Weighted Utility and Optimism/Pessimism: A Decision-Theoretic Foundation of Various Stochastic Dominance Orders

By Ehud Lehrer and Tao Wang*

April 26, 2023

Abstract

We show that a probability distribution likelihood-ratio dominates another distribution if and only if, for every weighted utility function, the former is preferred over the latter. Likewise, a probability distribution hazard-rate (or reverse hazard-rate) dominates another distribution if and only if, the former is preferred by every optimistic (or pessimistic) decision maker.

Keywords: Likelihood-ratio dominance; (Reverse) hazard-rate dominance; Weighted utility theory; Optimism and pessimism

JEL Codes: D80, D81.

---

*Lehrer: Department of Economics, Durham University, Mill Hill Lane, Durham DH1 3LB, UK (e-mail: ehud.m.lehrer@durham.ac.uk); Wang (corresponding author): International School of Economics and Management, Capital University of Economics and Business, Beijing 100070, China (e-mail: tao.wang.nau@hotmail.com). We are grateful for the helpful comments from the editor and two anonymous referees. The authors are responsible for any errors. Lehrer acknowledges the support of ISF 591/21 and DFG KA 5609/1-1. Wang acknowledges the support of NSFC Grant #11761141007.

1 Introduction

In this paper, we investigate the decision-theoretic foundations for likelihood-ratio dominance and (reverse) hazard-rate dominance relations. Specifically, we demonstrate that likelihood-ratio dominance is closely related to the weighted utility theory. The weighted utility theory (Chew, 1983) is an important generalization of the expected utility theory. By replacing the controversial independence axiom with a weaker one, called weak substitution axiom, one obtains the weighted utility theory, which accommodates various violations of the expected utility model, such as the Allais' paradox. The weighted utility admits a convenient functional representation: there exists a weight function that assigns weights to different prizes and a monotone utility function over prizes such that a lottery with higher weighted utility is preferred. The weight function could be used to capture a decision maker's perceptive distortion of the likelihood of different prizes when confronted with risk.

We find that for two lotteries F and G, F is preferred to G under every weighted utility if and only if, F likelihood-ratio dominates G. The prize space can be either continuum or finite. The result provides a behavioral interpretation of the well-known likelihood-ratio dominance relation. Our result contributes to the literature that studies the connection between utility theory and stochastic dominance. Early literature (e.g., Quirk & Saposnik, 1962; Hadar & Russell, 1969) focuses on the first- and second-order stochastic dominance and expected utility. However, the behavioral foundation of the likelihood-ratio order has not been studied until Mihm & Siga (2021). In Mihm & Siga (2021), the authors provide a different characterization of the likelihood-ratio order using the *betweenness preference relation* (Dekel, 1986). They show that two lotteries F and G over a *finite* set of prizes satisfy F likelihood-ratio dominates G if and only if, F is preferred to G under every *betweenness preference* relation. It is known that the weighted utility theory generates a strictly smaller class of preferences than the betweenness preferences. In relation to Mihm & Siga (2021), our result implies that one can guarantee likelihood-ratio dominance by using a strictly smaller class of preferences. Moreover, weighted utility allows for a more explicit functional representation than the betweenness preferences.

Weighted utility theory also provides a framework in which optimism/pessimism of the decision maker (henceforth, DM) concerning risk can be conveniently modeled. According to Karni & Schmeidler (1991), a DM is optimistic/pessimistic if he/she has an increasing/decreasing weight function over prizes. It means that when faced with a risky prospect (i.e., a lottery), an optimistic/pessimistic DM subjectively distorts the chances of different prizes, and bases his/her decision on the distorted lottery. The weight function may capture such a distortion: an optimistic DM overestimates the chances of good prizes and underestimates those of the bad prizes. A pessimistic DM would distort the chances in the opposite way.

The second subject we discuss is how optimism and pessimism relate to stochastic orders. We find that every optimistic (resp., pessimistic) DM prefers one lottery F to another one G if and only if, F hazard-rate dominates (resp., reverse hazard-rate dominates) G. The former statement is phrased in behavioral terms, while the latter in probabilistic ones. Our result connects these different notions and therefore provides a behavioral foundation for certain types of stochastic dominance, such as (reverse) hazard-rate order. Hazard-rate dominance and reverse hazard-rate dominance are stronger than first-order stochastic dominance and weaker than likelihood-ratio dominance. To the best of our knowledge, the behavioral aspect of hazard-rated dominance and reverse hazard-rate dominance has not been studied in the previous literature.

Optimism and pessimism have been considered in other settings as well. Related to our approach, the rank-dependent expected utility theory (Quiggin, 1982; Yaari, 1987) provides a different way to model optimism and pessimism. Instead of focusing on the weight function, the rank-dependent expected utility theory focuses on the transformation functions that apply directly to the CDFs of lotteries. In the rank-dependent expected utility framework, pessimism and optimism refer to the convexity and concavity of the transformation function, respectively. Generally speaking, convex/concave transformations of CDFs do not preserve the likelihood-ratio dominance relation and (reverse) hazard-rate dominance relation.[1] By contrast, optimism/pessimism defined via the weight function does have the property of preserving the likelihood-ratio dominance relations.

Preserving the likelihood-ratio dominance relation is a desirable property in the study of conditioning and belief updating. This is because if two prior beliefs satisfy the likelihood-ratio dominance relation, then any posterior beliefs conditional on an event with positive probability (such as a history of noisy signals about the state of nature) will also satisfy the same relation. This property ensures that the process of updating beliefs does not result in a preference reversal. For more examples of the applications of this likelihood-ratio dominance preservation property, refer to

[1]More precisely, let F and G be the CDFs of two lotteries, and let $\phi : [0,1] \to [0,1]$ be an increasing convex function. If F likelihood-ratio dominates G, then it is not necessarily true that $\phi(F)$ likelihood-ratio dominates $\phi(G)$. See Section 4 for further discussions.

Bikhchandani et al. (1992) and Section 5 of our working paper version Lehrer & Wang (2022).

The remaining part of the paper is organized as follows. Section 2 discusses several useful and well-known properties of likelihood-ratio dominance and (reverse) hazard-rate dominance. In Section 3, we provide a decision-theoretic foundation for likelihood-ratio dominance by establishing its connection to the weighted utility theory. We then demonstrate in Section 4 that hazard-rate and reverse hazard-rate dominance relations are related to optimistic and pessimistic attitudes towards risk, respectively. Section 5 concludes. Omitted proofs can be found in the Appendix.

2 A Background on Stochastic Dominance

Likelihood-ratio Dominance. Consider two random variables on $\mathbb{R}$ and let F and G be their cumulative distribution functions (CDF). We say that F *likelihood-ratio dominates* G, and write $F \succeq_{\mathrm{LR}} G$, if the probability measures induced by F and G have densities f and g (Radon-Nikodym derivatives), respectively, w.r.t. some dominating measure (not necessarily the Lebesgue measure) such that $f(x_1)g(x_2) \geq f(x_2)g(x_1)$, for any $x_1 > x_2$.

Note that this definition applies not only to continuous or discrete distributions, but also to general distributions that might have atoms. In fact, if we let $\mathbb{P}_F$ and $\mathbb{P}_G$ be the probability measures induced by F and G, respectively, then they are absolutely continuous[2] w.r.t. $\mathbb{P}_F + \mathbb{P}_G$ (Proposition 3.11, Folland, 1999). It is well-known that likelihood-ratio dominance is stronger than *first-order stochastic dominance* (FOSD).

The following proposition gives a few well-known conditions that are equivalent to likelihood-ratio dominance, expressed in terms of CDFs. They are used subsequently to derive some key results (e.g., Proposition 4).

Proposition 1. [Whitt, 1980; Shaked & Shanthikumar, 2007]. *The following conditions are equivalent to $F \succeq_{\mathrm{LR}} G$:*

1. $[F(x_1) - F(x_2)][G(x_1) - G(x_3)] \geq [F(x_1) - F(x_3)][G(x_1) - G(x_2)]$ *for any* $x_1 > x_2 > x_3$*;*

[2] A measure ν is *absolutely continuous* w.r.t. another measure μ if for any measurable set A, $\mu(A) = 0$ implies $\nu(A) = 0$.

2. $F|A \succeq_{\text{FOSD}} G|A$ for any measurable set A with a positive probability w.r.t. both F and G;

Condition 1 allows us to express LRD in terms of distribution functions. It implies a graphical property of the likelihood-ratio dominance: If $F \succeq_{\text{LR}} G$, then there is a convex function $\varphi : [0,1] \to [0,1]$ such that F can be obtained from G via the convex transformation φ, i.e., $F = \varphi \circ G$. Figure 1 depicts the graph of such a transformation φ. Convexity implies that if we take *any* two points A and B on the graph of φ, then the line segment that connects A and B lies above φ. Condition 2 is the well-known *uniform conditional stochastic order* due to Whitt (1980).

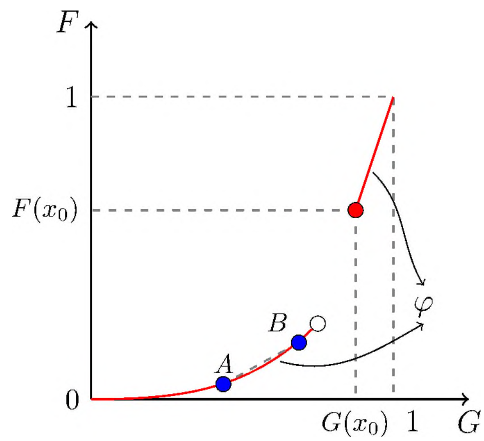


Figure 1: The transformation φ s.t. $F(x) = \varphi(G(p))$ when $F \succeq_{\text{LR}} G$

$F \succeq_{\text{LR}} G$ also implies a certain absolute continuity relation between the probability measures $\mathbb{P}_F$ and $\mathbb{P}_G$. Let $\underline{x} := \inf\{x \in \mathbb{R} | F(x) > 0\}$ and let $D := (\underline{x}, +\infty)$.

Proposition 2. *If $F \succeq_{\text{LR}} G$, then $\mathbb{P}_G(\cdot|D)$ is absolutely continuous w.r.t. $\mathbb{P}_F(\cdot|D)$. Moreover, the Radon-Nikodym derivative $\frac{\mathrm{d}\mathbb{P}_G(\cdot|D)}{\mathrm{d}\mathbb{P}_F(\cdot|D)}$ is non-increasing, except on a set of measure 0 w.r.t. $\mathbb{P}_F(\cdot|D)$.*

The proposition holds when $\underline{x}$ is finite and when $\underline{x} = -\infty$. The proposition implies that if $F \succeq_{\text{LR}} G$, then in the interior of the support[3] of F, an atom of G must also be an atom of F. For instance, in Figure 1, x_0 is an atom of both F and G. There are examples where G has an atom at or before $\underline{x}$, which is not an atom

[3]The support of a probability measure is the smallest closed set with measure 1.

of F. For instance, consider F and G with support $[0,1]$, where $F(x)=x,\ x\in[0,1]$ and $G(x)=\begin{cases}\frac{1}{2}, & \text{if } x=0,\\ \frac{1}{2}+\frac{1}{2}x, & \text{if } x\in(0,1].\end{cases}$

(Reverse) Hazard-Rate Dominance. Consider two CDFs F, G, and let f, g be their Radon-Nikodym derivatives w.r.t. some common dominating measure, not necessarily the Lebesgue measure (cases in which F and G have atoms are also included). We know that F *hazard-rate dominates* G (written $F \succeq_{\text{HR}} G$) if $\frac{f(x)}{1-F(x)} \leq \frac{g(x)}{1-G(x)}$. Similarly, F *reverse hazard-rate dominates* G (written $F \succeq_{\text{RHR}} G$) if $\frac{f(x)}{F(x)} \geq \frac{g(x)}{G(x)}$.

Similar to LRD, hazard-rate dominance (HRD) and reverse hazard-rate dominance (RHRD) also have equivalent conditions phrased in terms of CDFs. The following result, which is the counterpart of Proposition 1, presents two such equivalent conditions.

Proposition 3. [Shaked & Shanthikumar, 2007]. *The following conditions are equivalent to* $F \succeq_{\text{HR}} G$ *(resp.,* $F \succeq_{\text{RHR}} G$*).*

1. *for any* $x_1 > x_2$, $[1-F(x_1)][1-G(x_2)] \geq [1-G(x_1)][1-F(x_2)]$ *(resp.,* $F(x_1)G(x_2) \geq G(x_1)F(x_2)$ *);*

2. *for any* x *such that* $\mathbb{P}_F((x,\infty)) > 0$ *and* $\mathbb{P}_G((x,\infty)) > 0$, $F|(p,\infty) \succeq_{\text{FOSD}} G|(p,\infty)$ *(resp., when* $\mathbb{P}_F((-\infty,x]) > 0$ *and* $\mathbb{P}_G((-\infty,x]) > 0$, $F|(-\infty,x] \succeq_{\text{FOSD}} G|(-\infty,x]$*).*

It is important to note that the second condition of Proposition 3 imposes additional restrictions on the sets being considered compared to Proposition 1. Consequently, the property of HRD does not imply the property of RHRD, and vice versa. In fact, both HRD and RHRD, individually or together, are weaker conditions than likelihood ratio dominance (LRD).

3 Weighted Utility and LRD

Famous stochastic dominance relations, such as FOSD and second-order stochastic dominance (SOSD), admit behavioral equivalent conditions expressed in terms of utility theory. It is well known that FOSD is related to the expected utility theory:

Quirk & Saposnik (1962) and Hadar & Russell (1969) show that a lottery F first-order stochastically dominates another lottery G, if and only if, under every expected utility function, F is preferred to G. In other words, $F \succeq_{\mathrm{FOSD}} G$ if and only if, $\mathbb{E}_F(u) \geq \mathbb{E}_G(u)$ for any increasing function u. Similarly, SOSD is related to risk aversion: A lottery F second-order stochastically dominates another lottery G if and only if, for every risk averse DM, F is preferred to G. More recently, Mihm & Siga (2021) made the insightful discovery that LRD is related to the *betweenness preference* developed by Chew (1983) and Dekel (1986). Specifically, they show that $F \succeq_{\mathrm{LR}} G$ if and only if, F is preferred to G for every betweenness preference.

In this section, we show that Whitt (1980)'s uniform conditional stochastic order (condition 2 of Proposition 1) conveniently implies the following characterization of the LRD order that can be interpreted as comparison of wealth distributions in terms of weighted utility in Chew (1983).[4] The proof is relegated to the Appendix.

Proposition 4. *Let F and G be two distribution functions that concentrate on some interval J of $\mathbb{R}$. Then, $F \succeq_{\mathrm{LR}} G$ if and only if, for every* continuous *and* strictly increasing *function u and every* continuous, positive *function w on J, we have*

$$\frac{\int_J uw \, \mathrm{d}F}{\int_J w \, \mathrm{d}F} \geq \frac{\int_J uw \, \mathrm{d}G}{\int_J w \, \mathrm{d}G}, \tag{1}$$

whenever the integrals exist.

Proposition 4 requires the functions u and w to be continuous. However, it should be noted that the "only if" direction of the proposition can be generalized. Specifically, if $F \succeq_{\mathrm{LR}} G$ holds, then it is true that Eq. (1) holds for every increasing[5] function u and non-negative w, provided that the integrals exist and the denominators are nonzero.

Another way to view the equivalence condition is as follows. Take any positive function w. Define two new CDFs, $\tilde{F}(p) := \frac{\mathbb{E}_F(w\mathbf{1}_{(-\infty,p]})}{\mathbb{E}_F(w)}$ and $\tilde{G}(p) := \frac{\mathbb{E}_G(w\mathbf{1}_{(-\infty,p]})}{\mathbb{E}_G(w)}$. Then for any increasing and measurable function u, $\mathbb{E}_{\tilde{F}}(u) \geq \mathbb{E}_{\tilde{G}}(u)$, or $\tilde{F} \succeq_{\mathrm{FOSD}} \tilde{G}$. It means that first-order stochastic dominance is preserved under change of measure. In the environment of Bayesian learning, this characterization also has natural applications, since the posterior beliefs and conditional expectations can be

[4]To the best of our knowledge, this characterization has not been stated and used explicitly in the literature. But a similar characterization for HRD and RHRD is known. It is due to Capéraà (1988). See Proposition 6.

[5]By increasing, we mean weakly increasing.

expressed in the form $\frac{\mathbb{E}_F(uw)}{\mathbb{E}_F(w)}$, where u and w are functions with certain properties. For an application, see Bikhchandani et al. (1992) and our working paper version Lehrer & Wang (2022).

Proposition 4 has a natural connection with the weighted utility theory (see, e.g., Chew, 1983; Chew, 1985) that generalizes the expected utility theory by replacing the independence axiom with a weaker one called "weak substitution axiom".[6] Let $J \subseteq \mathbb{R}$ be an interval (the space of prizes). Denote by ΔJ the set of all probability distributions (lotteries) on J. Chew (1983) presents a representation result for a preference relation $\succeq$ on ΔJ that satisfies the axioms for weighted utility theory: There exists a continuous function $u : \mathbb{R} \to \mathbb{R}$ that is strictly increasing on J and a continuous function $w : \mathbb{R} \to \mathbb{R}$ that is positive on J such that $\succeq$ can be represented by the functional (the weighted utility function) $V_{u,w} : \Delta J \to \mathbb{R}$ defined by

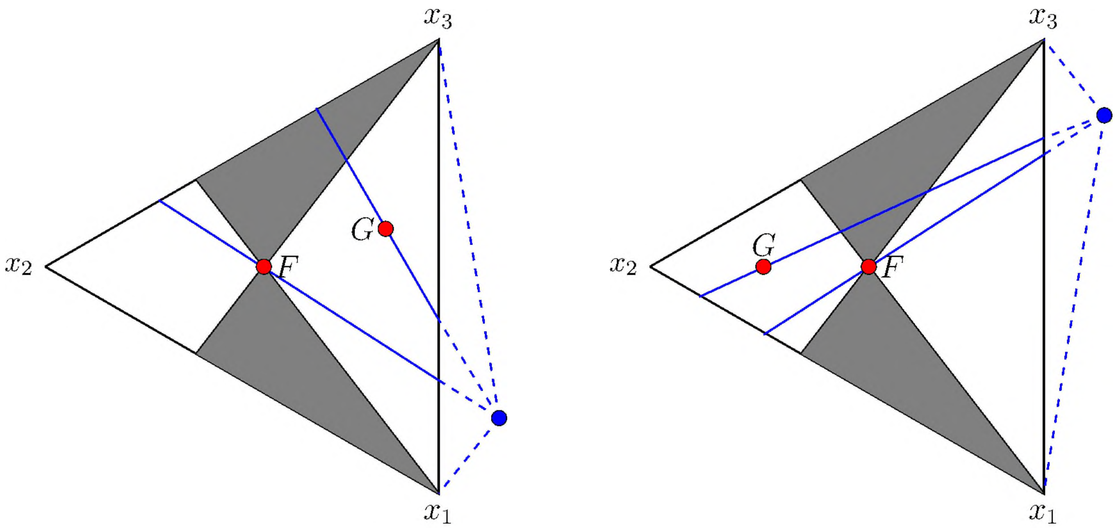
$$V_{u,w}(F) \equiv \frac{\int_J uw \, \mathrm{d}F}{\int_J w \, \mathrm{d}F}. \tag{2}$$

When $w(\cdot)$ is a constant, $V_{u,w}$ reduces to the standard expected utility function. Note that the functional form of the representation is exactly the same as Eq. (1). Hence Proposition 4 yields the main result of this section.

Theorem 1. *Let F and G be two lotteries in ΔJ. Then $F \succeq_{\mathrm{LR}} G$ if and only if, for every weighted utility function $V_{u,w}$, $V_{u,w}(F) \geq V_{u,w}(G)$.*

Theorem 1 therefore provides a decision-theoretic foundation for LRD. Figure 2 is an illustration of the connection between likelihood-ratio dominance and the weighted utility. Consider the case with three prizes $\{x_1, x_2, x_3\}$, where x_1 and x_3 are the worst and best prizes, respectively. Suppose we fix a lottery F in the simplex. The set of lotteries that likelihood-ratio dominates F corresponds to the gray area above F, and the set of lotteries dominated by F in the sense of likelihood-ratio is the gray area beneath F. Now consider the class of weighted utility functions. Clearly, the indifference curves are linear. In addition, Chew (1989) shows that all indifference curves intersect at the same point outside the simplex. Indifference curves closer to x_3 correspond to higher utility levels. To see why greater weighted utility under F implies that $F \succeq_{\mathrm{LR}} G$, suppose, by contradiction, that $F \not\succeq_{\mathrm{LR}} G$. Take any lottery G in the white region inside the simplex in Figure 2. One can always find a weighted utility function (the indifference curves are illustrated in the figure) such that G is preferred to F.

[6]The weak substitution axiom says that for two lotteries F and G, $F \sim G$ implies that for any $\beta \in (0,1)$, there exists $\gamma \in (0,1)$ such that for any other lottery H, $\beta F + (1-\beta)H \sim \gamma G + (1-\gamma)H$.



Panel (a): $F \dot{\succeq}_{\text{LR}} G$ Panel (b): $F \dot{\succeq}_{\text{LR}} G$

Figure 2: A graphical illustration of Theorem 1

Mihm & Siga (2021) establish a different decision-theoretic foundation for LRD. They focus on the *betweenness preference* instead of the weighted utility. The betweenness preference generalizes the expected utility theory by substituting the independence axiom with a weaker one called *betweenness axiom*.[7] A binary relation $\succeq$ is a *betweenness preference* if it satisfies the weak order, mixture continuity and the betweenness axioms. Mihm & Siga (2021) show that when there is a finite number of prizes, two lotteries F, G satisfy $F \succeq_{\text{LR}} G$ if and only if, for every betweenness preference, F is preferred to G.

Since the weighted utility theory satisfies the betweenness axiom, it generates a strictly smaller class of preferences than the betweenness preferences and admits more explicit functional representations. Hence, compared to Mihm & Siga (2021), our Theorem 1 shows that one can guarantee LRD by using a strictly smaller class of preferences. Besides, Mihm & Siga (2021)'s result requires finite support, while our result also covers the continuum support case.

Finally, it should be noted that this paper uses a distinct argument from the one presented in Mihm & Siga (2021). Specifically, the approach in Mihm & Siga (2021) leverages the geometric properties of the likelihood-ratio order in the probability

[7]The axiom states that if $F \succ G$, then $F \succ \alpha F + (1-\alpha)G \succ G$, for any $\alpha \in (0,1)$; and if $F \sim G$, then $F \sim \alpha F + (1-\alpha)G \sim G$, for any $\alpha \in (0,1)$.

simplex. In contrast, the current paper relies on the properties of integrals under the likelihood-ratio order.

4 Optimism/Pessimism and (Reverse) Hazard-Rate Dominance

When DMs are faced with the same risk, they often exhibit different attitudes: some are optimistic and they tend to systematically overestimate the chance of a higher prize over a lower one; some are pessimistic and tend to overestimate the chance of lower prizes. In this subsection, we demonstrate that optimism and pessimism can be conveniently captured in the weighted utility framework. We also consider when one lottery is preferred to another by optimistic/pessimistic DMs.

Again, let $J \subseteq \mathbb{R}$ be the prize space. Consider the weighted utility function representation Eq. (2). Following Karni & Schmeidler (1991) (Section 3.3.10), one can define optimism and pessimism in terms of whether the weight function w in Eq. (2) is increasing or decreasing:[8]

Definition 1. *A* DM *with weighted utility* $V_{u,w} : \Delta J \to \mathbb{R}$ *is* optimistic (pessimistic) *if the weight function* w *is increasing (decreasing).*

The weight function w captures the DM's perceptive distortion when faced with risk. If F is the CDF of the true lottery, then the DM will perceive F as $\tilde{F}$, where

$$\mathrm{d}\tilde{F} = \frac{w}{\int_J w \mathrm{d}F} \,\mathrm{d}F. \tag{3}$$

Since the weight function w is positive, $\tilde{F}$ and F are mutually absolutely continuous. It means that an event has positive chance in terms of the DM's perception if and only if, it has positive chance under the true lottery. When the weight function w is increasing (decreasing) on J, the Radon-Nikodym derivative $\frac{w}{\int_J w \mathrm{d}F}$ is increasing (decreasing), Eq. (3) implies that for every lottery and any two prizes with positive probabilities, the DM will overestimate (underestimate) the chance of the higher prize relative to the lower one. Another implication of Eq. (3) is that for an optimistic DM, the perceived lottery $\tilde{F}$ likelihood-ratio dominates the true lottery F.

[8]To the best of our knowledge, no extant literature has investigated the connection between optimism/pessisim and stochastic dominance in the weighted utility theory framework, which is the focus of our paper.

Definition 1 generates a partial order of DMs' levels of optimism. One DM is said to be more optimistic than another DM, if there exists a positive increasing function w on J such that the perceived lottery of the former can be obtained from the latter's perceived lottery via a transformation using w, as in Eq. (3).

Definition 1 also implies that an optimistic/pessimistic DM's perception preserves LRD and HRD/RHRD relations among true lotteries.

Proposition 5. *Suppose that two lotteries F and G satisfy $F \succeq_{\mathrm{HR}} G$ (or $F \succeq_{\mathrm{LR}} G$). Then, for an optimistic* DM, *the perceived lotteries $\tilde{F}$ and $\tilde{G}$ also satisfy $\tilde{F} \succeq_{\mathrm{HR}} \tilde{G}$ (or $\tilde{F} \succeq_{\mathrm{LR}} \tilde{G}$). Similarly, if $F \succeq_{\mathrm{RHR}} G$ (or $F \succeq_{\mathrm{LR}} G$), then for a pessimistic* DM, *$\tilde{F} \succeq_{\mathrm{RHR}} \tilde{G}$ (or $\tilde{F} \succeq_{\mathrm{LR}} \tilde{G}$).*

To establish the connections between optimism/pessimism and HRD/RHRD, we present a useful characterization of HRD and RHRD.

Proposition 6. [Capéraà, 1988]. *Let F and G be two* CDFs *on J, $J \subseteq \mathbb{R}$. Then, $F \succeq_{\mathrm{RHR}} G$ if and only if,*

$$\frac{\int_J uw \, \mathrm{d}F}{\int_J w \, \mathrm{d}F} \geq \frac{\int_J uw \, \mathrm{d}G}{\int_J w \, \mathrm{d}G} \tag{4}$$

for all functions u and w for which the expectations exist and such that w is positive, continuous and decreasing, *u is continuous and strictly increasing.*

Similarly, $F \succeq_{\mathrm{HR}} G$ if and only if, Eq. *(4) holds for all functions u and w for which the expectations exist and such that w is positive, continuous and* increasing, *u is continuous and strictly increasing.*

This result is analogous to Proposition 4. Compared to Proposition 4, the equivalent conditions for HRD and RHRD impose additional monotonicity requirements for the function w. In Capéraà (1988), the function u is assumed to be increasing instead of continuous and strictly increasing, and w is non-negative instead of continuous and positive. But Proposition 3, combined with a similar proof as that of Proposition 4, demonstrates that the equivalence result also holds under slightly more restrictive conditions, as in Proposition 6.

The characterization provided in Eq. (4) is useful within the Bayesian updating framework. Specifically, it implies that if two prior beliefs F and G satisfy $F \succeq_{\mathrm{HR}} G$ (or $F \succeq_{\mathrm{RHR}} G$), then the posterior beliefs will also preserve the hazard-rate dominance (or reverse hazard-rate dominance) relation, conditional on any event E whose probability increases (or decreases) with the state of nature. In other words,

$F|E \succeq_{\text{HR}} G|E$ (or $F|E \succeq_{\text{RHR}} G|E$). For a formal proof using Proposition 6 and applications of the result, refer to our working paper version Lehrer & Wang (2022).

More importantly, Proposition 6 and Definition 1 provide insight into optimism and pessimism as follows:

Theorem 2. *Let F and G be two lotteries in ΔJ. Then $F \succeq_{\text{HR}} G$ (resp. $F \succeq_{\text{RHR}} G$) if and only if, every optimistic (resp. pessimistic) weighted utility maximizer prefers F to G.*

On the one hand, HRD and RHRD are purely probabilistic relations that capture when one probability distribution is more favorable than another; on the other hand, optimism and pessimism captures DMs' behavioral attitudes when faced with risk. Theorem 2 establishes their connections, and therefore, it provides a behavioral characterization for HRD and RHRD. The theorem is the counterpart of Theorem 1 and the result obtained by Mihm & Siga (2021) regarding LRD.

Related notions of optimism/pessimism. There are other ways to model pessimism and optimism when a DM faces risk, based on the rank-dependent expected utility theory (Quiggin, 1982). A rank-dependent expected utility functional $V : \Delta J \to \mathbb{R}$ is defined for all $F \in \Delta J$ as $V(F) = \int_J u(x) \, \mathrm{d}\phi(F(x))$, where $u : J \to \mathbb{R}$ is increasing, and $\phi : [0, 1] \to [0, 1]$ is a continuous, strictly increasing surjective transformation function. In one approach (e.g., Quiggin, 1993, Section 6.2, Lemma 6.1), a DM is said to be optimistic (resp., pessimistic), if $\phi(y) \leq y$ (resp., $\phi(y) \geq y$) for every $y \in [0, 1]$. In other words, when the DM is optimistic, $\phi(F) \succeq_{\text{FOSD}} F$ for any F. In a different approach (e.g., Yaari, 1987, page 108; Diecidue & Wakker, 2001, Section 4), optimism/pessimism requires the transformation ϕ to be convex/concave.[9] By Proposition 1, it follows that for any lottery F, the transformed lottery $\phi(F)$ satisfies $\phi(F) \succeq_{\text{LR}} F$ for an optimistic DM.

In our paper, as in the second approach, an optimistic DM's perception also satisfies $\tilde{F} \succeq_{\text{LR}} F$, for any F. But the transformations that we consider are different from those in the rank-dependent utility theory. In rank-dependent utility setup, the transformation ϕ applies directly to the CDF, while in the weighted utility setup, a transformation is defined via a weight function w, as in Eq. (3). To see the difference, note that in rank-dependent utility theory, as long as $F(x) = G(x')$, the

[9]Related to this approach, Wakker (1990) characterizes the convexity and concavity of the transformation function ϕ in terms of comonotonicity.

transformed lotteries satisfy $\phi(F(x)) = \phi(G(x'))$. But in our case, $F(x) = G(x')$ does not imply that $\tilde{F}(x) = \tilde{G}(x')$.[10]

Another difference lies in that for a convex transformation ϕ, we may not have $\phi(F) \succeq_{\text{LR}} \phi(G)$. It means that, generally speaking, convex transformations do not preserve the LRD relation.[11] Similarly, HRD relation is not preserved under a convex transformation. By contrast, optimism/pessimism defined via the weight function preserves the LRD and HRD relations (Proposition 5). Preservation of LRD is important when studying conditioning, since if two prior beliefs satisfy LRD, then the posteriors conditional on any event with positive probability also satisfy LRD. It means that updating does not lead to preference reversal. For an application of this property, see our working paper version Lehrer & Wang (2022).

Optimism and pessimism have also been studied in the Savage framework by Dillenberger et al. (2017). They consider the situation in which the probability distributions over states depend on the payoffs that acts generate (i.e., probability distributions are stake-dependent). In their setup, optimism and pessimism are related to concavity and convexity of the transformation applied to the utility function.

5 Conclusion

In this paper, we investigate the behavioral implications of several well-known stochastic dominance relations that compare the magnitude or location of random variables in decision-theoretic framework. These stochastic dominance relations include likelihood-ratio dominance (LRD), hazard-rate dominance (HRD) and reverse

[10] To see an example, consider $x \in [0,1]$, $w(x) = x$, $F(x) = x$ and $G(x) = x^2$. Let $x = 0.16$, $x' = 0.4$. Then $F(x) = G(x') = 0.16$, but $\tilde{F}(x) = \frac{\int_0^{0.16} w(x)\mathrm{d}F(x)}{\int_0^1 w(x)\mathrm{d}F(x)} = (0.16)^2$, while $\tilde{G}(x') = \frac{\int_0^{0.4} w(x)\mathrm{d}G(x)}{\int_0^1 w(x)\mathrm{d}G(x)} = (0.4)^3$.

[11] To see the reason, let $\varphi : [0,1] \to [0,1]$ be the convex transformation s.t. $F(x) = \varphi(G(x))$. Assume that both ϕ and φ are twice continuously differentiable. To check whether $\phi(\varphi(G)) \succeq_{\text{LR}} \phi(G)$, it suffices to check whether $\frac{\phi'(\varphi(G))\varphi'(G)}{\phi'(G)}$ is increasing in x. Consider the derivative of this expression:

$$\text{sign}\left\{\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\phi'(\varphi(G))\varphi'(G)}{\phi'(G)}\right)\right\} = \text{sign}\left\{\phi''(\varphi(G))\phi'(G)[\varphi'(G)]^2 + \phi'(\varphi(G))\varphi''(G)\phi'(G) - \phi''(G)\phi'(\varphi(G))\varphi'(G)\right\}.$$

The second term can be made arbitrarily close to 0 by setting $\varphi''(G)$ sufficiently close to 0. Compare the first and the third terms. The first term is not always greater than the third, since the first term involves a square component.

hazard-rate dominance (RHRD).

In decision-theoretic framework, we show that LRD is related to the weighted utility theory: A lottery F LRD another lottery G if and only if, F is preferred to G under every weighted utility function. This result complements a recent result by Mihm & Siga (2021), which establishes the connection between the LRD relation of lotteries and the betweenness preference relation.

Regarding HRD and RHRD, we find that they are related to DM's optimistic and pessimistic attitudes toward risk in the weighted utility theory framework. When faced with risk, an optimistic/pessimistic DM systematically overestimates (underestimates) the likelihood of higher prize relative to a lower one. We provide a behavioral characterization for the (R)HRD relation by showing that one lottery F HRD/RHRD another lottery G if and only if, F is preferred to G by every optimistic/pessimistic DM.

References

Bikhchandani, S., Segal, U., & Sharma, S. (1992). Stochastic dominance under Bayesian learning. *Journal of Economic Theory*, *56*(2), 352–377.

Capéraà, P. (1988). Tail ordering and asymptotic efficiency of rank tests. *The Annals of Statistics*, *16*(1), 470–478.

Chew, S. H. (1983). A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the Allais paradox. *Econometrica*, *51*(4), 1065–1092.

Chew, S. H. (1985). An axiomatization of rank dependent quasilinear mean generalizing the Gini mean and the quasilinear mean. *Johns Hopkins University Working Paper No. 156*.

Chew, S. H. (1989). Axiomatic utility theories with the betweenness property. *Annals of Operations Research*, *19*(1), 273–298.

Dekel, E. (1986). An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom. *Journal of Economic Theory*, *40*(2), 304–318.

Diecidue, E. & Wakker, P. P. (2001). On the intuition of rank-dependent utility. *Journal of Risk and Uncertainty*, *23*(3), 281–298.

Dillenberger, D., Postlewaite, A., & Rozen, K. (2017). Optimism and pessimism with expected utility. *Journal of the European Economic Association, 15*(5), 1158–1175.

Folland, G. B. (1999). *Real analysis: Modern techniques and their applications.* John Wiley & Sons.

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

Karni, E. & Schmeidler, D. (1991). *Utility theory with uncertainty*, chapter in Handbook of Mathematical Economics, Volume 4, (pp. 1765–1831). North Holland.

Lehrer, E. & Wang, T. (2022). Weighted utility and subjective options: Some behavioral aspects of stochastic dominance. *Working Paper.*

Mihm, M. & Siga, L. (2021). A behavioral characterization of the likelihood ratio order. *American Economic Review: Insights, 3*(3), 353–366.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization, 3*(4), 323–343.

Quiggin, J. (1993). *Generalized expected utility theory: The rank-dependent model.* Kluwer Academic Publishers.

Quirk, J. P. & Saposnik, R. (1962). Admissibility and measurable utility functions. *The Review of Economic Studies, 29*(2), 140–146.

Shaked, M. & Shanthikumar, J. G. (2007). *Stochastic orders.* Springer Science & Business Media.

Wakker, P. (1990). Characterizing optimism and pessimism directly through comonotonicity. *Journal of Economic Theory, 52*(2), 453–463.

Whitt, W. (1980). Uniform conditional stochastic order. *Journal of Applied Probability, 17*(1), 112–123.

Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica, 55*(1), 95–115.

A Appendix

A.1 Proof of Proposition 2

By the definition of LRD, there exists a dominating measure λ on $\mathbb{R}$ such that (i) $\frac{\mathrm{d}F}{\mathrm{d}\lambda} = f$, $\frac{\mathrm{d}G}{\mathrm{d}\lambda} = g$, (ii) there exists a subset A with $\lambda(A) = 0$ such that for any $x' > x''$ in A^C, we have

$$f(x')g(x'') \geq f(x'')g(x'). \tag{5}$$

We show that for any measurable set $B \subseteq D$ such that $\int_B f \mathrm{d}\lambda = 0$, we have $\int_B g \mathrm{d}\lambda = 0$. Suppose the claim does not hold. Then there exists a measurable set $B \subseteq D$ with $\lambda(B) > 0$ such that $\int_B f \mathrm{d}\lambda = 0$, but $\int_B g \mathrm{d}\lambda > 0$. It implies that one can find $x' \in B \bigcap A^C$ (hence $x' > \underline{x}$) such that $g(x') > 0$ and $f(x') = 0$. Note that by definition of $\underline{x}$, $\int_{(\underline{x}, x') \cap A^C} f \mathrm{d}\lambda > 0$, hence one can find $x'' \in (\underline{x}, x') \cap A^C$ such that $f(x'') > 0$. It follows that $f(x'')g(x') > 0$, while $f(x')g(x'') = 0$, which contradicts Eq. (5). This completes the proof that $\mathbb{P}_G(\cdot|D) \ll \mathbb{P}_F(\cdot|D)$.

To show the second part, note that since $\mathbb{P}_G(\cdot|D) \ll \mathbb{P}_F(\cdot|D) \ll \lambda$, by the chain rule for Radon-Nikodym derivative, we have

$$\frac{\mathrm{d}\mathbb{P}_G(\cdot|D)}{\mathrm{d}\lambda} = \frac{\mathbb{P}_G(\cdot|D)}{\mathrm{d}\mathbb{P}_F(\cdot|D)} \frac{\mathrm{d}\mathbb{P}_F(\cdot|D)}{\mathrm{d}\lambda}.$$

Hence $\frac{\mathbb{P}_G(\cdot|D)}{\mathbb{P}_F(\cdot|D)} = \frac{g}{f}$, which is decreasing.

A.2 Proof of Propositions 4

To show the "only if" direction of Proposition 4, take any non-negative measurable function w such that $\mathbb{E}_F(h) > 0$ and $\mathbb{E}_G(h) > 0$. Consider probability measures $\tilde{F}$ and $\tilde{G}$ with $\mathrm{d}\tilde{F} = \frac{w}{\mathbb{E}_F(w)} \mathrm{d}F$ and $\mathrm{d}\tilde{G} = \frac{w}{\mathbb{E}_G(w)} \mathrm{d}G$. It follows from $F \succeq_{\mathrm{LR}} G$ that $\tilde{F} \succeq_{\mathrm{LR}} \tilde{G}$. Thus we have $\mathbb{E}_{\tilde{F}}(u) \geq \mathbb{E}_{\tilde{G}}(u)$ for any measurable and increasing function u, so Eq. (1) holds. The "only if" direction also holds if we require w to be continuous and positive, and u to be continuous and strictly increasing.

Now we prove the "if" direction. Suppose, by contradiction, that $F \not\succeq_{\mathrm{LR}} G$. Then, by condition 1 of Proposition 1, there exist x_1, x_2, x_3 in the interval $J = [\underline{x}, \bar{x}]$ with $x_1 > x_2 > x_3$ such that

$$[F(x_1) - F(x_2)][G(x_1) - G(x_3)] < [G(x_1) - G(x_2)][F(x_1) - F(x_3)]. \tag{6}$$

For a positive constant ε, define

$$u_\varepsilon(x) = \begin{cases} \frac{\varepsilon}{\bar{x}-(x_2+\varepsilon)}[x-(x_2+\varepsilon)]+1, & \text{if } x \in (x_2+\varepsilon, \bar{x}], \\ \frac{1-\varepsilon}{\varepsilon}[x-x_2]+\varepsilon, & \text{if } x \in (x_2, x_2+\varepsilon], \\ \frac{\varepsilon}{x_2-\underline{x}}(x-\underline{x}), & \text{if } x \in [\underline{x}, x_2], \end{cases}$$

$$w_\varepsilon(x) = \begin{cases} \frac{\varepsilon-1}{\varepsilon}[x-(x_1+\varepsilon)]+\varepsilon, & \text{if } x \in (x_1, x_1+\varepsilon], \\ 1, & \text{if } x \in (x_3+\varepsilon, x_1], \\ \frac{1-\varepsilon}{\varepsilon}(x-x_3)+\varepsilon, & \text{if } x \in (x_3, x_3+\varepsilon), \\ \varepsilon, & \text{if } x \in [\underline{x}, x_3] \cup (x_1+\varepsilon, \bar{x}]. \end{cases}$$

For ε sufficiently small, u_ε is continuous and strictly increasing, and w_ε is continuous and positive. As $\varepsilon \to 0$, $u_\varepsilon \to 1_{(x_2,\bar{x}]}$ and $w_\varepsilon \to 1_{(x_3,x_1]}$. By assumption, for every ε, $\mathbb{E}_F(u_\varepsilon w_\varepsilon)\mathbb{E}_G(w_\varepsilon) \geq \mathbb{E}_G(u_\varepsilon w_\varepsilon)\mathbb{E}_F(w_\varepsilon)$. By the dominated convergence theorem, $\lim_{\varepsilon\to 0}\mathbb{E}_F(w_\varepsilon) = F(x_1) - F(x_3)$, $\lim_{\varepsilon\to 0}\mathbb{E}_G(w_\varepsilon) = G(x_1) - G(x_3)$, and that $\lim_{\varepsilon\to 0}\mathbb{E}_F(u_\varepsilon w_\varepsilon) = F(x_1) - F(x_2)$, $\lim_{\varepsilon\to 0}\mathbb{E}_G(u_\varepsilon w_\varepsilon) = G(x_1) - G(x_2)$. It follows that $[F(x_1)-F(x_2)][G(x_1)-G(x_3)] \geq [G(x_1)-G(x_2)][F(x_1)-F(x_3)]$, which contradicts Eq. (6).

This completes the proof of Proposition 4.

A.3 Proof of Proposition 5

We show that for an optimistic DM, $F \succeq_{\text{HR}} G$ implies that $\tilde{F} \succeq_{\text{HR}} \tilde{G}$. By Definition 1, an optimistic DM can be characterized by a continuous and increasing weight function w. For any positive, continuous and increasing function w' and any continuous and strictly increasing function u measurable w.r.t. both F and G, $w'w$ is positive, continuous and increasing, so by Proposition 6,

$$\begin{aligned} \frac{\mathbb{E}_{\tilde{F}}(uw')}{\mathbb{E}_{\tilde{F}}(w')} &= \frac{\mathbb{E}_F(uw'w)}{\mathbb{E}_F(w'w)} \\ &\geq \frac{\mathbb{E}_G(uw'w)}{\mathbb{E}_G(w'w)} \\ &= \frac{\mathbb{E}_{\tilde{G}}(uw')}{\mathbb{E}_{\tilde{G}}(w')}. \end{aligned}$$

It follows that $\tilde{F} \succeq_{\text{HR}} \tilde{G}$. Similarly, by using Proposition 4, one can show that $F \succeq_{\text{LR}} G$ implies that $\tilde{F} \succeq_{\text{LR}} \tilde{G}$. The proof for pessimistic DMs is analogous.

Durham University
Durham Research Online

To cite this article: Lehrer, E., & Wang, T. (in press). Weighted Utility and Optimism/Pessimism: A Decision-Theoretic Foundation of Various Stochastic Dominance Orders. American Economic Journal: Microeconomics

Durham Research Online URL: https://durham-repository.worktribe.com/output/1718281

Copyright statement: This content can be used for non-commercial, personal study.