/\varepsilon_2}$, then*

$$X_1^{H_1} \geq_{\varepsilon_1\text{-}AFSD} X_2^{H_1} \Longrightarrow X_1^{H_2} \geq_{\varepsilon_2\text{-}AFSD} X_2^{H_2}.$$

Proof. Since $X_1^{H_1} \geq_{\varepsilon_1}$-AFSD $X_2^{H_1}$ and $H_2 \circ H_1^{-1} \in \mathscr{H}_1^{\varepsilon_1/\varepsilon_2}$, we have

$$X_1^{(H_2 \circ H_1^{-1}) \circ H_1} \geq_{\varepsilon_2\text{-AFSD}} X_2^{(H_2 \circ H_1^{-1}) \circ H_1}$$

by Proposition 3.3, that is, $X_1^{H_2} \geq_{\varepsilon_2\text{-AFSD}} X_2^{H_2}$. This completes the proof. □

Proposition 3.5. *For any $0 < \varepsilon < 1/2$, if the distortion functions H_1 and H_2 satisfy $H_1(u) \leq H_2(u)$ for any $u \in [0,1]$, then*

$$X_1^{H_1} \geq_{\varepsilon\text{-}AFSD} X_2^{H_1} \Longrightarrow X_1^{H_1} \geq_{\varepsilon\text{-}AFSD} X_2^{H_2}.$$

Proof. Since $X_1^{H_1} \geq_{\varepsilon\text{-AFSD}} X_2^{H_1}$, it follows from Theorem 3.2 that

$$\int_0^1 F_1^{-1}(H_1^{-1}(u))\,\mathrm{d}H(u) \geq \int_0^1 F_2^{-1}(H_1^{-1}(u))\,\mathrm{d}H(u) \tag{3.5}$$

for any $H \in \mathscr{H}_1^{\varepsilon}$. Since $H_1(u) \leq H_2(u)$ for any $u \in [0,1]$, we have $H_1 \circ F_2(x) \leq H_2 \circ F_2(x)$ for any $x \in \mathbb{R}$, and hence, $F_2^{-1} \circ H_1^{-1}(u) \geq F_2^{-1} \circ H_2^{-1}(u)$ for $u \in [0,1]$. Therefore, for any $H \in \mathscr{H}_1^{\varepsilon}$,

$$\int_0^1 F_2^{-1} \circ H_1^{-1}(u)\,\mathrm{d}H(u) \geq \int_0^1 F_2^{-1} \circ H_2^{-1}(u)\,\mathrm{d}H(u). \tag{3.6}$$

https://doi.org/10.1017/S0269964822000250 Published online by Cambridge University Press

Combining (3.5) with (3.6), we have

$$\int_0^1 F_1^{-1} \circ H_1^{-1}(u)\,\mathrm{d}H(u) \geq \int_0^1 F_2^{-1} \circ H_2^{-1}(u)\,\mathrm{d}H(u)$$

for any $H \in \mathscr{H}_1^{\varepsilon}$, that is, $X_1^{H_1} \geq_{\varepsilon\text{-AFSD}} X_2^{H_2}$. This completes the proof. □

Proposition 3.6. *For two cdfs F_1 and F_2, if F_1 singly crosses F_2 from below, and H is a concave distortion function, then for $0 < \varepsilon < 1/2$,*

$$F_1 \geq_{\varepsilon\text{-AFSD}} F_2 \implies H \circ F_1 \geq_{\varepsilon\text{-AFSD}} H \circ F_2.$$

Proof. By Lemma 3.1, $F_1 \geq_{\varepsilon\text{-AFSD}} F_2$ if and only if

$$\int_0^1 \left\{[F_2^{-1}(u) - F_1^{-1}(u)]_+ - \frac{\varepsilon}{1-\varepsilon}[F_1^{-1}(u) - F_2^{-1}(u)]_+\right\}\,\mathrm{d}u \leq 0.$$

To prove $H \circ F_1 \geq_{\varepsilon\text{-AFSD}} H \circ F_2$, it is sufficient to prove

$$\int_0^1 \left\{[F_2^{-1} \circ H^{-1}(u) - F_1^{-1} \circ H^{-1}(u)]_+ - \frac{\varepsilon}{1-\varepsilon}[F_1^{-1} \circ H^{-1}(u) - F_2^{-1} \circ H^{-1}(u)]_+\right\}\mathrm{d}u \leq 0$$

or, equivalently,

$$\int_0^1 \left\{[F_2^{-1}(v) - F_1^{-1}(v)]_+ - \frac{\varepsilon}{1-\varepsilon}[F_1^{-1}(v) - F_2^{-1}(v)]_+\right\}\,\mathrm{d}H(v) \leq 0$$

by setting $v = H^{-1}(u)$. Define

$$\Delta_1(v) := \int_0^v \left\{[F_2^{-1}(u) - F_1^{-1}(u)]_+ - \frac{\varepsilon}{1-\varepsilon}[F_1^{-1}(u) - F_2^{-1}(u)]_+\right\}\,\mathrm{d}u, \quad v \in [0, 1].$$

Since F_1 singly crosses F_2 from below, it follows that $\Delta_1(v) \leq \Delta_1(1) \leq 0$ for all $v \in [0, 1]$. Since H is concave, we have

$$\begin{aligned}
&\int_0^1 \left\{[F_2^{-1}(v) - F_1^{-1}(v)]_+ - \frac{\varepsilon}{1-\varepsilon}[F_1^{-1}(v) - F_2^{-1}(v)]_+\right\}\,\mathrm{d}H(v)\\
&= \int_0^1 H'_+(v)\,\mathrm{d}\Delta_1(v)\\
&= H'_+(1)\Delta_1(1) - \int_0^1 \Delta_1(v)\,\mathrm{d}H'_+(v) \leq 0.
\end{aligned}$$

This completes the proof. □

Proposition 3.7. *Let F_1 and F_2 be two cdfs with a common support (a, b), where $-\infty \leq a < b \leq +\infty$. If F_1 singly crosses F_2 from below, and H is an increasing and convex function such that $F_1 \circ H$ and $F_2 \circ H$ are also two cdfs, then for $0 < \varepsilon < 1/2$,*

$$F_1 \geq_{\varepsilon\text{-AFSD}} F_2 \implies F_1 \circ H \geq_{\varepsilon\text{-AFSD}} F_2 \circ H.$$

Proof. By Definition 2.1, $F_1 \geq_{\varepsilon\text{-AFSD}} F_2$ if and only if

$$\int_a^b [F_1(x) - F_2(x)]_+\,\mathrm{d}x \leq \frac{\varepsilon}{1-\varepsilon}\int_a^b [F_2(x) - F_1(x)]_+\,\mathrm{d}x.$$

https://doi.org/10.1017/S0269964822000250 Published online by Cambridge University Press

To prove $F_1 \circ H \geq_{\varepsilon\text{-AFSD}} F_2 \circ H$, it suffices to prove

$$\int_{-\infty}^{\infty} [F_1 \circ H(x) - F_2 \circ H(x)]_+ \, \mathrm{d}x \leq \frac{\varepsilon}{1-\varepsilon} \int_{-\infty}^{\infty} [F_2 \circ H(x) - F_1 \circ H(x)]_+ \, \mathrm{d}x$$

or, equivalently,

$$\int_a^b [F_1(x) - F_2(x)]_+ \, \mathrm{d}H^{-1}(x) \leq \frac{\varepsilon}{1-\varepsilon} \int_a^b [F_2(x) - F_1(x)]_+ \, \mathrm{d}H^{-1}(x).$$

Define

$$\Delta_2(y) := \int_a^y \left\{ [F_1(x) - F_2(x)]_+ - \frac{\varepsilon}{1-\varepsilon} [F_2(x) - F_1(x)]_+ \right\} \mathrm{d}x, \quad y \in [a, b].$$

Since F_1 singly crosses F_2 from below, it follows that $\Delta_2(y) \leq \Delta_2(b) \leq 0$ for all $y \in [a, b]$. Define $\eta(y) = H^{-1}(y)$ for $y \in [a, b]$, and let $\eta'_+(y)$ denote the right derivative of η at y. Therefore,

$$\begin{aligned} &\int_a^b \left\{ [F_1(x) - F_2(x)]_+ - \frac{\varepsilon}{1-\varepsilon} [F_2(x) - F_1(x)]_+ \right\} \mathrm{d}H^{-1}(x) \\ &= \int_a^b \eta'_+(x) \, \mathrm{d}\Delta_2(x) \\ &= \eta'_+(b)\Delta_2(b) - \int_a^b \Delta_2(x) \, \mathrm{d}\eta'_+(x) \leq 0, \end{aligned}$$

where $\eta'_+(b) = \lim_{y \to b-} \eta'_+(y)$, and the last inequality follows from the convexity of H. This completes the proof. □

4. Applications

In this section, we will apply our results to establish stochastic comparisons of order statistics and ROC curves with respect to AFSD.

4.1. Stochastic comparisons of order statistics

Order statistics are widely used in reliability, data analysis, risk management, auction theory, statistical inference, and many other applied areas. Let $X_1, \ldots, X_n$ be a random sample of size n from a distribution F. Denote by $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ its order statistics. If $X_1, \ldots, X_n$ are independent, it is well known that the cdf of $X_{k:n}$ is given by $F_B \circ F$, where F_B is the beta distribution with parameters k and $n - k + 1$.

Order statistics have a close connection with the lifetimes of k-out-of-n systems. In reliability theory, a k-out-of-n system is the system consisting of n components and normally operating if and only if at least k of the n components work. Considering a k-out-of-n system with all components having a common cdf F, then, its lifetime $T_{k:n}$ is the same as $X_{(n-k+1):n}$.

A significant body of literature exists on stochastic comparisons of order statistics in the past two decades. One may refer to Shaked and Shanthikumar [30] and references therein. In the following, we deal with the problem of comparing the lifetimes of k-out-of-n systems with respect to AFSD. First, we recall some popular ageing notions.

Definition 4.1 [16]**.** *Let X be a random variable with cdf F. Then, X or F is said to be*

- *Convex if F is convex on its support, denoted by $F \in \mathscr{F}_{\mathrm{C}}$;*
- *Logit-convex if $\log\left(F/\bar{F}\right)$ is convex on the support of F, denoted by $F \in \mathscr{F}_{\mathrm{CL}}$;*

https://doi.org/10.1017/S0269964822000250 Published online by Cambridge University Press

- *Odds-convex if $F/\bar{F}$ is convex on the support of F, denoted by $F \in \mathscr{F}_{\mathrm{CO}}$;*
- *IFR (increasing failure rate) if $\bar{F}(x)$ is log-concave, denoted by $F \in \mathscr{F}_{\mathrm{IFR}}$.*

If X has a probability density function f, then $F \in \mathscr{F}_{\mathrm{C}}$ if and only if f is increasing. Define $\lambda(x) = f(x)/\bar{F}(x)$ on $\{x : F(x) < 1\}$, which is called the failure rate function of F. Then, $F \in \mathscr{F}_{\mathrm{IFR}}$ if and only if $\lambda(x)$ is increasing on $\{x : F(x) < 1\}$. $\mathscr{F}_{\mathrm{IFR}}$ is an important class in reliability theory and contains many relevant models. It is clear that $\mathscr{F}_{\mathrm{C}} \subset \mathscr{F}_{\mathrm{IFR}}$.

The ageing distribution set $\mathscr{F}_{\mathrm{CL}}$ contains distributions with unbounded support, which has been studied by Zimmer *et al.* [36], Sankaran and Jayakumar [29] and Navarro *et al.* [27]. $F \in \mathscr{F}_{\mathrm{CL}}$ if and only if the log-odds rate $\lambda(x)/F(x)$ is increasing. Therefore, $\mathscr{F}_{\mathrm{CL}} \subset \mathscr{F}_{\mathrm{IFR}}$.

The ageing distribution set $\mathscr{F}_{CO}$ can be characterized as follows: $F \in \mathscr{F}_{\mathrm{CO}}$ if and only if the ratio between the failure rate and the survival function, $\lambda(x)/\bar{F}(x)$, is increasing [15]. Therefore, $\mathscr{F}_{\mathrm{CO}}$ is the widest ageing distribution set and $\mathscr{F}_{\mathrm{IFR}} \subset \mathscr{F}_{\mathrm{CO}}$. Several examples of distributions belonging to $\mathscr{F}_{\mathrm{C}}$, $\mathscr{F}_{\mathrm{CL}}$, $\mathscr{F}_{\mathrm{CO}}$ and $\mathscr{F}_{\mathrm{IFR}}$ can be found in Lando *et al.* [16].

To study ageing patterns of lifetimes of k-out-of-n systems, Lando *et al.* [16] compared different order statistics with respect to SSD and derived sufficient dominance conditions by identifying the class of component lifetimes. In this subsection, we will explore stochastic comparison of order statistics via AFSD. The following lemma is due to Theorem 7 in Müller *et al.* [26] since $(1+\gamma, 1+\gamma)$-SD is equivalent to ε-AFSD with $\varepsilon = \gamma/(1+\gamma)$ (see [25], Definition 4.2]).

Lemma 4.1 [26]. *Let X_1 and X_2 be random variables with $\mathbb{E}(X_1) = \mu_1$, $\mathbb{E}(X_2) = \mu_2$,* $\mathrm{Var}(X_1) = \sigma_1^2$, $\mathrm{Var}(X_2) = \sigma_2^2$ *and $\mu_1 > \mu_2$. Define*

$$t := \frac{\mu_1 - \mu_2}{\sigma_2 + \sigma_1},$$

and

$$\varepsilon^*(t) = \frac{1}{2 + 2t(t + \sqrt{t^2+1})}.$$

Then $X_1 \geq_{\varepsilon\text{-}AFSD} X_2$ for $\varepsilon^*(t) < \varepsilon < 1/2$.

Let $B_{i,n} \sim \mathrm{beta}(i, n-i+1)$. Denote

$$\mu_B(i,n) := \mathbb{E}[B_{i,n}] = \frac{i}{n+1} \quad \text{and} \quad \sigma_B^2(i,n) := \mathrm{Var}(B_{i,n}) = \frac{i(n-i+1)}{(n+1)^2(n+2)}.$$

For the sake of simplification, denote

$$t_1 = \frac{\mu_B(i,n) - \mu_B(j,m)}{\sigma_B(i,n) + \sigma_B(j,m)}.$$

It is known that if $j \geq i$ and $n - m \geq i - j$, then $X_{j:m} \geq_{\mathrm{FSD}} X_{i:n}$ (see, e.g., [4]).

The next two propositions give the conditions under which $X_{j:m}$ and $X_{i:n}$ can be ordered by ε-AFSD when the condition $j \geq i$ is violated and F belongs to any one of $\mathscr{F}_{\mathrm{C}}$, $\mathscr{F}_{\mathrm{CO}}$, $\mathscr{F}_{\mathrm{IFR}}$ and $\mathscr{F}_{\mathrm{CL}}$.

Proposition 4.2. *For any $F \in \mathscr{F}_{\mathrm{C}}$, if $n-m > i-j > 0$ and $i/(n+1) > j/(m+1)$, then $X_{i:n} \geq_{\varepsilon\text{-}AFSD} X_{j:m}$ for $\varepsilon^*(t_1) \leq \varepsilon < 1/2$.*

Proof. Let $B_{i,n} \sim beta(i, n-i+1)$ and $B_{j,m} \sim beta(j, m-j+1)$. Since $n-m > i-j > 0$, the cdf $F_{B_{i,n}}$ of $B_{i,n}$ singly crosses the cdf $F_{B_{j,m}}$ of $B_{j,m}$ from below by Lemma A.1. Since $i/(n+1) > j/(m+1)$, we have $\mu_B(i,n) := \mathbb{E}[B_{i,n}] > \mathbb{E}[B_{j,m}] =: \mu_B(j,m)$. Therefore, from Lemma 4.1, we have $B_{i,n} \geq_{\varepsilon\text{-AFSD}} B_{j,m}$ for $\varepsilon^*(t_1) \leq \varepsilon < 1/2$.

https://doi.org/10.1017/S0269964822000250 Published online by Cambridge University Press

On the other hand, the cdf of $X_{i:n}$ can be formulated as $F_{B_{i,n}} \circ F$. Since F is convex, it follows from Proposition 3.7 that $F_{B_{i,n}} \circ F \geq_{\varepsilon\text{-AFSD}} F_{B_{j,m}} \circ F$. That is, $X_{i:n} \geq_{\varepsilon\text{-AFSD}} X_{j:m}$. This completes the proof of this proposition. □

Define

$$t_\ell = \frac{\mu_{h_\ell}(i,n) - \mu_{h_\ell}(j,m)}{\sigma_{h_\ell}(i,n) + \sigma_{h_\ell}(j,m)}, \quad \ell = 2, 3, 4,$$

where $\mu_{h_\ell}(\cdot,\cdot)$ and $\sigma_{h_\ell}(\cdot,\cdot)$ are defined in Lemmas A.2 and A.3.

Proposition 4.3. *Assume that $n - m > i - j > 0$, and define $\psi(x) = \Gamma'(x)/\Gamma(x)$ for $x > 0$.*

(1) *For any $F \in \mathscr{F}_{\text{CO}}$, if $i/(n-i) > j/(m-j)$, then*

$$X_{i:n} \geq_{\varepsilon\text{-AFSD}} X_{j:m} \quad \text{for } \varepsilon^*(t_2) \leq \varepsilon < \tfrac{1}{2};$$

(2) *For any $F \in \mathscr{F}_{\text{CL}}$, if $\psi(i) - \psi(n-i+1) > \psi(j) - \psi(m-j+1)$, then*

$$X_{i:n} \geq_{\varepsilon\text{-AFSD}} X_{j:m} \quad \text{for } \varepsilon^*(t_3) \leq \varepsilon < \tfrac{1}{2};$$

(3) *For any $F \in \mathscr{F}_{\text{IFR}}$, if $\psi(n+1) - \psi(n-i+1) > \psi(m+1) - \psi(m-j+1)$, then*

$$X_{i:n} \geq_{\varepsilon\text{-AFSD}} X_{j:m} \quad \text{for } \varepsilon^*(t_4) \leq \varepsilon < \tfrac{1}{2}.$$

Proof. (1) Since $h_2(p) = p/(1-p)$ for $p \in (0,1)$, its inverse function $h_2^{-1}(x) = x/(1+x)$ is increasing in $x \in \mathbb{R}_+$. Note that $F_{B_{i,n}} \circ h_2^{-1}$ is the cdf of random variable $h_2(B_{i,n})$. From Lemmas A.2 and 4.1, it follows that, under the condition $i/(n-i) > j/(m-j)$,

$$F_{B_{i,n}} \circ h_2^{-1} \geq_{\varepsilon\text{-AFSD}} F_{B_{j,m}} \circ h_2^{-1} \quad \text{for } \varepsilon^*(t_2) \leq \varepsilon < \tfrac{1}{2}.$$

By Lemma A.1, the condition $n - m > i - j > 0$ implies that $F_{B_{i,n}}$ singly crosses $F_{B_{j,m}}$ from below. Hence, $F_{B_{i,n}} \circ h_2^{-1}$ singly crosses $F_{B_{j,m}} \circ h_2^{-1}$ from below. On the other hand, since $F \in \mathscr{F}_{\text{CO}}$, we have $h_2 \circ F$ is convex. Therefore, from Proposition 3.7, we have

$$F_{B_{i,n}} \circ h_2^{-1} \circ h_2 \circ F \geq_{\varepsilon\text{-AFSD}} F_{B_{j,m}} \circ h_2^{-1} \circ h_2 \circ F,$$

for $\varepsilon^*(t_2) \leq \varepsilon < 1/2$. That is, $X_{i:n} \geq_{\varepsilon\text{-AFSD}} X_{j:m}$ for $\varepsilon^*(t_2) \leq \varepsilon < 1/2$.

(2) and (3) The proof is similar to part (1) by applying Lemmas A.1, A.3 and 4.1. This completes the proof of the proposition. □

4.2. Stochastic comparison of ROC curves

The *Receiver Operating Characteristic* (ROC) curve is one of the most common statistical tools to assess classifier performance [22]. It is generated by plotting the fraction of true positive rate (TPR) against false positive rate (FPR) at various operating points as the decision threshold or misclassification cost [7]. Let us consider a binary classification tool by assigning a real-valued score to classify items into two categories: good or bad. Let random variables X_B and X_G represent the respective scores of the bad population with cdf F_B and good population with cdf F_G. Then, ROC curve can be defined as

$$\text{ROC}_X(u) = F_B \circ F_G^{-1}(u) \quad \text{for } u \in (0,1).$$

However, the selection of the best classifier is quite challenging when ROC curves intersect. The *Area Under the Curve* (AUC) is one of the most common measures to evaluate the classifier performance, but it has well-understood weakness when comparing ROC curves which cross. Gigliarano *et al.* [9] proposed

https://doi.org/10.1017/S0269964822000250 Published online by Cambridge University Press

a novel approach of ROC comparison and investigated the relationships between ROC orderings and integer-degree stochastic dominance in a theoretical framework. In this subsection, we will focus on extending the methodological approach to AFSD.

Proposition 4.4. *Let F_1, F_2 and G be three cdfs such that $F_1 \circ G^{-1}$ and $F_2 \circ G^{-1}$ are also cdfs. If F_1 singly crosses F_2 from below, and G is concave, then for any $0 < \varepsilon < 1/2$,*

$$F_1 \geq_{\varepsilon\text{-}AFSD} F_2 \Longrightarrow F_1 \circ G^{-1} \geq_{\varepsilon\text{-}AFSD} F_2 \circ G^{-1}.$$

Proof. The desired result follows from Proposition 3.7 by observing that G^{-1} is increasing and convex. □

Proposition 4.5. *Let G_1, G_2 and F be three cdfs such that $F \circ G_1^{-1}$ and $F \circ G_2^{-1}$ are also cdfs. If G_2 singly crosses G_1 from below, and F is concave, then for any $0 < \varepsilon < 1/2$,*

$$G_2 \geq_{\varepsilon\text{-}AFSD} G_1 \implies F \circ G_1^{-1} \geq_{\varepsilon\text{-}AFSD} F \circ G_2^{-1}.$$

Proof. Without loss of generality, let the single crossing point of G_1 and G_2 be x_0. Define

$$\varphi_1(t) = \int_{-\infty}^{t} \left\{[G_2(x) - G_1(x)]_+ - \frac{\varepsilon}{1-\varepsilon}[G_1(x) - G_2(x)]_+\right\} \mathrm{d}x$$

for $t \in \mathbb{R}$. Then $\varphi_1(t)$ is decreasing on $(-\infty, x_0]$ and increasing on (x_0, ∞). Since $G_2 \geq_{\varepsilon\text{-AFSD}} G_1$, we get $\varphi_1(t) \leq 0$ for any $t \in \mathbb{R}$. It is clear that

$$\begin{aligned}
&\int_0^1 \left\{[F \circ G_1^{-1}(u) - F \circ G_2^{-1}(u)]_+ - \frac{\varepsilon}{1-\varepsilon}[F \circ G_2^{-1}(u) - F \circ G_1^{-1}(u)]_+\right\} \mathrm{d}u \\
&= \int_{G_1(x_0)}^{1} [F \circ G_1^{-1}(u) - F \circ G_2^{-1}(u)]\,\mathrm{d}u - \frac{\varepsilon}{1-\varepsilon}\int_0^{G_1(x_0)} [F \circ G_2^{-1}(u) - F \circ G_1^{-1}(u)]\,\mathrm{d}u \\
&= \int_{x_0}^{\infty} F(x)\,\mathrm{d}[G_1(x) - G_2(x)] - \frac{\varepsilon}{1-\varepsilon}\int_{-\infty}^{x_0} F(x)\,\mathrm{d}[G_2(x) - G_1(x)] \\
&= \int_{x_0}^{\infty} [G_2(x) - G_1(x)]\,\mathrm{d}F(x) - \frac{\varepsilon}{1-\varepsilon}\int_{-\infty}^{x_0} [G_1(x) - G_2(x)]\,\mathrm{d}F(x) \\
&= \int_{-\infty}^{\infty} \left\{[G_2(x) - G_1(x)]_+ - \frac{\varepsilon}{1-\varepsilon}[G_1(x) - G_2(x)]_+\right\} \mathrm{d}F(x) \\
&= \int_{-\infty}^{\infty} F'_+(x)\,\mathrm{d}\varphi_1(x) \\
&= \varphi_1(+\infty)F'_+(+\infty) - \int_{-\infty}^{\infty} \varphi_1(x)\,\mathrm{d}F'(x) \leq 0,
\end{aligned}$$

where the last inequality follows from the decreasing property of F'_+. Therefore, $F \circ G_1^{-1} \geq_{\varepsilon\text{-AFSD}} F \circ G_2^{-1}$ for any $0 < \varepsilon < 1/2$. □

Proposition 4.6. *Let G_i and F_i be cdfs such that $F_i \circ G_j^{-1}$ is also a cdf for any $i, j \in \{1, 2\}$. If G_2 singly crosses G_1 from below, F_1 singly crosses F_2 form below, and F_2 and G_1 are both concave (or F_1 and G_2 are both concave), then for any $0 < \varepsilon < 1/2$,*

$$G_2 \geq_{\varepsilon\text{-}AFSD} G_1, \quad F_1 \geq_{\varepsilon\text{-}AFSD} F_2 \implies F_1 \circ G_1^{-1} \geq_{\varepsilon\text{-}AFSD} F_2 \circ G_2^{-1}.$$

https://doi.org/10.1017/S0269964822000250 Published online by Cambridge University Press

Proof. It suffices to prove that

$$\begin{aligned}
&(1-\varepsilon)\int_0^1 [F_1\circ G_1^{-1}(u) - F_2\circ G_2^{-1}(u)]_+\,\mathrm{d}u - \varepsilon\int_0^1 [F_2\circ G_2^{-1}(u) - F_1\circ G_1^{-1}(u)]_+\,\mathrm{d}u \\
&\quad = (1-2\varepsilon)\int_0^1 [F_1\circ G_1^{-1}(u) - F_2\circ G_2^{-1}(u)]_+\,\mathrm{d}u - \varepsilon\int_0^1 [F_2\circ G_2^{-1}(u) - F_1\circ G_1^{-1}(u)]\,\mathrm{d}u \\
&\quad =: I_1 - I_2 \le 0.
\end{aligned}$$

Since

$$\begin{aligned}
I_1 &= (1-2\varepsilon)\int_0^1 [F_1\circ G_1^{-1}(u) - F_2\circ G_1^{-1}(u) + F_2\circ G_1^{-1}(u) - F_2\circ G_2^{-1}(u)]_+\,\mathrm{d}u \\
&\le (1-2\varepsilon)\int_0^1 [F_1\circ G_1^{-1}(u) - F_2\circ G_1^{-1}(u)]_+\,\mathrm{d}u + (1-2\varepsilon)\int_0^1 [F_2\circ G_1^{-1}(u) - F_2\circ G_2^{-1}(u)]_+\,\mathrm{d}u,
\end{aligned}$$

we have

$$\begin{aligned}
I_1 - I_2 &\le (1-2\varepsilon)\int_0^1 [F_1\circ G_1^{-1}(u) - F_2\circ G_1^{-1}(u)]_+\,\mathrm{d}u - \varepsilon\int_0^1 [F_2\circ G_1^{-1}(u) - F_1\circ G_1^{-1}(u)]\,\mathrm{d}u \\
&\qquad + (1-2\varepsilon)\int_0^1 [F_2\circ G_1^{-1}(u) - F_2\circ G_2^{-1}(u)]_+\,\mathrm{d}u - \varepsilon\int_0^1 [F_2\circ G_2^{-1}(u) - F_2\circ G_1^{-1}(u)] \\
&= (1-\varepsilon)\int_0^1 [F_1\circ G_1^{-1}(u) - F_2\circ G_1^{-1}(u)] + \mathrm{d}u - \varepsilon\int_0^1 [F_2\circ G_1^{-1}(u) - F_1\circ G_1^{-1}(u)]_+\,\mathrm{d}u \\
&\qquad + (1-\varepsilon)\int_0^1 [F_2\circ G_1^{-1}(u) - F_2\circ G_2^{-1}(u)]_+\,\mathrm{d}u - \varepsilon\int_0^1 [F_2\circ G_2^{-1}(u) - F_2\circ G_1^{-1}(u)]_+\,\mathrm{d}u \\
&\le 0,
\end{aligned}$$

where the last inequality follows from Propositions 4.4 and 4.5. This completes the proof. □

We will end this section by presenting two examples to illustrate the applications of Proposition 4.6.

Example 4.1. Assume that $X \sim \mathcal{P}(a,b)$, the Pareto distribution with parameters $a > 0$ and $b > 0$, and with the cdf given by

$$F(x;a,b) = 1 - \left(\frac{b}{x}\right)^a, \quad x > b.$$

Since $F'(x;a,b) = ab^a x^{-a-1} > 0$ and $F''(x;a,b) = -a(a+1)b^a x^{-a-2} < 0$, $F(x)$ is a concave function. Furthermore, assume that $X_1 \sim F(\cdot;a_1,b_1)$ and $X_2 \sim F(\cdot;a_2,b_2)$. Denote $F_i = F(\cdot;a_i,b_i)$ for $i = 1,2$. We claim that if

$$b_1 > b_2 > 0, \quad a_1 > a_2 > 1 \quad \text{and} \quad \frac{a_2 b_2}{1-a_2} > \frac{a_1 b_1}{1-a_1}, \tag{4.1}$$

then $X_1 \ge_{\varepsilon\text{-AFSD}} X_2$ for $\varepsilon(a_1,a_2,b_1,b_2) < \varepsilon < 1/2$, where $\varepsilon(a_1,a_2,b_1,b_2)$ is to be determined later.

Now, we assume that (4.1) holds. First, for $a_1 > a_2 > 0$, it can be checked that F_1 singly crosses F_2 from below with crossing point $x_0 := (b_1^{a_1}/b_2^{a_2})^{\frac{1}{a_1-a_2}}$. Next, note that, for any $0 < \varepsilon < 1/2$, $X_1 \ge_{\varepsilon\text{-AFSD}} X_2$ if and only if

$$\int_{-\infty}^{+\infty} [F_1(x;a_1,b_1) - F_2(x;a_2,b_2)]_+\,\mathrm{d}x \le \frac{\varepsilon}{1-\varepsilon}\int_{-\infty}^{+\infty} [F_2(x;a_2,b_2) - F_1(x;a_1,b_1)]_+\,\mathrm{d}x$$

https://doi.org/10.1017/S0269964822000250 Published online by Cambridge University Press

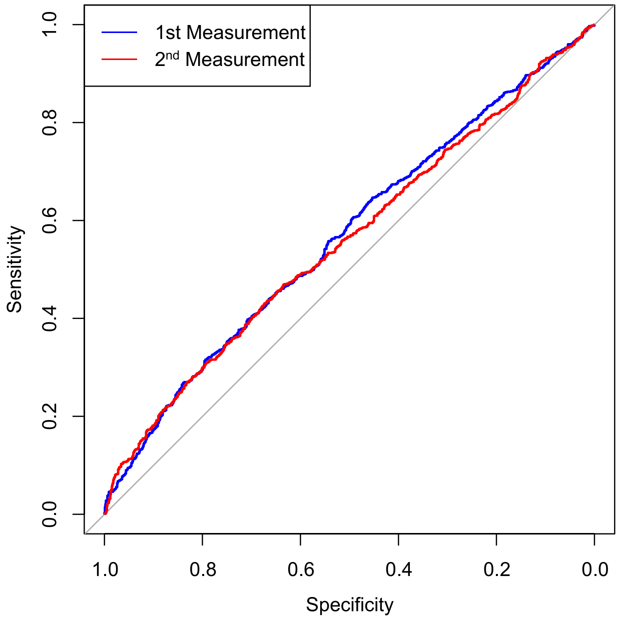

Figure 1. *The ROC curves for the* 1st *and* 2nd *measurements.*

or, equivalently,

$$\int_{x_0}^{+\infty} [F_1(x; a_1, b_1) - F_2(x; a_2, b_2)]\, \mathrm{d}x \le \frac{\varepsilon}{1-\varepsilon} \int_{-\infty}^{x_0} [F_2(x; a_2, b_2) - F_1(x; a_1, b_1)]\, \mathrm{d}x. \tag{4.2}$$

Define

$$A(a_1, a_2, b_1, b_2) := \int_{x_0}^{+\infty} [F_1(x; a_1, b_1) - F_2(x; a_2, b_2)]\, \mathrm{d}x,$$
$$B(a_1, a_2, b_1, b_2) := \int_{-\infty}^{x_0} [F_2(x; a_2, b_2) - F_1(x; a_1, b_1)]\, \mathrm{d}x.$$

It can be checked that

$$A(a_1, a_2, b_1, b_2) = \frac{b_1^{a_1}}{1-a_1} x_0^{1-a_1} - \frac{b_2^{a_2}}{1-a_2} x_0^{1-a_2},$$

and

$$\begin{aligned} B(a_1, a_2, b_1, b_2) &= A(a_1, a_2, b_1, b_2) + b_1 - b_2 + \frac{b_2}{1-a_2} - \frac{b_1}{1-a_1} \\ &= A(a_1, a_2, b_1, b_2) + \frac{a_2 b_2}{1-a_2} - \frac{a_1 b_1}{1-a_1}. \end{aligned}$$

Define

$$\varepsilon(a_1, a_2, b_1, b_2) = \frac{A(a_1, a_2, b_1, b_2)}{A(a_1, a_2, b_1, b_2) + B(a_1, a_2, b_1, b_2)}.$$

Then, $0 < \varepsilon(a_1, a_2, b_1, b_2) < 1/2$ in view of (4.1). So, (4.2) holds if $\varepsilon(a_1, a_2, b_1, b_2) < \varepsilon < 1/2$. This proves $X_1 \ge_{\varepsilon\text{-AFSD}} X_2$.

https://doi.org/10.1017/S0269964822000250 Published online by Cambridge University Press

Example 4.2. Suppose that we have measurements from two diagnostic tests. Let X_i denote the measurement from test i for diseased subject, and let Y_i denote the corresponding measurement on the healthy subject, $i = 1, 2$. Assume that $X_1 \sim \mathcal{P}(8, 3.5)$, $X_2 \sim \mathcal{P}(7, 3)$, $Y_1 \sim \mathcal{P}(4.5, 2.5)$ and $Y_2 \sim \mathcal{P}(5, 3)$. We plot the ROC curves of the two measurements in Figure 1. The two curves intersect with many crossing points. AUC of the 1st measurement is 0.5652 and the 2nd is 0.5655, which are almost the same. Therefore, we can not compare the accuracy of these two diagnostic tests through the classic ROC comparison method. But, we can use AFSD rules to rank the two ROC curves.

To see it, denote $X_i \sim F_i$ and $Y_i \sim G_i$, and let $\varepsilon^* < \varepsilon < 1/2$, where

$$\varepsilon^* = \max\{\varepsilon(8, 7, 3.5, 3),\ \varepsilon(5, 4.5, 3, 2.5)\}.$$

From Example 4.1, we have $X_1 \geq_{\varepsilon\text{-AFSD}} X_2$ and $Y_2 \geq_{\varepsilon\text{-AFSD}} Y_1$, and cdfs F_i and G_i are both concave. By Proposition 4.6, we get $F_1 \cdot G_1^{-1} \geq_{\varepsilon\text{-AFSD}} F_2 \circ G_2^{-1}$, which also implies that the AUC of 1st measurement is smaller than the AUC of 2nd measurement.

Acknowledgments. J. Yang was supported by the NNSF of China (No. 11701518, 12071438), Zhejiang Provincial Natural Science Foundation (No. LQ17A010011) and Zhejiang SCI-TECH University Foundation (No. 16062097-Y). W. Zhuang was supported by the NNSF of China (No. 71971204) and Excellent Youth Foundation of Anhui Scientific Committee (No. 2208085J43).

References

[1] Atkinson, A.B. (2008). More on the measurement of inequality. *Journal of Economic Inequality* 6(3): 277–283.
[2] Balbás, A., Garrido, J., & Mayoral, S. (2009). Properties of distortion risk measures. *Methodology & Computing in Applied Probability* 11(3): 385–399.
[3] Bali, T.G., Demirtas, K.O., Levy, H., & Wolf, A. (2009). Bonds versus stocks: Investors' age and risk taking. *Journal of Monetary Economics* 56(6): 817–830.
[4] Boland, P.J., Hu, T., Shaked, M., & Shanthikumar, J.G. (2002). Stochastic ordering of order statistics II. In M. Dror, P. L'Ecuyer, & F. Szidarovszky (eds), *Modeling uncertainty: An examination of stochastic theory, methods, and applications*. Boston: Kluwer Academic Publishers, pp. 607–623.
[5] Denuit, M., Dhaene, J., Goovaerts, M., & Kaas, R. (2005). *Actuarial theory for dependent risks: Measures, orders and models*. West Sussex: John Wiley & Sons, Ltd.
[6] Donaldson, D. & Weymark, J.A. (1983). Ethically flexible Gini indices for income distributions in continuum. *Journal of Economic Theory* 29(2): 353–358.
[7] Fawcett, T. (2005). An introduction to ROC analysis. *Pattern Recognition Letters* 27(8): 861–874.
[8] Föllmer, H. & Schied, A (2016). *Stochastic finance: An introduction in discrete time*. 4th ed. Berlin: Walter de Gruyter.
[9] Gigliarano, C., Figini, S., & Muliere, P. (2014). Making classifier performance comparisons when ROC curves intersect. *Computational Statistics and Data Analysis* 77: 300–312.
[10] Guo, X., Zhu, X., Wong, W.-K., & Zhu, L. (2013). A note on almost stochastic dominance. *Economics Letters* 121: 252–256.
[11] Guo, X., Post, T., Wong, W.-K., & Zhu, L. (2014). Moment conditions for almost stochastic dominance. *Economics Letters* 124: 163–167.
[12] Guo, D., Hu, Y., Wang, S., & Zhao, L. (2016). Comparing risks with reference points: A stochastic dominance approach. *Insurance: Mathematics and Economics* 70: 105–116.
[13] Hadar, J. & Russell, W. (1969). Rules for ordering uncertain prospects. *American Economic Review* 59: 25–34.
[14] Hanoch, G. & Levy, H. (1969). The efficiency analysis of choice involving risk. *The Review of Economic Studies* 36(3): 335–346.
[15] Kirmani, S. & Gupta, R. (2001). On the proportional odds model in survival analysis. *Annals of the Institute of Statistical Mathematics* 53(2): 203–216.
[16] Lando, T., Arab, I., & Oliveira, P.E. (2021). Second-order stochastic comparisons of order statistics. *Statistics* 55(3): 561–579.
[17] Leshno, M. & Levy, H. (2002). Prefered by all and preferred by most decision makers: Almost stochastic dominance. *Management Science* 48: 1074–1085.
[18] Levy, M. (2012). Almost stochastic dominance and efficient investment sets. *American Journal of Operations Research* 2: 313–321.
[19] Levy, H. (2016). *Stochastic dominance*, 3rd ed. New York: Springer.
[20] Levy, H. & Wiener, Z. (1998). Stochastic dominance and prospect dominance with subjective weighting functions. *Journal of Risk and Uncertainty* 16(2): 147–163.
[21] Levy, H., Leshno, M., & Leibovitch, B. (2010). Economically relevant preferences for all observed epsilon. *Annals of Operations Research* 176: 153–178.
[22] Lusted, L. (1971). Signal detectability and medical decision-making. *Science* 171(3977): 1217–1219.

https://doi.org/10.1017/S0269964822000250 Published online by Cambridge University Press

[23] Muliere, P. & Scarsini, M. (1989). A note on stochastic dominance and inequality measures. *Journal of Economic Theory* 49(2): 314–323.
[24] Müller, A. & Stoyan, D. (2002). *Comparison methods for stochastic models and risks*. Chichester, UK: John Wiley & Sons.
[25] Müller, A., Scarsini, M., Tsetlin, I., & Winkler, R. (2017). Between first- and second-order stochatic dominance. *Management Science* 63: 2933–2974.
[26] Müller, A., Scarsini, M., Tsetlin, I., & Winkler, R. (2021). Ranking distributions when only means and variances are known. *Operations Research*. doi:10.1287/opre.2020.2072
[27] Navarro, J., Ruiz, J.M., & Aguila, Y.D. (2008). Characterizations and ordering properties based on log-odds functions. *Statistics* 42(4): 313–328.
[28] Rothschild, M. & Stiglitz, J.E. (1970). Increasing risk: I. A definition. *Journal of Economic Theory* 2: 225–243.
[29] Sankaran, P.G. & Jayakumar, K. (2008). On proportional odds models. *Statistical Papers* 49(4): 779–789.
[30] Shaked, M. & Shanthikumar, J.G. (2007). *Stochastic orders*. New York: Springer.
[31] Tzeng, L.Y., Huang, R.J., & Shih, P.T. (2013). Revisiting almost second-degree stochastic dominance. *Management Science* 59: 1250–1254.
[32] Wang, S. (1995). Insurance pricing and increased limits ratemaking by propositional hazards transforms. *Insurance: Mathematics and Economics* 17: 43–54.
[33] Wang, S. (2000). A class of distortion operators for pricing financial and insurance risk. *Journal of Risk and Insurance* 67: 15–36.
[34] Wang, S.S. & Young, V.R. (1998). Ordering risks: Expected utility theory versus Yaari's dual theory of risk. *Insurance Mathematics and Economics* 22(2): 145–161.
[35] Yaari, M.E. (1987). The dual theory of choice under risk. *Econometrica* 9: 5–115.
[36] Zimmer, W.J., Wang, Y., & Pathak, P.K. (1998). Log-odds rate and monotone log-odds rate distributions. *Journal of Quality Technology* 30(4): 376–385.

Appendix

Lemma A.1. *Let $B_1 \sim beta(\alpha_1, \beta_1)$ and $B_2 \sim beta(\alpha_2, \beta_2)$ with $\alpha_i > 0$ and $\beta_i > 0$ for $i = 1, 2$. If $\alpha_1 > \alpha_2$ and $\beta_1 > \beta_2$, then F_{B_1} singly crosses F_{B_2} from below.*

Proof. Let f_{B_1} and f_{B_2} be the respective probability density functions of B_1 and B_2, and let F_{B_1} and F_{B_2} be the respective cdfs of B_1 and B_2. Then

$$\ell(x) := \frac{f_{B_1}(x)}{f_{B_2}(x)} = \frac{B(\alpha_2, \beta_2)}{B(\alpha_1, \beta_1)} x^{\alpha_1 - \alpha_2}(1-x)^{\beta_1 - \beta_2}, \quad x \in (0, 1).$$

Taking the derivative of $\ell(x)$, we have

$$\ell'(x) = \frac{B(\alpha_2, \beta_2)}{B(\alpha_1, \beta_1)} x^{\alpha_1 - \alpha_2 - 1}(1-x)^{\beta_1 - \beta_2 - 1}[(\alpha_2 - \alpha_1 + \beta_2 - \beta_1)x + \alpha_1 - \alpha_2].$$

Denote the number of sign changes of the function $a(x)$ in $\mathbb{R}$ by $S^-(a)$. Then, if $\alpha_1 > \alpha_2$ and $\beta_1 > \beta_2$, we have $S^-(\ell'(x)) = 1$ and the sign sequence is $+, -$. This implies that $\ell(x)$ is first increasing and then decreasing with $\ell(0) = \ell(1) = 0$. It is clear that $S^-(f_{B_1} - f_{B_2}) = S^-(\ell - 1) = 2$ and the sign sequence is $-, +, -$. Therefore, $S^-(F_{B_1} - F_{B_2}) = 1$ and the sign sequence is $-, +$. This completes the proof of the lemma. □

Lemma A.2. *Let $B_{i,n} \sim beta(i, n-i+1)$, where $1 \le i \le n$. Then,*

$$\mu_{h_2}(i, n) := \mathbb{E}[h_2(B_{i,n})] = \frac{i}{n-i} \quad \textit{and} \quad \sigma^2_{h_2}(i, n) := \mathrm{Var}(h(B_{i,n})) = \frac{-4i^2 + 3ni + 2n - 2i}{(n-i+1)(n-i)^2},$$

where $h_2(p) = p/(1-p)$.

Proof. Note that

$$\mathbb{E}[h_2(B_{i,n})] = \int_0^1 \frac{y}{1-y} \cdot \frac{y^{i-1}(1-y)^{n-i}}{B(i,n-i+1)}\,\mathrm{d}y = \frac{B(i+1,n-i)}{B(i,n-i+1)} = \frac{i}{n-i}$$

and

$$\mathbb{E}[h_2^2(B_{i,n})] = \int_0^1 \left(\frac{y}{1-y}\right)^2 \frac{y^{i-1}(1-y)^{n-i}}{B(i,n-i+1)}\,\mathrm{d}y = \frac{1}{B(i,n-i+1)}\int_0^1 y^{i+1}(1-y)^{n-i-1}\,\mathrm{d}y = \frac{B(i+2,n-i-1)}{B(i,n-i+1)} = \frac{(i+1)(i+2)}{(n-i+1)(n-i)}.$$

Therefore,

$$\mathrm{Var}\left(h_2(B_{i,n})\right) = \mathbb{E}[h_2^2(B_{i,n})] - [\mathbb{E}(h_2(B_{i,n}))]^2 = \frac{-4i^2+3ni+2n-2i}{(n-i+1)(n-i)^2}.$$

This completes the proof. □

Lemma A.3. *Let $B_{i,n} \sim beta(i,n-i+1)$, $1 \le i \le n$, and define $\psi(x) = \Gamma'(x)/\Gamma(x)$ for $x > 0$.*

(1) *If $h_3(p) = \log[p/(1-p)]$, then $\mu_{h_3}(i,n) := \mathbb{E}[h_3(B_{i,n})] = \psi(i) - \psi(n-i+1)$ and*

$$\sigma^2_{h_3}(i,n) := \mathrm{Var}(h_3(B_{i,n})) = \psi'(i) + \psi'(n-i+1);$$

(2) *If $h_4(p) := -\log(1-p)$, then $\mu_{h_4}(i,n) := \mathbb{E}[h_4(B_{i,n})] = \psi(n+1) - \psi(n-i+1)$ and*

$$\sigma^2_{h_4}(i,n) := \mathrm{Var}(h_4(B_{i,n})) = \psi'(n-i+1) - \psi'(n+1).$$

Proof. (1) Note that the moment generating function of $h_3(B_{i,n})$ is

$$M(\omega) = \mathbb{E}[e^{\omega h_3 \circ B_{i,n}}] = \int_0^1 \left(\frac{t}{1-t}\right)^{\omega} \frac{t^{i-1}(1-t)^{n-i}}{B(i,n-i+1)}\,\mathrm{d}t = \frac{B(i+\omega,n-i+1-\omega)}{B(i,n-i+1)} = \frac{\Gamma(i+\omega)\Gamma(n-i+1-\omega)}{\Gamma(i)\Gamma(n-i+1)}.$$

Then,

$$\mathbb{E}[h_3(B_{i,n})] = M'(0) = \frac{\Gamma'(i)\Gamma(n-i+1) - \Gamma(i)\Gamma'(n-i+1)}{\Gamma(i)\Gamma(n-i+1)} = \psi(i) - \psi(n-i+1)$$

and

$$\mathbb{E}[h_3^2(B_{i,n})] = M''(0) = \frac{\Gamma''(i)\Gamma(n-i+1) - 2\Gamma'(i)\Gamma'(n-i+1) + \Gamma(i)\Gamma''(n-i+1)}{\Gamma(i)\Gamma(n-i+1)}.$$

https://doi.org/10.1017/S0269964822000250 Published online by Cambridge University Press

Therefore,

$$\begin{aligned}
\mathrm{Var}(h_3(B_{i,n})) &= \mathbb{E}[h_3^2(B_{i,n})] - [\mathbb{E}(h_3(B_{i,n}))]^2 \\
&= \frac{\Gamma''(i)\Gamma(n-i+1) - 2\Gamma'(i)\Gamma'(n-i+1) + \Gamma(i)\Gamma''(n-i+1)}{\Gamma(i)\Gamma(n-i+1)} \\
&\quad - \frac{(\Gamma'(i)\Gamma(n-i+1) - \Gamma(i)\Gamma'(n-i+1))^2}{\Gamma^2(i)\Gamma^2(n-i+1)} \\
&= \frac{\Gamma''(i)\Gamma(i) - \Gamma'^2(i)}{\Gamma^2(i)} + \frac{\Gamma''(n-i+1)\Gamma(n-i+1) - \Gamma'^2(n-i+1)}{\Gamma^2(n-i+1)} \\
&= \psi'(i) + \psi'(n-i+1).
\end{aligned}$$

(2) The proof is similar to part (1), and hence is omitted. This completes the proof. □

Cite this article: Yang J, Zhou T and Zhuang W (2023). Almost first-order stochastic dominance by distorted expectations. *Probability in the Engineering and Informational Sciences* **37**, 888–906. https://doi.org/10.1017/S0269964822000250