



# Rational risk-aversion: Good things come to those who weight

Christopher Bottomley<sup>1</sup>  | Timothy Luke Williamson<sup>2</sup> 

<sup>1</sup>Australian Productivity Commission,  
Barton, ACT, Australia

<sup>2</sup>Global Priorities Institute, Faculty of  
Philosophy, University of Oxford, Oxford,  
UK

## Correspondence

Timothy Luke Williamson, Global  
Priorities Institute, Faculty of Philosophy,  
University of Oxford, Trajan House, Mill  
St Oxford, OX2 0DJ, UK.

Email: [timothy.williamson2@philosophy.ox.ac.uk](mailto:timothy.williamson2@philosophy.ox.ac.uk)

## Abstract

No existing normative decision theory adequately handles risk. Expected Utility Theory is overly restrictive in prohibiting a range of reasonable preferences. And theories designed to accommodate such preferences (for example, Buchak's (2013) Risk-Weighted Expected Utility Theory) violate the *Betweenness* axiom, which requires that you are indifferent to randomizing over two options between which you are already indifferent. Betweenness has been overlooked by philosophers, and we argue that it is a compelling normative constraint. Furthermore, neither Expected nor Risk-Weighted Expected Utility Theory allow for *stakes-sensitive* risk-attitudes—they require that risk matters in the same way whether you are gambling for loose change or millions of dollars. We provide a novel normative interpretation of *Weighted-Linear Utility Theory* that solves all of these problems.

## KEYWORDS

Decision theory, Risk aversion, Risk-weighted Expected Utility

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Philosophy and Phenomenological Research* published by Wiley Periodicals LLC on behalf of Philosophy and Phenomenological Research LLC.

## 1 | INTRODUCTION

Existing decision theories fall short when it comes to risk. The dominant view, *Expected Utility Theory* (henceforth EU), prohibits a range of reasonable preferences such as risk-averse preferences in the well-known Allais Paradox. That is, many think that EU violates:

**Condition 1.** Allow for reasonable risk-sensitive preferences.

In response, some philosophers endorse alternatives to EU, such as Lara Buchak's *Risk-Weighted Expected Utility Theory* (henceforth REU). But this theory has a strange consequence. It can require that you pay to randomize over, or to avoid randomizing over, two options between which you are indifferent. That is, REU violates:

**Condition 2.** Satisfy the Betweenness axiom.<sup>1</sup>

Moreover, both EU and REU require that risk matter to you in the same way regardless of the stakes of the decision that you face—your risk-attitudes are insensitive to the magnitude of the utilities you stand to gain or lose. That is, both EU and REU violate:

**Condition 3.** Allow for stakes-sensitive risk-attitudes.

We know of no theory defended in the philosophical literature that satisfies all three conditions—indeed, neither EU nor REU satisfies more than one. In this paper, we defend a view that satisfies all three.<sup>2</sup>

The argument proceeds as follows. After some scene-setting in Section 2, Section 3 defends Betweenness, our second condition.<sup>3</sup> Section 4 outlines a paradigmatic Betweenness-satisfying theory, *Weighted-Linear Utility Theory* (henceforth WLU, first introduced as a descriptive theory in pioneering papers by Chew 1983, 1989). We provide a novel normative interpretation of WLU and argue that it achieves everything we want from a normative decision theory. This serves as proof of concept that Betweenness-satisfying theories can do serious normative work and shows that philosophers have overlooked an important class of views in their theorizing about risk. Section 5 shows that WLU permits stakes-sensitive risk-attitudes in a natural way. Section 6 considers possible counterexamples to Betweenness and argues that they do not undermine our view. We conclude that good things really do come to those who weight (in line with WLU).

## 2 | TAKING RISKS

In order to talk clearly about instrumental rationality, we introduce a formal decision-theoretic framework.

<sup>1</sup> A terminological point: for ease of expression, we often talk about *theories* satisfying (or violating) preference axioms. This is shorthand for those theories' permitting preferences that satisfy (or violate) those axioms.

<sup>2</sup> One caveat: Stefánsson and Bradley (2019) defend a variant of EU that could perhaps be augmented to satisfy all three conditions. We discuss this strategy in contrast to our own and provide some critical comments in Footnotes 48 and 61. We also discuss interpretations of EU that may allow for Condition 3 in Footnote 49.

<sup>3</sup> Condition 1 has been elsewhere defended by Buchak (2013) and Condition 3 by Armendt (2014) and Hájek (2003).

We are interested in which *acts* you may take to achieve your *ends*. Let  $\mathcal{X}$  be a set of outcomes, which we think of as various possible ends (think of these as the propositions of intrinsic concern, or descriptions that include all normatively relevant properties of a potential state of affairs).<sup>4</sup> Since the world is risky, we are concerned with *lotteries*, which are probability distributions over outcomes—call the set of such lotteries  $\mathcal{L}$ . (We here think of *options* as lotteries that you can bring about at will—a simplified account that will do for present purposes.) We think of each outcome as the degenerate lottery that assigns probability 1 to that outcome. We then define the compound lottery  $pA + (1 - p)B$  as the lottery that yields (possibly non-degenerate) lotteries  $A$  with probability  $p$  and  $B$  with probability  $(1 - p)$ . Instrumental rationality concerns your preferences over lotteries. So, let  $\geq$  be a *weak preference* relation over  $\mathcal{L}$ . We assume that each act is weakly preferred to itself. You are indifferent between two lotteries  $A$  and  $B$ , denoted  $A \sim B$ , if and only if  $A \geq B$  and  $B \geq A$ . Similarly, you strictly prefer  $A$  to  $B$ , denoted  $A > B$ , if and only if  $A \geq B$  and  $\neg(B \geq A)$ .<sup>5</sup>

The final ingredient in decision theory is a *utility function*. This is a measure of how much you desire each outcome and so is intrinsically linked to your preferences. A utility function  $u : \mathcal{X} \rightarrow \mathbb{R}$  represents your preferences, in the sense that  $u(X) \geq u(Y)$  if and only if  $X \geq Y$ . Let  $p_A$  be the probability distribution induced by lottery  $A$ . Then if  $u$  is unique (up to positive affine transformation), define the *expected utility* of  $A$  as follows:

$$EU(A) = \sum_{x_i \in \mathcal{X}} p_A(x_i) \cdot u(x_i)$$

EU theorists say that lotteries are ranked in line with their EU.

This raises two questions. First, when is a unique  $u$  guaranteed to exist? Second, under what conditions does an agent maximize EU? Von Neumann and Morgenstern (1953) answer both questions with their celebrated *representation theorem*.<sup>6</sup> You can be represented as maximizing expected utility relative to a unique (up to positive affine transformation) utility function if and only if, for all  $A, B, C \in \mathcal{L}$ , you satisfy:

<sup>4</sup> In distinguishing between acts and ends, we follow von Neumann and Morgenstern (1953). (Though Joyce 1999 makes a similar distinction in a different framework.) Some attempt to resolve the paradoxes we discuss here by adopting a more flexible interpretation of outcomes while maintaining EU. While we return to this point in Section 6, two initial comments are in order: (i) Even someone like Jeffrey with an extremely flexible notion of outcomes takes these paradoxes to be incompatible with EU (see his discussion in (Jeffrey, 1982)). (ii) Even if EU (plus an appropriate theory of outcomes) can solve the paradoxes, that is good news for our project. Our primary goal is to jointly satisfy Betweenness and solve the paradoxes that plague canonical versions of EU.

<sup>5</sup> We make the following assumptions: (i) Each lottery is *finite*, assigning positive probability to a finite number of outcomes. (ii)  $\mathcal{L}$  is convex: if two lotteries are elements of  $\mathcal{L}$ , then so is any compound lottery over them. (iii) We assume the existence of a probability function, though remain neutral about its interpretation. We do not here consider representation theorems that allow for simultaneous elicitation of probability and utility functions (though Grant et al. (2000) provide such a representation for a general Betweenness-satisfying theory). Grant et al.'s key axiom is *Decomposability*, which is the analog of Betweenness in the Savage framework (just as the *Sure-Thing Principle* is the analog of Independence in the Savage framework). For the Savage aficionados, Decomposability captures the idea that for act  $f$  'if the agent is made better off by substituting  $g$  for  $f$  on [event]  $A$ , and she is also made better off by substituting  $g$  for  $f$  on [the complement of  $A$ ], then she unconditionally prefers  $g$  to  $f$ .' (Grant et al., 2000, p. 175) (iv) We set aside cases of Knightian uncertainty in which outcomes lack unique probabilities.

<sup>6</sup> The axioms we use are actually due to Marschak (1950). Von Neumann and Morgenstern's initial presentation is equivalent but less intuitive.

1. **Completeness:** Either  $A \geq B$  or  $B \geq A$ .
2. **Transitivity:** If  $A > B > C$ , then  $A > C$ ; if  $A \geq B \geq C$ , then  $A \geq C$ .
3. **Continuity:** If  $A \geq B \geq C$ , then there exists a  $p$  such that  $B \sim pA + (1 - p)C$ .
4. **Independence:** If  $A \geq B$ , then for all  $p$ ,  $pA + (1 - p)C \geq pB + (1 - p)C$ .

This result establishes an important connection between EU-maximization and qualitative constraints on behavior. Instrumental rationality consists in maximizing expected utility and this means obeying the vNM axioms—so the story goes.

## 2.1 | Taking risks seriously

EU, the Independence axiom in particular, has been subject to extensive criticism.<sup>7</sup> For example, it seems reasonable to violate Independence in the following case:<sup>8</sup>

COMMON-RATIO: Jane likes London but loves Rome. She strictly prefers going to London for sure over taking a risk on an 80% chance of going to Rome (and a 20% chance of going nowhere). On the other hand, she strictly prefers a 20% chance of going to Rome (and an 80% chance of going nowhere) over a 25% chance of going to London (and a 75% chance of going nowhere).

Jane's preferences are easy to rationalize. Given a high probability of going to London, the increased chance of losing out on a holiday altogether by taking a risk on Rome looms large. Jane plays it safe when the background conditions are favorable. On the other hand, when the probability of getting a holiday is low, she is happy to bear more risk and take slightly worse odds on a Roman holiday.<sup>9</sup>

While reasonable, Jane's risk-sensitive preferences violate Independence. Let lotteries between Rome, London and No Holiday be represented by triples of the form  $(p_{\text{Rome}}, p_{\text{London}}, p_{\text{None}})$ . Jane's preferences are then:

$$(1, 0, 0) > (0, 1, 0) \quad (1)$$

$$(0, 1, 0) > (0.8, 0, 0.2) \quad (2)$$

$$(0.2, 0, 0.8) > (0, 0.25, 0.75) \quad (3)$$

<sup>7</sup> By focusing on Independence, we do not thereby endorse the remaining axioms. We take Continuity and Completeness as structural assumptions. Both might reasonably be challenged (e.g., Hare (2010) on Completeness; cases like Pascal's Wager put pressure on Continuity, cf. Hájek (2003)). We are happy to first theorize about agents who meet these structural assumptions and ask later how our theory generalizes.

<sup>8</sup> Hájek (2021) distinguishes between 'rationality' and 'reasonableness', where rationality is a matter of satisfying coherence constraints, while reasonableness is a matter of satisfying some additional substantive constraints. We use reasonableness in a slightly different sense: reasonable preferences are those such that a clear-thinking agent could hold onto them on reflection and so, at least *prima facie*, ought to be permitted by the correct decision theory.

<sup>9</sup> If you are worried that Jane's preferences only seem reasonable because of the certainty effect, it is easy to simultaneously rationalize Jane's preferring a 95% chance of London to a 76% chance of Rome.

Independence requires that Jane’s preferences remain fixed when we substitute in a third lottery. In particular, (2) remains fixed when mixed with a 75% chance of nothing-for-sure:

$$0.25(0, 1, 0) + 0.75(0, 0, 1) > 0.25(0.8, 0, 0.2) + 0.75(0, 0, 1)$$

Which is equivalent to:

$$(0, 0.25, 0.75) > (0.2, 0, 0.8)$$

Which contradicts (3). So, Jane’s reasonable preferences violate Independence and hence EU. The most famous violation of Independence is the ‘Allais Paradox’ (Allais, 1953):

ALLAIS: You are considering a hypothetical choice between *A* and *B* as well as a hypothetical choice between *C* and *D*:

	Tickets 1 – 89	Ticket 90	Tickets 91 – 100
<i>A</i>	\$ 1 million	\$ 1 million	\$ 1 million
<i>B</i>	\$ 1 million	\$ 0	\$ 5 million

	Tickets 1 – 89	Ticket 90	Tickets 91 – 100
<i>C</i>	\$ 0	\$ 1 million	\$ 1 million
<i>D</i>	\$ 0	\$ 0	\$ 5 million

In this case, many have the preferences  $A > B$  and  $D > C$ . These preferences can be rationalized along similar lines to Jane’s: when the odds of becoming a millionaire are high, we are averse to even small increases in the probability of becoming a pauper; when the odds of becoming a millionaire are low, we are happy to accept a small increase in the probability of becoming a pauper in order to secure a non-negligible chance of becoming a super-millionaire. It is perfectly reasonable to play it safe when choosing between the first pair of lotteries while taking a risk when choosing between the second pair.

But again, these preferences violate EU. Independence requires that preferences are fixed when substituting out ‘common consequences’. This means that preferences are settled only by the consequences of receiving a ticket labeled 90-100 in both pairs of lotteries. And since each pair is identical on these tickets, Independence requires that  $A > B$  if and only if  $C > D$ , contrary to the Allais preferences.

A lot of ink has been spilled on the Common-Ratio and Allais preferences. Defenders of EU have suggested various ways of accommodating these preferences within their preferred framework.<sup>10</sup> Though it deserves a lengthier discussion, we side with those who think that EU is

<sup>10</sup> See, among others, Bradley (2017, Section 9.5.1), Broome (1991, Chapter 5), Pettigrew (2015), Stefánsson and Bradley (2019), and Weirich (2020). The most prominent strategy involves *re-description* of outcomes. For example, Broome treats ‘getting nothing’ and ‘getting nothing when you were likely to get a million’ as distinct outcomes. Broadly, we find Buchak’s (2013, Chapter 4) responses compelling. Roughly, an agent can have the Allais preferences without the re-individuated outcomes ‘being legitimate descriptions of the choice problem he takes himself to face.’ (p. 123) We return to this point in Section 6.

incompatible with canonical risk-sensitive preferences. Similarly, there are arguments for Independence.<sup>11</sup> Though this again deserves lengthier discussion, we can here only state that we side with those who think that such arguments fail. Those on the fence can take our argument to be a conditional one: if risk-sensitive preferences are rational, then they ought to be captured in a Betweenness-satisfying framework.

## 2.2 | Implications of risk

So, we reject Independence to allow for risk-sensitivity. But Independence can be decomposed into weaker principles, not all of which need be rejected. Burghart (2020), for example, shows that given a finite outcome space, Independence is equivalent to the conjunction of the following constraints:

**Betweenness:** For all  $A, B \in \mathcal{L}$  and  $p \in [0, 1]$ , if  $A \sim B$ , then  $A \sim pA + (1 - p)B \sim B$ .<sup>12</sup>

**Homotheticity:** Let  $Z$  be the degenerate lottery that assigns probability 1 to the worst outcome in  $\mathcal{X}$ . For all  $A, B \in \mathcal{L}$  and any  $p \in [0, 1]$ , if  $A \sim B$ , then  $pA + (1 - p)Z \sim pB + (1 - p)Z$ .<sup>13</sup>

Intuitively, and as we argue at greater length in Section 3, Betweenness is extremely compelling. Randomizing over two equally desirable lotteries makes things neither better nor worse. And remarkably, Betweenness is compatible with the full range of risk-sensitive preferences considered above.<sup>14</sup>

Homotheticity, like Independence, is a kind of substitution-insensitivity constraint. Independence requires that your preference between  $A$  and  $B$  is preserved under mixtures with all distinct lotteries  $C$ ; Homotheticity only requires that your preference between  $A$  and  $B$  is insensitive to substituting in the worst degenerate lottery. And it turns out that Homotheticity is incompatible with the Common-Ratio preferences (cf. Burghart p. 573). Moreover, the Common-Ratio and Allais preferences show that substitution-sensitivity is rationally permissible. From a normative perspective then, it is unclear why anyone who accepts substitution-sensitivity in those cases would reject it in the special case prescribed by Homotheticity.

So, canonical risk-sensitive preferences show that Independence must go. But they do not speak against all weakenings of Independence, Betweenness in particular. We think this creates a strong *prima facie* case that the correct decision theory satisfies Betweenness.

<sup>11</sup> See, among others, Ahmed (2016), Briggs (2015), Hammond (1988), Pettigrew (2015), and Thoma (2019).

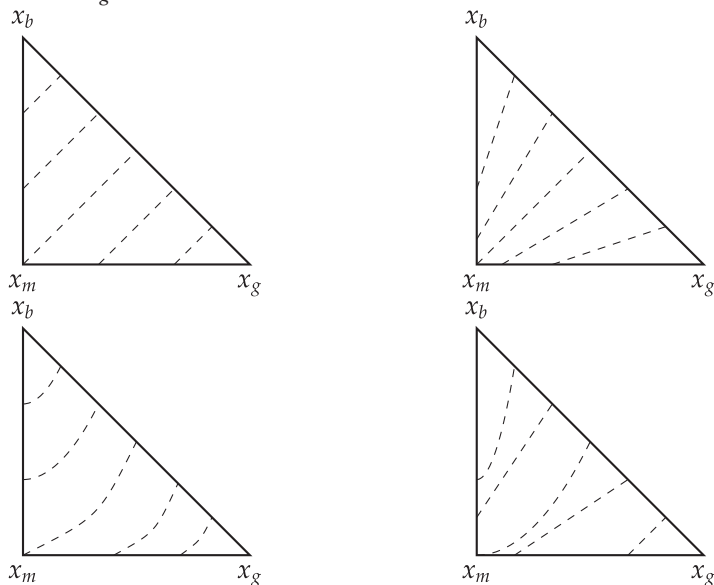
<sup>12</sup> This is Betweenness of Indifference. Note that in the presence of the other axioms, Betweenness of Strong and Weak Preference follows from Betweenness of Indifference.

<sup>13</sup> See Burghart (p. 569) for a discussion of existing descriptive decision theories that satisfy Homotheticity but not Betweenness.

<sup>14</sup> The failure of philosophers to take Betweenness seriously is especially surprising given that von Neumann and Morgenstern (1953) adopted structurally similar axioms to Betweenness (3:B:a and 3:B:b), as did Marschak (1950) (Postulate  $IV_1$ ) and Jeffrey (1983) (the *Averaging* axiom). We take a commitment to Betweenness to cross-cut various debates among EU theorists.

## 2.3 | Triangulating risk

To illustrate the effects of preference axioms, economists often employ *Marschak-Machina Triangles*. Take a set of three ordered outcomes,  $x_g > x_m > x_b$ . Points in the triangle correspond to lotteries (e.g., halfway along the hypotenuse corresponds to the lottery  $.5x_b + .5x_g$ ). Each *indifference curve*, represented by a dotted line, indicates a set of lotteries between which the agent is indifferent. Preference increases as we shift down and to the right of the triangles, toward the best outcome  $x_g$ :



Betweenness requires that there be no premiums or discounts for randomization, forcing indifference curves to be straight (as in the top two triangles). Homotheticity, on the other hand, requires that preferences are invariant to mixing with the worst possible outcome,  $x_b$ . So, if two lotteries  $L_1$  and  $L_2$  lie on the same indifference curve and we construct new lotteries by mixing each with some probability of  $x_b$ , then the new lotteries,  $pL_1 + (1 - p)x_b$  and  $pL_2 + (1 - p)x_b$ , are on the same indifference curve as each other. This corresponds to the indifference curves being parallel (left triangles).

EU satisfies Betweenness and Homotheticity and so requires indifference curves to be straight and parallel, as in the top left diagram. The top right diagram illustrates preferences that satisfy Betweenness, and so are straight, but violate Homotheticity, and so are not parallel.<sup>15</sup> The bottom left illustrates indifference curves that satisfy Homotheticity, and so are parallel, but that violate Betweenness, and so are not all straight.<sup>16</sup> Finally, the bottom right diagram illustrates indifference curves that satisfy neither Betweenness nor Homotheticity, and so are not all straight, nor are they parallel.

<sup>15</sup> These particular indifference curves uniformly ‘fan out’ as more probability mass is concentrated on extremal outcomes. This is characteristic of WLU (see Section 4) but is not entailed by Betweenness itself (see Gul’s (1991) *Disappointment Aversion Theory*).

<sup>16</sup> Such indifference curves result from some rank-dependent theories coupled with a ‘power weighting function’ (Diecidue et al. (2009)). One such theory is Buchak’s REU, outlined below, coupled with  $r(p) = p^2$ . That this function cannot accommodate the Common-Ratio effect shows that it is risk-sensitive only in a narrowly constrained sense.

These diagrams illustrate the space of possibilities for risk-sensitive decision theory. Philosophers have seriously explored the top-left triangle, preserving much structure, and the bottom-right triangle, preserving far less structure, as theories of instrumental rationality.<sup>17</sup> We think that the top-right triangle preserves just the right amount of structure.

## 2.4 | Risk-weighted expected utility

Responding to the shortcomings of EU, Buchak (2013) proposes Risk-Weighted Expected Utility Theory, which allows for risk-sensitivity.<sup>18</sup> In addition to utility and probability, REU introduces a *risk-weighting function*,  $r : [0, 1] \rightarrow [0, 1]$ , that captures risk-attitudes by transforming probabilities. We require that  $r$  be increasing with  $r(0) = 0$  and  $r(1) = 1$ . REU is *rank-dependent*, so we take outcomes to be ordered such that  $x_1 \leq \dots \leq x_n$ . Lotteries are then ranked by their *risk-weighted expected utility*:

$$REU(A) = u(x_1) + \sum_{i=2}^n r\left(\sum_{j=i}^n p_A(x_j)\right) \cdot [u(x_i) - u(x_{i-1})]$$

For each outcome  $x_i$ , REU tells you to weight how much better  $x_i$  is than the next worse outcome by the risk-weighted probability of doing at least as well as  $x_i$ .

Say that  $u(\$x) = \sqrt{x}$  and let  $r(p) = p^2$ . This  $r$  increases the weight given to the worst outcomes in a lottery, rationalizing the Allais preferences:

$$\begin{aligned} REU(A) &= u(\$1m) &&= 1000 \\ REU(B) &= u(\$0) + 0.99^2[u(\$1m) - u(\$0)] + 0.10^2[u(\$5m) - u(\$1m)] &&\approx 992.5 \\ REU(C) &= u(\$0) + 0.11^2[u(\$1m) - u(\$0)] &&\approx 12.1 \\ REU(D) &= u(\$0) + 0.10^2[u(\$5m) - u(\$0)] &&\approx 22.4 \end{aligned}$$

This means that  $A \succ B$  and  $D \succ C$ . In recent years, REU has received a lot of philosophical attention and perhaps deserves to be called the orthodox alternative to orthodoxy for dealing with risk.

## 2.5 | REU violates betweenness

Consider the outcomes  $\{\$1, \$4, \$25\}$ , again with  $u(\$x) = \sqrt{x}$  and  $r(p) = p^2$ . Then each of the following lotteries has REU of 2:

$$\begin{aligned} \text{Gamble A} &: \left(\frac{1}{2}, 0, \frac{1}{2}\right) \\ \text{Gamble B} &: (0, 1, 0) \end{aligned}$$

<sup>17</sup> As mentioned, we might strengthen Buchak's theory to land in the bottom-left by insisting that  $r(p) = p^x$  and  $x \neq 1$ . As far as we know, nobody has defended that strengthened version of REU.

<sup>18</sup> See Buchak Chapter 2 for an exposition of REU. Henceforth, all references to 'Buchak' refer to (2013), unless otherwise specified.

But consider:

Gamble C: A coin flip between Gamble A and Gamble B.

$$\begin{aligned} REU(C) &= REU\left(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right) \\ &= 1 + \left(\frac{1}{2} + \frac{1}{4}\right)^2 [2 - 1] + \left(\frac{1}{4}\right)^2 [5 - 2] \\ &= \frac{7}{4} < 2 \end{aligned}$$

So, REU coupled with a standard risk-weighting function therefore violates Betweenness.<sup>19</sup> Note how odd this is! Imagine being indifferent between *A* and *B* and simplifying your decision by tossing a coin or closing your eyes and picking blindly. REU says that this is impermissible (you effectively choose *C* over *A* or *B*, contra REU). Moreover, note that Gamble *C* is a coin toss between Gamble *A* and its own *certainty equivalent*. So REU tells you that *A*'s risks and benefits are precisely offset by a sure thing of \$4, but then tells you that some probability of *A* or that very certainty equivalent is *worse* than *A*. Introducing some chance of the certainty equivalent intuitively makes things no more risky (indeed, doing so decreases the chance of ending up with the worst outcome). REU, however, coupled with a paradigmatically *risk-averse* *r*, says that it makes things worse. These results strike us as odd and are a good indication that REU's handling of risk isn't quite right—we flesh this point out in the next section.

### 3 | BOXING UP BETWEENNESS

We have our Betweenness-violating preferences for an REU-maximizer. This raises a question—can we provide an argument for Betweenness that does not implicitly appeal to Independence, or reasoning that the risk-sensitive theorist is likely to reject? Such an argument has not been given.<sup>20</sup> Indeed, Camerer and Ho (1994) suggest that no such argument can be given. We address this shortcoming by providing one.<sup>21</sup> As part of a process of reflective equilibrium, Betweenness is more compelling than both Independence and the rival kind of rank-dependent intuitions that underwrite REU.

<sup>19</sup> The more general point is well-known: the only strictly increasing risk-weighting function that satisfies Betweenness is  $r(p) = p$ , in which case REU reduces to EU. An easy way of seeing this is to note Camerer's (1989, pp. 77–78) result that rank-dependent models exhibit hypotenuse parallelism (the tangents to indifference curves along the hypotenuse of any MM-triangle are parallel). Since Betweenness requires that indifference curves be straight, REU coupled with Betweenness results in straight and parallel indifference curves, which is just EU.

<sup>20</sup> Grant et al. (2000) argue that Decomposability (see our Footnote 5) is normatively compelling because it captures the correct interpretation of the informal event-wise reasoning Savage (1954) offered to motivate the Sure-Thing Principle. However, those tempted by Buchak's defense of REU, in particular her discussion (2013, Chapter 1) of *global properties*, will not be motivated to uphold that informal reasoning (on either Savage's or Grant et al.'s interpretation). We take the following discussion to vindicate Betweenness, and hence the informal reasoning that motivated Savage.

<sup>21</sup> Gul and Lantto (1990) show that Betweenness is equivalent to various *prima facie* plausible dynamic constraints. But their constraints are normatively plausible only if we antecedently accept Betweenness as a synchronic constraint. Since our argument is synchronic, it provides the basis needed to underwrite Gul and Lantto's constraints. Steele (2020) discusses cases in which Betweenness-violations lead to further complications in the dynamic setting.

Firstly, note how natural reasoning in line with Betweenness is. In the above case, the reason that  $C$  is worse than \$4 must be because of the possibility of \$1. But  $C$  introduces equal chances of \$1 and \$25. And the initial preference  $A \sim B$  tells us that the agent is prepared to sacrifice \$4 for equal chances of \$1 and \$25. This is our first-pass motivation for endorsing Betweenness: in declaring a certainty equivalent for some gamble, the agent thereby tells us what chances of loss they are prepared to live with for some chance of gain. As we have seen, the mere possibility of risk-sensitivity does not undermine the *prima facie* acceptability of this reasoning—it is a natural starting point as we go about our normative theorizing.

Very well, the opponent of Betweenness might reply. Risk-sensitive preferences tell us that there can be *interaction effects* between outcomes. Just because you judge \$4-for-sure to be as good as a coin-toss on \$1 and \$25, that does not mean that you must judge \$4 to be offset by equal chances of \$1 and \$25 in all gambles. So, you are getting something if you pay to randomize—you are getting a lottery with different global properties.

In response to this concern, consider the following case (like all good arguments in decision theory, it involves opaque and transparent boxes):

THREE BOXES: You get to choose one of the three boxes in front of you. Two are transparent and one is opaque. Transparent box  $A$  contains lottery Ticket  $A$ . Transparent box  $B$  contains lottery Ticket  $B$ . You are indifferent between Ticket  $A$  and Ticket  $B$ . Inside the Opaque Box is another copy of either Ticket  $A$  or Ticket  $B$ , determined by a fair coin-toss.<sup>22</sup>

	Heads	Tails
Box $A$	$A$	$A$
Box $B$	$B$	$B$
Opaque	$A$	$B$

Anyone whose preferences violate Betweenness by, say, dispreferring the Opaque Box owes us some story about why this is permissible.<sup>23</sup>

Recall that we can violate Independence because of interaction effects: your preference between  $A$  and  $B$  is not decisive in determining your preference between  $pA + (1 - p)C$  and  $pB + (1 - p)C$ . This is because a  $(1 - p)$  probability of  $C$  affects how  $pA$  and  $pB$  contribute to your overall evaluation of gambles. In other words, the fact that you could receive  $C$  with probability  $(1 - p)$  affects how you assess the risks and benefits of a  $p$  probability of receiving  $A$  or  $B$ . Independence incorrectly requires that what happens on the  $(1 - p)$ -region of the probability space have no bearing on how you evaluate risks borne on the  $p$ -region of the probability space. Such reasoning might explain why it is rational to violate Betweenness in THREE BOXES—perhaps the fact that you could have either  $A$  or  $B$  affects the contributive value of the other.

<sup>22</sup> We assume the coin is fair but nothing hinges on this.

<sup>23</sup> Of course, THREE BOXES is a single case. And as a referee points out, one example of unreasonable Betweenness-violations may not establish that all such violations are unreasonable. We should emphasize that THREE BOXES is *schematic*: For any Betweenness-violator, we can construct a THREE BOXES case and ask how they would behave in that case, which allows consideration of the standard moves for and against the principle to be explored. And, as we discuss below, we take this schema to show that any Betweenness-violator faces a difficult explanatory challenge.

But such a story is mysterious. Say that you choose the opaque box in THREE BOXES. Then you are guaranteed to get either *A* or *B*, both of which you are prepared to choose to begin with. If the opaque box contains *A*, you get something that you were perfectly happy to take outright. If the opaque box contains *B*, you similarly get something that you were perfectly happy to take outright. There can therefore be no ‘rational dissatisfaction’ on discovering what the box contains—you are simply getting something you were happy to have, while not getting something you were prepared to turn down. Indifference between *A* and *B* means that you do not care how the uncertainty is resolved.

Contrast this with Allais, a paradigmatic case where interaction effects *can* affect your rational preferences. In that case, learning whether one of Tickets 1 – 89 or Tickets 90 – 100 is drawn makes the world of difference. In the first pair of Allais lotteries, learning that one of Tickets 1 – 89 was drawn is great news (you are a millionaire), while in the second pair learning that one of Tickets 1 – 89 was drawn is terrible news (you are a pauper). So, what happens on the 1 – 89 portion of the probability space affects how you evaluate what happens on the rest of the probability space. If 1 – 89 is bad, you might take more of a risk on the remaining tickets. If it is good, you might take less of a risk on those tickets. The crucial feature here is that what happens on Tickets 1 – 89 is either relatively good or relatively bad.

THREE BOXES lacks this feature. Finding out how the coin lands brings no good or bad news. Because you are indifferent between the resolutions of event-wise uncertainty, you have no reason to evaluate one region of probability space differently to another. You are simply finding out which of two options you receive, both of which you are equally happy to have when the other is available. This is why reasoning in line with Betweenness is not a mere covert appeal to Independence.<sup>24</sup>

Let’s say you dig your heels in and insist that the opaque box really is worse than either *A* or *B*. Then you might be reasoning as follows: ‘If I get Ticket *A*, the fact that I could have had *B* with probability .5 makes a difference to the contributive value of .5*A*’. But that is incoherent! You are treating *A* differently because you could have had *B* though you are perfectly happy to take *A* when you could have had *B*. Similarly, you are prepared to take *B* when you could have had *A*, so there is no justification for treating *B* differently given the possibility of *A*. The crucial point is that you are already happy to trade *A* for *B* and vice-versa, so interaction effects which purport to justify Betweenness-violations cannot be explained in the same way as more plausible Independence-violations.

Here, the defender of REU might insist that substituting equivalent lotteries can change the risk profile of a lottery. And of course, if you already buy the rank-dependent account of risk, you will think that randomizing in the cases we have given is rationally significant (say, by affecting the ranking of outcomes).

But this response begs the question. The very issue at stake is which properties of a lottery are rightly called its risk profile. And whether rank matters in the way REU says is precisely what is up for grabs here. Ultimately, we think that normative theorizing requires us to engage in reflective equilibrium. So, the question is whether REU or some Betweenness-satisfying theory represents the better trade-off between principles and intuitions about cases. In light of the discussion so

<sup>24</sup> Here is another way of thinking about the difference. When you violate Independence in the Allais case, it is because learning that some event holds (e.g., that Tickets 1-89 were drawn) tells you that you are facing a risk which is not compensated by a chance of an equally valuable gain. However, in THREE BOXES anything you might learn only tells you that you face risks that are precisely compensated for by chances of gains. So, there is no news-value on learning how the coin lands in THREE BOXES: you do not learn anything that makes a difference to you. Thanks to Alan Hájek for this way of thinking about the reasoning here.

far, we see no reason to privilege REU (or a rank-dependent story more broadly). That theory tells risk-averse agents not to hedge their bets by accepting a chance of the certainty equivalent—counterintuitive. And it prohibits risk-averse agents from tossing a coin to make up their minds when they are indifferent between prospects—also counterintuitive. And, as we will see in Section 5 (and as Armendt (2014) notes independently), it does not let risk-attitudes depend on the magnitudes of the gains and losses you face—highly counterintuitive. Taken together, these considerations speak strongly against taking rank-dependence as our starting point when theorizing about risk.

We think that Betweenness is pretheoretically compelling and that it loses none of its appeal when considering canonical risk-sensitive preferences. We see no reason to reject it, as the defender of REU does. Of course, even if you are unpersuaded, then a weaker conclusion is surely acceptable: it is surely at least *reasonable* to satisfy Betweenness while displaying risk-sensitivity. Risk- and randomization-sensitivity are conceptually distinct phenomena. Therefore, the risk-sensitive agent ought to at least be *allowed* to conform their preferences to Betweenness and reason along the lines we gave in response to THREE BOXES. So, even if you maintain that some agents can rationalize Betweenness-violations, most of what we say in this paper is compatible with a more ecumenical conclusion: the correct decision theory should allow risk-sensitivity while being *compatible* with Betweenness. And that is still an important conclusion—it entails that REU cannot be the correct theory of rationality because it requires randomization-sensitivity to get risk-sensitivity.<sup>25</sup>

## 4 | RISK WITHOUT IRRATIONALITY

So far, so good for Betweenness. But if Betweenness turned out to be incompatible with other important normative constraints, then we would have to reconsider our assessment. However, as we now show, Betweenness is compatible with what many take to be the non-negotiable constraints of decision theory.<sup>26</sup>

Following Buchak (p. 37), any theory that deserves to be called normative must at least:

1. Satisfy First-Order Stochastic Dominance (FOSD).<sup>27</sup>
2. Satisfy Transitivity.
3. Be compatible with Second-Order Stochastic Dominance (SOSD).<sup>28</sup>

<sup>25</sup> Though it would take us too far afield, there are different ways of fleshing this idea out. One is to adopt a pluralist stance on which there are multiple, legitimate normative responses to risk. Or, we might search for an overarching perspective that allows rational agents to reason along rank-dependent or Betweenness-satisfying lines. In either case, Betweenness is still normative in the sense that it is one of the principles that reasonable agents may consistently conform their preferences to. And while that is not the way we go here, it is still a novel position overlooked by philosophers, and the interpretive work we do in the following sections will still play an important role on that approach.

<sup>26</sup> Again, think of the methodology here as a kind of reflective equilibrium: we are seeking to show that Betweenness sits well with the rest of our normative commitments.

<sup>27</sup> For  $A, B \in \mathcal{L}$ ,  $A$  first-order stochastically dominates  $B$  just in case, assuming outcomes are ordered  $x_1 \leq x_2 \leq \dots \leq x_n$ , for all  $i$ :  $\sum_{j=i}^n p_A(x_j) \geq \sum_{j=i}^n p_B(x_j)$ . We say that a theory satisfies First-Order Stochastic Dominance if  $A \geq B$  whenever  $A$  first-order stochastically dominates  $B$ . Strictly speaking, this is *Weak Stochastic Dominance*. *Strong Stochastic Dominance* is more controversial and, while many theories satisfy it, we will not insist on it here.

<sup>28</sup>  $A$  second-order stochastically dominates  $B$  if and only if  $B$  is a mean-preserving spread of  $A$ . That is,  $B$  can be obtained from  $A$  from a sequence of operations that shift pairs of probabilities on either side of the utility-mean farther away, leaving

FOSD captures an uncontroversial sufficient condition for one lottery to be preferable to another. Transitivity is satisfied by all standard theories. We are not aware of any other principles enjoying such wide-spread endorsement.<sup>29</sup> SOSD is more controversial but amounts to compatibility with risk-aversion, so we take it for granted here. Together, we call this package of commitments the *normative core*.

## 4.1 | Weighted-linear utility theory

Betweenness is compatible with the normative core—indeed, it forms the basis of at least one attractive decision theory. To demonstrate this, we outline *Weighted-Linear Utility Theory* (WLU), a paradigmatic Betweenness-satisfying theory first introduced by Chew (1983) as a descriptive economic theory. We show that with appropriate modifications and a novel interpretation (which we give in Section 4.2), it satisfies our normative core.<sup>30</sup> WLU therefore demonstrates that we do not face the kind of trade-off that might require us to abandon Betweenness (though of course others may wish to explore features and interpretations of other Betweenness-satisfying theories).

Like REU, WLU introduces a weighting function in addition to utility and probability functions. Unlike REU, WLU's weighting function takes *outcomes* as its domain,  $w : \mathcal{X} \rightarrow (0, \infty)$ .<sup>31</sup> This weighting function will rationalize risk-sensitivity without transforming probabilities like REU.

For lottery  $A \in \mathcal{L}$ , define  $A$ 's *weighted-linear utility* as:

$$WLU(A) = \sum_{i=1}^n \left( \frac{w(x_i)}{\sum_{j=1}^n w(x_j) p_A(x_j)} \right) p_A(x_i) u(x_i)$$

We will refer to the term  $\frac{w(x_i)}{\sum_{j=1}^n w(x_j) p_A(x_j)}$  as outcome  $x_i$ 's *relative weight* in  $A$ . The relative weight of an outcome is that outcome's weight divided by the (probability-weighted) average weight of the lottery. WLU says that:

$$A \succeq B \text{ if and only if } WLU(A) \geq WLU(B)$$

Think of WLU as weighting each outcome's utility by its probability and its relative weight.<sup>32</sup> An outcome whose relative weight is greater than 1 (i.e., whose weight is greater than the average) will have greater contributive value than its probability-weighted utility, while an outcome whose relative weight is less than 1 will contribute less than its probability-weighted utility. In Section 4.2 we provide an interpretation of  $w$ , but we first demonstrate how WLU rationalizes reasonable risk-sensitivity.

---

the utility-mean unchanged. A theory is compatible with SOSD if it permits  $A \succeq B$  whenever  $A$  second-order stochastically dominates  $B$ . See Stiglitz and Rothschild (1970) for further discussion.

<sup>29</sup> Some reject FOSD when lotteries involve preference gaps over outcomes (e.g., Schoenfield (2014)). And strictly speaking, we wish to remain open to the possibility of Transitivity-violations (e.g., see Williamson (2023)). But those issues would take us too far afield, so we set them aside here and focus on cases where these constraints are satisfied.

<sup>30</sup> For detailed expositions of the formal properties of this theory see Fishburn (1988, 1983).

<sup>31</sup> Implicit in our framework is that indifferent outcomes are assigned equal weight.

<sup>32</sup> When  $w$  is constant WLU reduces to EU. Since we do not impose constraints on  $u$  and  $w$  beyond the normative core, rational risk-neutrality is permitted on our view. You might think that morality can place additional constraints on risk-attitudes (cf. Buchak (2017, 2019)). That is an important issue that we set aside here.

Consider a weight function defined on positive dollar payoffs.<sup>33</sup> Money again has diminishing marginal utility  $u(\$x) = \sqrt{x}$ , and our weight function is everywhere decreasing:

$$w(\$x) = \frac{1}{1 + \sqrt[4]{x}}$$

Intuitively, an agent with this weight function places greater significance on worse outcomes compared to better ones—rags loom large in comparison to riches.

Now consider lotteries involving the outcomes:

$$\{\$0, \$100, \$10\,000, \$20\,000\}$$

Note the contrast between WLU and EU when considering fair coin-tosses on the worst two outcomes:

$$WLU((.5, .5, 0, 0)) \approx 1.94$$

$$EU((.5, .5, 0, 0)) = 5$$

Since  $u(x) = \sqrt{x}$ , the EU-maximizer pays up to \$25 for this lottery, while the WLU-maximizer is significantly more cautious—they only pay \$3.75. On the other hand, contrast WLU and EU in the case of a coin toss on the best outcomes:

$$WLU((0, 0, .5, .5)) \approx 119.07$$

$$EU((0, 0, .5, .5)) \approx 120.71$$

The WLU-maximizer pays up to \$14 177.80 for this lottery, while the EU-maximizer pays only slightly more, \$14 571.07. This illustrates an interesting feature of our chosen weight function: because  $w$  approximates a constant function for large  $x$ , our WLU-maximizer ‘looks like’ an EU-maximizer when the worst-case scenario is good. Things change, however, on incorporating even a small probability of a bad outcome:

$$WLU((.1, 0, .45, .45)) \approx 51.35$$

$$EU((.1, 0, .45, .45)) \approx 108.64$$

The WLU-maximizer dramatically reduces their valuation to \$2 636.35, while the EU-maximizer is far less sensitive and pays up to \$11 802.56. While this particular weight function is merely illustrative, it has some appealing features: a high degree of responsiveness to bad outcomes coupled with an almost risk-neutral attitude towards safe gambles.

And our chosen  $w$  rationalizes the Allais preferences. Recall the Allais case:

<sup>33</sup> See Chew (1983) for a discussion of WLU on more general outcome sets. Note that Chew does not discuss explicit functional forms for  $w$ .

	$u(x)$	$w(x)$	$p_A(x)$	$p_B(x)$	$p_C(x)$	$p_D(x)$
\$0	0	1	0	0.01	0.89	0.9
\$1m	1000	0.031	1	0.89	0.11	0
\$5m	2236.068	0.021	0	0.1	0	0.1

With:

$$WLU(A) = 1000$$

$$WLU(B) = 810.94$$

$$WLU(C) = 3.77$$

$$WLU(D) = 5.13$$

This gives the Allais preferences,  $A > B$  but  $D > C$ , as desired.

From an axiomatic perspective, WLU only requires relaxing Independence from the vNM axioms. Chew proves that there exists a  $u$  and  $w$  (unique up to fractional linear transformation) such that you can be represented as a WLU-maximizer if you satisfy Completeness, Transitivity, Continuity, and the following weakening of Independence:<sup>34</sup>

1. **Weak Independence:** For all  $A, B \in \mathcal{L}$ , if  $A \sim B$ , then for every  $p \in (0, 1)$  there is a  $q \in (0, 1)$  such that for every  $C \in \mathcal{L}$ :  $pA + (1 - p)C \sim qB + (1 - q)C$ .

One way of thinking about Weak Independence is that for any two indifferent lotteries,  $p$ -substitutions of the first are always indifferent to *some* fixed  $q$ -substitution of the second. Independence is the special case where  $q = p$  for all  $p$ . In the presence of the other axioms, Weak Independence trivially entails Betweenness. We will not defend Weak Independence here, in part because it is one of many possible axioms underwriting WLU, and because we wish to leave open other Betweenness-satisfying theories. For now, what matters is that WLU satisfies our normative core, is well-axiomatized, and has some attractive properties. In what follows, we provide a normative interpretation of WLU—it is more than merely formally attractive. Though again, recall that we focus on WLU as proof of concept and leave it for another time to consider normative interpretations of other Betweenness-satisfying theories.<sup>35</sup>

<sup>34</sup> Formally, uniqueness up to fractional linear transformation means that if  $u(\cdot)$  and  $w(\cdot)$  together represent your preferences, then  $u'(\cdot)$  and  $w'(\cdot)$  together also represent your preferences if and only if there exist real numbers  $a, b, c, d$  such that  $u'(\cdot) = au(\cdot) + bw(\cdot)$  and  $w'(\cdot) = cu(\cdot) + dw(\cdot)$ , with  $ad - bc > 0$ . If  $w(\cdot)$  is constant, this amounts to a positive affine transformation of  $u$ , as in EU theory. See Jeffrey (1983) for discussion of a similar relationship between utility and probability in his theory. And see Fishburn (1988, p. 65) for alternative axiomatizations of WLU.

<sup>35</sup> Importantly, there are interesting candidate normative theories that respect Betweenness and not some stronger axiom like Weak Independence. See, for example, Fishburn's (1983, Section 3) *Ordvex Theory*, and Grant et al.'s (2000) *Decomposability-satisfying Theory*. By way of looking forward, we want to raise one complication for those theories. Unlike WLU, they do not typically provide an explicit representation of preferences in terms of cleanly distinguishable desires (represented by  $u$ ) and risk-attitudes (represented by some other function, such as  $w$ ). It may therefore not be possible to interpret such agents in line with familiar folk psychology, as Buchak does for REU and as we do for WLU in Section 4.2. An important question is whether this would constitute an objection to such theories or a complication for the project of folk-psychological interpretation.

One last word about the status of WLU: though we have not given a first-principles argument for Weak Independence, we have not been able to find WLU-violating preferences that have the obvious and intuitive appeal of the Allais preferences.

Before moving on, we must add one constraint to standard WLU. While our choice of  $w$  satisfies FOSD and SOSD, Chew's WLU permits violations of FOSD for some pairs of weight and utility functions.<sup>36</sup> We therefore insist that WLU-maximizers are rational only if they satisfy FOSD (call this WLU\* if you like).<sup>37</sup>

## 4.2 | What are weights?

Weight functions allow us to rationalize risk-sensitivity. But how should we understand  $w$ ? If WLU is to serve as a normative theory, then it had better be more than a mere formal fix. To this end, we interpret  $w$  as a normatively relevant *significance* function.

We begin with two natural interpretations of  $w$  that would disqualify WLU as a normative theory.

Firstly, you might interpret relative weight as measuring the degree to which an outcome has some intrinsically (dis)valuable property. Outcomes with relative weight greater than 1 might induce (intrinsically valuable) elation due to thrill, surprise, and so on. Similarly, outcomes with relative weight less than 1 might induce (intrinsically disvaluable) disappointment due to regret, remorse, anxiety, or so on. We, however, accept that the utility function captures every intrinsically valuable feature of an outcome.<sup>38</sup> Broadly speaking, this means that we commit to a view along the lines sketched by Broome (1991, p. 103), on which two outcomes are distinct just in case if they differ with respect to properties that you rationally care about. Pettit adopts a similar but distinct view (1991) on which two outcomes are distinct just in case they differ with respect to desirable properties. Regardless of how we fill in the details of such views, it is clear that relative weights cannot be interpreted as a measure of some intrinsically valuable property. If they were, then the one outcome could have different relative weights in different lotteries, meaning the one outcome would actually be two outcomes—that is incoherent.

A second interpretation is due to Camerer (1989, p. 60), who interprets relative weights as subjective distortions of probabilities. The WLU-maximizer is therefore guilty of misrepresenting probabilities by treating outcomes with relative weight greater than 1 as more probable than they actually are (and outcomes with relative weights less than 1 as less probable than they actually are). Of course, it may be that global properties do interact with biases such that we sometimes

---

Indeed, we think that WLU strikes a good (and perhaps the best) balance between (i) Permitting reasonable risk-sensitive preferences, while (ii) Structuring and constraining our attitudes toward risk in a coherent framework (moreover, a conservative one that departs minimally from EU). Ultimately, we are inclined to say that WLU represents a stable equilibrium as we jockey between intuitions about principles and cases. Of course, claims about reflective equilibrium are notoriously hard to make precise, and so we hold loosely to that claim here. Indeed, those suspicious of Weak Independence *qua* universal constraint (the kind of thing that all rational agents must obey) should still agree that WLU represents *one* way in which a reasonable Betweenness-satisfier might consistently respond to risk. That makes it a worthy object of normative discussion, which is enough for our purposes here.

<sup>36</sup> Chew (1983) provides conditions under which the WLU-maximizer satisfies FOSD and SOSD. He assumes a smooth preference constraint—in the Appendix, we provide sufficient conditions that do not rely on smooth preferences.

<sup>37</sup> We could either (i) directly restrict the set of admissible functions to those satisfying FOSD, or (ii) introduce an additional preference axiom that captures first-order stochastic dominance such as Stochastic Monotonicity (see Grant et al. (2008)).

<sup>38</sup> Note also Buchak's argument that the Allais preferences should not be rationalized by intrinsic concern for global properties (Section 1.4.1).

misrepresent probabilities. But if such biases are the only explanation for risk-sensitive preferences, then they have no place in an ideal theory of rationality.<sup>39</sup>

So, we cannot think of weights as representing desirable properties of outcomes, nor subjective distortions of probabilities. There is an important analogy with the development of REU here. Prior to Buchak (see pp. 43–47), many interpreted the third component of rank-dependent theories (Buchak's  $r$ ) as a measure of optimism and pessimism.<sup>40</sup> One of Buchak's important innovations is to interpret  $r$  as weighting undistorted probabilities such that each outcome's contributive value depends on something other than its probability and utility. We follow Buchak in interpreting risk-sensitive theories not as involving psychological distortions of probabilities, but as involving a third component distinct from utility and probability.

We therefore think of  $w$  as representing the significance of outcomes: a measure of how much you are concerned with securing some outcomes compared to others.<sup>41</sup> Rational agents can agree on the desirability and probability of outcomes but still disagree about the value of lotteries because they disagree about the relative significance of the outcomes involved.<sup>42</sup> Buchak expresses a similar point (*italics ours*):

...it is plausible to think that some people are more concerned with the worst-case scenario than others, again, for purely instrumental reasons: because they think that

<sup>39</sup> We could imagine a similar argument related to utilities: relative weights measure your disposition to misrepresent values, reasons, and so on. Again, such misrepresentation would be irrational.

<sup>40</sup> For example, Quiggin (1982, p. 333) takes  $r$  to represent an agent's attitudes toward probabilities, rather than as capturing the way that probabilities determine contributive value, as does Buchak.

<sup>41</sup> Of course, you might be tempted to ask why we should think that  $w$  really tracks significance (thanks to a referee for raising this question). Note first that this question only makes sense if we assume some form of realism about decision theory, which we are inclined to do. (Non-realist approaches simply take terms in decision theory to be convenient manners of speaking, meaning there is no further question about the deeper interpretation of theoretical terms like  $w$ .) Indeed, we are inclined to accept a broadly functionalist form of realism: theoretical terms earn their keep by playing the right kind of role in rationalizing and explaining behavior and by their relationship to other theoretical terms. In answer to the question 'Does  $u$  really capture the attitude that we call desire?', we can do no better than note that  $u$  plays the right kind of role to do so. So, the question is whether  $w$  plays the right kind of role to be called significance.

What then is significance? We introduce talk of significance to capture the intuitive idea that the risk-sensitive agent may be more concerned with avoiding some outcomes than others (in a way that goes beyond the mere goodness or desirability of those outcomes). In particular, any attitude that is not reducible to the attitude represented by utility, but that explains why particularly bad outcomes count for more when assessing lotteries, might be called significance. And this is precisely what  $w$  does: as outcomes get better (as captured by  $u$ ), the additional amount that they contribute to the instrumental value of a lottery diminishes (as captured by diminishing  $w$ ). So,  $w$  interacts with  $u$  to explain risk-sensitive preferences, and thus, in virtue of playing the right kind of role, deserves the name significance.

<sup>42</sup> Is it possible to interpret significance itself as a kind of (or component of) goodness (or desirability)? Thanks to a referee for raising this thought. Significance differs from goodness as traditionally understood in three ways: (1) If it is to explain canonical risk-sensitive preferences, it must *decrease* as outcomes get better, (2) It does not affect our willingness to pay for or trade between outcomes, but only makes a difference to how an agent responds to risk, and (3) It contributes to the instrumental value of lotteries in an essentially modal way—the extent to which an outcome's weight matters in some lottery cannot be separated from the probabilities and weights of other outcomes in that lottery. Contrast this with standard accounts of desirability, represented by  $u$ , on which a fixed probability of a fixed utility contributes the same to each lottery, regardless of the probabilities and utilities of the other outcomes in that lottery. So, significance (as captured and explicated by  $w$ ) rationalizes risk-sensitive preferences and behaves in a distinctively modal way: this is different enough from goodness (as usually understood) that it is worth calling it something else. We think it natural to call this 'something else' a risk-attitude, not least because it explains aversion to mean-preserving spreads and other canonical risk-sensitive preferences. Whether it could additionally (or instead) be called a novel kind or component of goodness seems to be a largely terminological point.

*guaranteeing* themselves something of moderate value is a better way to satisfy their general aim of getting some of the things that they value than is making something of very high value merely possible ... Alternatively an agent might be concerned with the best-case scenario: the maximum weighs more heavily in the estimation of a gamble's value than the minimum does, even if these two outcomes are equally likely ... Thus, in addition to having different attitudes towards outcomes and different evaluations of likelihoods, *two agents might have different attitudes towards some way of potentially obtaining some of these outcomes.* (Buchak, p. 49)

On Buchak's view, rational agents can have different attitudes towards various ways of realizing the same goals. We agree: instrumental rationality is permissive enough to allow an outcome with fixed probability and utility to have different contributive values in different lotteries.

While we accept this overall point about the permissiveness of risk-attitudes, we differ on a crucial point with Buchak. For Buchak, an agent's risk-attitude concerns how important it is to avoid the *worst* ranked outcome, seek the *best* ranked outcome, and so on. Terms like 'worst' and 'best' here are purely ordinal—they refer to the ranking of outcomes in the specific lottery under consideration. The WLU-maximizer, however, does not care about avoiding the worst outcome simply because it is the worst outcome in the lottery, seeking the best outcome simply because it is the best outcome in the lottery, and so on. Rather, because *w* takes outcomes as its domain, the relative weight of an outcome depends on the particular outcomes involved in the lottery. Our WLU-maximizer attaches high significance to avoiding bad outcomes relative to good ones—the relative weight of the worst outcome depends on how bad that worst outcome is (and what the other outcomes are), not just on its rank order.<sup>43</sup>

Our disagreement with Buchak therefore boils down to a disagreement about how rational agents structure 'the potential realization of some of [their] aims'. Buchak takes this structuring of the possible realization of ends to depend on the rank-order of those ends, whatever those ends are. It does not matter whether the worst outcome is losing your life or winning \$1 000. By contrast, we think that this structuring of the potential realization of your ends may depend on the specific ends involved in the lottery, not just their rank. We further explore the advantages of this in the next section.

Before moving on note that, setting aside our commitment to Betweenness, there is an independent and important challenge for the defender of REU here. Buchak starts with compelling commitments about folk psychology (that rational agents can have different risk-attitudes which are not reducible to beliefs or desires) and intuitions (that the Allais preferences are rational). She then argues that REU is the correct theory of instrumental rationality, taking the rank-dependent story as a natural fit but without discussing alternatives along the lines defended here. For

<sup>43</sup> Note that Diecidue and Wakker (2001) explicitly defend the intuition behind rank-dependence. One of their defenses is the standard descriptive story that people over- and under-weight probabilities in line with pessimism and optimism. For reasons already given, we set this defense aside given our normative goals. They do, however, give a normative motivation for rank-dependence. They claim that people 'may decide that unfavorable outcomes are especially important in decision making and therefore should receive more attention than equally likely favorable outcomes' (p. 284). But note that this intuition does not specifically support the rank-dependent approach—WLU with a decreasing weight function also captures the idea that 'unfavorable outcomes should receive more attention'. Finally, they claim that the 'intuition of rank-dependence entails that the attention given to an outcome depends not only on the probability of the outcome but also on the favorability of the outcome in comparison to other outcomes' (p. 284). But again note that this intuition is just as well captured by WLU as REU. So, none of the intuitions Diecidue and Wakker provide favor REU, or the rank-dependent approach more generally, over WLU.

example, she says that 'a natural way to interpret ... different attitudes is to postulate that different decision-makers take the fact that they might improve over the minimum to be a more or less important consideration in evaluating a gamble' (p. 49). But our discussion so far shows that the broad folk-psychological commitments that Buchak has, along with the intuitions about cases she endorses, underdetermine the correct theory of rationality.<sup>44</sup> REU says that rational risk-aversion is only a matter of rank-dependent probability weighting. But it seems reasonable to act in line with WLU and take features beyond rank to matter. And it seems reasonable to be risk-sensitive without being sensitive to randomization. And yet REU says that WLU-maximizers are irrational<sup>45</sup> and that risk-sensitivity entails randomization-sensitivity. The defender of REU therefore owes us two stories—one is why rank order matters in the first place, the other is why it is the *only* thing that matters.

We have defended a novel version of WLU. Betweenness is not just a plausible principle in the abstract—it serves as the basis for a compelling decision theory (one that structures our approach to risk in a way that coheres with a well-motivated account and explication of rational attitudes).<sup>46</sup>

## 5 | STAKES-SENSITIVE RISK-ATTITUDES

We began with three desiderata and have so far focused on the first two, allowing reasonable risk-sensitivity while satisfying Betweenness. WLU satisfies our final condition as well:

**Condition 3.** Allow for stakes-sensitive risk-attitudes.

This is remarkable because neither orthodox EU nor popular, recently defended alternatives like REU allow for stakes-sensitive risk-attitudes.<sup>47</sup>

<sup>44</sup> Buchak does (Section 1.4) consider some alternatives to REU. She argues that they either violate plausible normative commitments or fall afoul of a plausible folk psychology. But our re-interpretation of WLU shows that there are further important options to consider.

<sup>45</sup> Except, of course, in the special case that the WLU-maximizer is risk-neutral.

<sup>46</sup> Note that while our interpretation of WLU is explicitly normative, Dekel and Lipman (2010) have argued that descriptive decision theory benefits from rationalization in psychologically plausible, intuitive terms. In light of that, we hope our interpretation will open up avenues for descriptive theorists as well.

<sup>47</sup> There is an interpretation of EU that distinguishes between utility and value, the latter of which is elicited independently of looking at an agent's preferences over risky lotteries. This reading of EU may allow for stakes-sensitive risk-attitudes. But it satisfies Independence and so falls afoul of our [Condition 1](#), which is to allow for reasonable risk-sensitive preferences like Common Ratio and Allais. For our purposes, it will not do. We should here mention an approach to risk-sensitivity recently defended by Orri Stefánsson and Richard Bradley (2017, 2019). Their theory adopts the Jeffrey framework and so satisfies Averaging, which is analogous to Betweenness in cases where the probabilities involved are subjective probabilities. Moreover, they attach values to outcomes and chance-propositions, allowing for risk-sensitivity. Stefánsson and Bradley's view deserves a paper-length treatment of its own. However, an initial comment about the structure of their view may help explain why we do not consider it at length here. Their view is extremely permissive because of the way that they take chance-propositions to be objects of intrinsic concern. If an agent faces risk in the form of an *objective chance distribution* over outcomes, then they may violate both FOSD and Betweenness. (For example, a fair coin toss may be valued above or below any of its outcomes, since the chance distribution the coin toss induces is itself an object of intrinsic concern.) It might be possible to restrict their model—in particular, the manner in which chance-propositions are valued—such that it allows for risk-sensitivity while respecting FOSD and Betweenness. And if they can allow for risk-sensitivity without requiring randomization-sensitivity, then that is certainly an interesting position worth exploring. At that stage we would question the benefit of their model over WLU, but our main goal here is to defend a Betweenness-satisfying approach

It often seems appropriate to take a risk on some low-stakes gamble while refusing to take a risk on a scaled-up version of that same gamble. Many of us would accept a coin-toss that yields a small win or loss without accepting a coin-toss that yields the same win or loss a thousand times over! And generally, you might be risk-averse in high-stakes scenarios (e.g., when buying life insurance) while simultaneously being risk-neutral in low-stakes scenarios (e.g., when buying vegetables).

Formally, we think of stakes-sensitivity in the following way. Let  $L$  be a lottery with outcomes  $l_i$  and  $kL$  be a lottery with outcomes  $kl_i$ , where  $kl_i$  is an outcome of utility  $k$ -times that of  $l_i$ . A theory is stakes-insensitive if it says:

$$L > M \text{ iff } kL > kM$$

REU and EU satisfy this property trivially. WLU, however, does not. To illustrate, consider a scaled down Allais case, in which each outcome has one-thousandth the utility it did in the original case:

	Tickets 1 – 89	Ticket 90	Tickets 91 – 100
$A^*$	\$ 1	\$ 1	\$ 1
$B^*$	\$ 1	\$ 0	\$ 5

	Tickets 1 – 89	Ticket 90	Tickets 91 – 100
$C^*$	\$ 0	\$ 1	\$ 1
$D^*$	\$ 0	\$ 0	\$ 5

	$u(x)$	$w(x)$	$p_A(x)$	$p_B(x)$	$p_C(x)$	$p_D(x)$
\$0	0	1	0	0.01	0.89	0.9
\$1	1	0.5	1	0.89	0.11	0
\$5	2.236	0.400	0	0.1	0	0.1

Our WLU-maximizer evaluates these lotteries as follows:

$$WLU(A^*) = 1$$

$$WLU(B^*) \approx 1.08$$

$$WLU(C^*) \approx 0.06$$

$$WLU(D^*) \approx 0.10$$

This gives the preferences  $B^* > A^*$  and  $D^* > C^*$ , breaking with the original high-stakes Allais preferences.

There are at least two reasons to think that stakes-sensitivity is a fruitful feature of WLU. First, stakes-sensitivity seems reasonable when we consider *specific cases*. Holding fixed an agent's utility function (or using outcomes that arguably have linear utility, such as lives), we have the strong

to risk, not to offer an exclusive defense of one particular model. WLU serves as proof of concept for us, so we leave internecine debates for another day. Note also that a recent model from Hill (2013) allows for stakes-sensitivity, though only in cases of ambiguity as opposed to risk.

intuition that it is permissible, say, to have the Allais preferences without their low-stakes counterparts (or more generally, to turn down some high-stakes lotteries while turning down their low-stakes counterparts; see Armendt (2014) for intuitions along these lines). A mere appeal to diminishing marginal utility cannot capture the Allais preferences in the first place, while REU cannot capture the kind of differential pattern that seems reasonable in such cases.<sup>48</sup>

Second, stakes-sensitivity seems reasonable when we consider *general principles*. Insofar as we care about risk itself, it seems permissible to do so in a way that is informed by, in a general sense, how much we stand to gain or lose (again, see Hájek (2021) for intuitions along these lines).

Indeed, stakes-insensitivity seems to be a substantive constraint. And so, especially in light of intuitions about cases, we should only insist on it given some good argument. But stakes-insensitivity does not seem to be entailed by either the mere idea of risk-sensitivity nor some other incontrovertible axiom. Accordingly, based on intuitions about both cases and principles, and on the grounds that we prefer not to impose substantive constraints without argument, we maintain that WLU's stakes-sensitivity constitutes an advantage over REU's stakes-insensitivity.<sup>49</sup>

WLU allows for natural stakes-sensitivity because  $w$  takes *outcomes* as its domain. By contrast, REU posits a single risk-attitude that is fixed independently of the magnitudes of utilities in the lottery at hand. By itself, relative rank is too coarse-grained to serve as the basis for a complete theory of risk-sensitivity. WLU allows that your concern with, say, with avoiding the worst outcome is informed by how bad that outcome is: you are risk-averse when buying parachutes not simply because death is worse than a broken leg, but because death is particularly bad.

<sup>48</sup> The defender of REU could push back with a kind of debunking story: our intuitions about the stakes-sensitivity of risk-attitudes are entirely explained by diminishing marginal utility (in quantities of goods), which gives us all the stakes-sensitivity we need. Thanks to a referee for raising this point. This move would require an odd kind of non-linearity in the utility function that undercuts much of the motivation for REU. Stakes-sensitivity seems reasonable in cases designed to screen off diminishing marginal utility. For example, say you are gambling over additional days in heaven and believe that days in heaven are so good that they are not subject to saturation effects (feel free to here substitute in any other quantity you like that is plausibly not subject to saturation effects, such as happy lives of equal duration). You are prepared to take a coin toss that yields either 0 or 2 additional days in heaven over 1 day for sure, but you are not prepared to take the risk of a coin toss that yields either 0 or 2 000 000 additional days in heaven over 1 000 000 days for sure. The REU-theorist *could* accommodate these entirely reasonable preferences if they accepted diminishing marginal utility with respect to days in heaven (or happy lives, or whatever else you like). But that would be to abandon the clean separation between risk-attitudes and utility that motivates REU in the first place. Indeed, anyone motivated to introduce risk-attitudes into their theory will have a hard time rationalizing plausible preferences if they only get stakes-sensitivity from diminishing marginal utility.

<sup>49</sup> A possible additional reason in favor of stakes-sensitive risk-attitudes comes from considering the role of risk in EU itself. Note that the interpretation of  $u$  (and its relationship with risk-attitudes) is itself contested among EU theorists. If we simply interpret a concave  $u$  as capturing saturation effects with respect to some good, then the only kind of risk-sensitivity that EU allows for is the kind entirely explained by desires—overly restrictive. But you might wonder whether  $u$  could additionally capture something like significance: perhaps an agent with a concave  $u$  places more weight on avoiding bad outcomes relative to good ones (where goodness would have to be purely ordinal or cardinalized independently of looking at preferences over lotteries). If so, the structure of  $u$  could capture an agent's response to risk itself. Call this the *risk-sensitive interpretation of EU* (see Wilkinson (2022) for discussion). This interpretation satisfies our Condition 2 and gives us one plausible kind of stakes-sensitivity—stakes-sensitivity in some quantity of interest (though not utility). The problem with the risk-sensitive interpretation of EU is that risk-attitudes cannot do the work they ought to do: they do not rationalize canonical risk-sensitive preferences like Allais. The key advantage of WLU over this interpretation of EU then is that we allow for risk-sensitivity *in the right kind of way*, which rationalizes the Allais preferences. Nonetheless, we might think of WLU as capturing the core insights from the risk-sensitive interpretation of EU, but more cleanly separating out our desires and risk-attitudes by introducing weights, and thereby allowing significance to play the full role that it ought to play. The key point is that if you are attracted to the risk-sensitive interpretation of EU, then this furnishes you with an additional argument that risk-attitudes ought be stakes-sensitive: stakes-sensitivity is a feature of the *core view* (EU) and so any generalization (WLU, REU, and so on) ought to preserve that feature.

## 6 | OBJECTIONS TO BETWEENNESS

We conclude by considering a well-known class of cases that are putative counterexamples to Betweenness. Consider the following (cf. Diamond (1967)):<sup>50</sup>

PARENT TRAP: Nour is deciding whether to give an indivisible lolly to Jharna or Jane. They do not prefer giving the lolly to one child over the other, but they do strictly prefer tossing a coin to decide who gets the lolly over giving the lolly to one child outright.

Nour seemingly prefers a lottery over both outcomes to either outcome and so violates Betweenness.<sup>51</sup>

We do not think that such cases undermine a commitment to Betweenness. To see this, note first that cases like Nour's are putative counterexamples to *any* theory that satisfies the normative core, including EU and REU. After all, PARENT TRAP seemingly involves a violation of Betweenness over degenerate lotteries, and trivially any theory that satisfies the normative core also satisfies Betweenness for degenerate lotteries. So, the lesson from PARENT TRAP is not that randomization-sensitivity is the best way of rationalizing risk-sensitive preferences, nor that Betweenness-violations *of the kind* REU generates are rational.<sup>52</sup> If PARENT TRAP does show anything, it is that any theory that upholds a minimal dominance constraint must go. This is an important debate, but it is orthogonal to the main debate we have been engaged in here—the structure of rational risk-sensitivity.<sup>53</sup>

<sup>50</sup> Diamond did not introduce such cases as explicit counterexamples to Betweenness. To preempt what follows, this indicates that Diamond's case does not put pressure specifically on Betweenness, but on a much wider class of views that satisfy a minimal Dominance constraint.

<sup>51</sup> There are readings of cases like PARENT TRAP that immediately rule out Betweenness-violations. Such cases often seem to involve the hallmarks of *incommensurability*, such as insensitivity to mild sweetening (cf. Hare (2010)). And Betweenness says nothing about randomization over incommensurable outcomes. For the sake of argument we set such readings aside and assume commensurability here.

<sup>52</sup> In fact, while cases like PARENT TRAP create problems for everyone, they may actually present *worse* problems for REU than other theories. After all, a risk-averse REU-maximizer has a preference against randomization, while canonical Diamond-cases, if they involve Betweenness-violations, would involve a preference *for* randomization. Thanks to a referee for this point.

<sup>53</sup> Another important but orthogonal question concerns whether ambiguity-averse agents violate Betweenness. Suppose one of Urns *A* and *B* contains all red balls and the other contains all black balls, but you have no other information. Consider a bet that yields \$1 if red is drawn, \$0 otherwise. Given your ignorance, you might reason that you should be indifferent between drawing from either urn. Nonetheless, you might prefer to toss a coin to decide which urn to draw from, since doing so guarantees a 50% chance of winning (something not guaranteed by drawing from either urn outright). This may seem like a reasonable Betweenness-violation.

But such cases do not raise problems for Betweenness. Chiefly, Betweenness here is a constraint on preferences over lotteries, which are precise probability distributions over outcomes. And if the above preferences are rational, it is because the urn case involves ambiguity, so options are not represented as lotteries (rather they induce something like an imprecise probability distribution over outcomes). And Betweenness says nothing about such cases.

Of course, you might wonder whether sensitivity to randomization in ambiguity-cases puts pressure on Betweenness in the precise setting. We think not.

Many (ourselves included) want to model ambiguity cases with indeterminate (or incomplete) preferences (see Rinard (2015)). This indeterminacy (or incompleteness) arises because you are not in a position to assign a unique, precise probability distribution to the outcomes of a single urn. And randomization resolves this indeterminacy, rather than breaking

Since PARENT TRAP is everyone's problem, we can draw on conclusions from a long-running discussion about how to model such cases. In particular, we lean towards a view on which randomizing can change the outcomes available to the decision-maker (so there is no genuine violation of the normative core). This is defended by Buchak (Chapter 4), who draws on Pettit (1991) (see also Broome (1991)).

Briefly, the key move is to ask why Nour prefers the coin-toss. The best explanation is that the coin-toss allows Nour to realize outcomes that they could not have realized without randomization. Say that tossing the coin eliminates the perception of favoritism.<sup>54</sup> Then without the coin toss, the outcomes are 'Give Jharna the lolly (and show favoritism to Jharna)' and 'Give Jane the lolly (and show favoritism to Jane)'. And with the coin-toss, the outcomes are 'Give Jharna the lolly (and show no favoritism to Jharna)' and 'Give Jane the lolly (and show no favoritism to Jane)'.<sup>55</sup> So, the outcomes Nour may realize depend on whether or not they toss the coin. For Nour,  $A \sim B$  and they prefer to randomize only in the sense that there exist distinct outcomes  $A'$  and  $B'$  such that  $pA' + (1 - p)B' > A$ . And of course Betweenness permits such preferences (recall that Betweenness, like all preference axioms, only applies to preferences over lotteries with fully described outcomes).

Indeed, here is an argument for explaining Nour's preferences with change-of-outcomes, and not a Betweenness-violation. Most people, we suspect, in Nour's situation would not pay to re-randomize, and justifiably so. (Nour gets everything they need with a single coin toss). But if Nour's preferences really do result from a Betweenness-violation, then they would be required to re-randomize. Simply note that—setting aside preference changes—the Betweenness-violator still prefers the randomized option on finding out how the coin-toss is resolved. And while we will not make empirical claims here about the structure of all agents' preferences in all cases similar to PARENT TRAP, we think it likely that many cases that involve a preference for randomization do not involve a similar preference for re-randomization. And this puts pressure on the idea that Betweenness-violations really are involved in such cases.

If someone genuinely is indifferent between outcomes, pays to re-randomize, and yet steadfastly denies that randomization allows them to access new outcomes, then we have no intuition that such an agent is rational. And we feel no pressure to accommodate such a rationally objectionable fetish for coin-tosses.<sup>56</sup>

---

indifference. For example, note the common move of modeling ambiguity with sets of admissible precisifications of your attitudes (e.g., sets of EU, REU or WLU functions). Randomization in the urn case results in a precise probability distribution over balls, and so resolves disagreement within this set of admissible precisifications. But this says nothing about the structure of the precisifications themselves, and in particular whether or not they should satisfy Betweenness. So, we should not conclude that randomization may break indifference in risky cases simply because it may resolve indeterminacy in cases of ambiguity.

<sup>54</sup> Of course, favoritism is only one possible explanation. You might think of this case instead through the lens of fairness, inequality, and so on. Crucially, each of these are well-understood, normatively salient properties, and so the approach here does not permit us to re-individuate outcomes however we like.

<sup>55</sup> Note that when we include outcomes described at this level of detail, we end up with *non-practical* preferences. For example, we must be able to talk about the lottery that yields 'Treat Jharna fairly' with probability 1, though you could never practically face such a lottery. We follow Broome (2009) here and do not take this to be a problem: we can make sense of claims like 'It is preferable to treat Jharna fairly for sure over treating Jane unfairly for sure', even if we could never face this choice in practice (see Broome (2009, Sections 6–8) for discussion).

<sup>56</sup> Non-ideal agents might have pragmatic reasons to randomize. Icard (2021) proves some important results about the value of randomization for agents with bounded memory. Such agents are not the focus of this paper, so we set that discussion aside. Ideal agents might have non-trivial credences in their own acts, particularly in cases of strategic interaction (as

At this point, you might worry that there is a tension in our approach between (i) Accepting that everyone must fine-grain outcomes to some degree in order to accommodate Diamond-cases, and (ii) Denying that outcomes may be fine-grained in a way that reconciles Allais with EU. Indeed, you might hope that a complete and mature theory of outcome individuation will dissolve the Allais paradox on terms favorable to EU.<sup>57</sup>

Of course, we cannot arbitrarily individuate outcomes on pain of trivializing decision theory (simply note that maximally fine-grained individuation would allow outcomes to change whenever they appear in a new lottery, and this would undermine any attempt to constrain the instrumental utility of lotteries based on the utilities of their outcomes; see Broome (2009, 1991) and Buchak Section 4.4). And we are sceptical that the substantive constraints on individuation in our complete and mature theory will allow the EU theorist to rationalize all reasonable risk-sensitivity.

First, note that all theories agree that outcomes count as distinct just in case they differ in some normatively significant way (though this may be understood in terms of reasons, desirable properties, what it is rational to care about, and so on). So, a good strategy for determining whether two outcomes are distinct is to ask whether they differ in a way that can be described using the language of some plausible normative theory. In Diamond-cases, the answer is clearly *yes*: fairness and similar concepts abound in plausible normative theories.

But what of Allais? Because that case involves only one person, it is clear that we cannot straightforwardly invoke the kinds of concepts that we invoked in Diamond-cases, such as fairness. The most obvious way then of reconciling EU is to introduce terms like ‘regret’ or ‘disappointment’ into outcome descriptions. But this approach faces two problems. First as Buchak (Section 1.3) notes, it does not seem to track the phenomenology of risk-taking. And while we should be careful about relying too much on introspection here, it seems that if we are to introduce subjective components to outcomes like regret or disappointment, we should give some weight to people’s subjective judgements as to whether they experience such emotions (in a way that fully explains their preferences).

Second, and more worryingly, a mere appeal to emotions and the like as the basis for outcome individuation does not avoid the charge of triviality. After all, if emotions fully explain the Allais preferences, then rational agents must be *very* emotional. Indeed, the amount an agent would need to be willing to pay to avoid regret seems implausibly high.<sup>58</sup> So, an appeal to emotions as the basis for re-individuation in Allais appears to be unconstrained—they appear to be little more than a black box. This means that we have made no progress towards resolving the basic dilemma for reconciling EU and Allais. Either emotions allow EU to accommodate all versions of Allais, in which case they play an unconstrained role and our theory retreats to triviality. Or we place sensible values on emotions, informed at least in part by introspection, in which case we doubt that all versions of Allais can be explained within the EU framework.

Certainly, none of this is proof that there is no reasonable theory of outcome individuation that does the work that the Allais-sympathetic EU theorist wants. But we hope we have said enough

in Skyrms (1990)). This is not a case of ideal agents preferring to randomize, but of reasoning their way to non-trivial credences in their own acts. Indeed, Skyrms assumes EU so satisfies Betweenness.

<sup>57</sup> Thanks to an anonymous referee for pressing this point.

<sup>58</sup> For example, say that in our original Allais case,  $A > B$ .  $B$  is better in monetary expectation than  $A$  by \$390 000. And indeed, if we scale up the case, then  $B$  may be better than  $A$  by an enormous amount. So, a .01 chance of regret in  $B$  must outweigh this very large difference in monetary expectation (to say nothing of the elation you would experience in  $B$  if you end up with \$5 000 000). While emotions plausibly matter, the emotions we would have to invoke here are costly indeed.

to show that there is a serious and unresolved explanatory challenge for such a theorist. At this point, we think it reasonable to rationalize Allais in the kinds of well-axiomatized, Independence-violating frameworks we have been working with.<sup>59</sup>

We think the correct lesson to draw from Diamond-cases is that everyone must finesse outcomes to some degree. And while we have not presented a full theory of outcome individuation here, we think it unlikely that the correct theory will reconcile EU with Allais. We maintain that risk-sensitivity is about how you structure the possible realization of your ends, while the re-individuation strategy makes risk-sensitivity all about your attitudes towards ends themselves.<sup>60</sup>

## 7 | CONCLUSION

We started by outlining three desiderata for the correct decision theory: accommodate risk-sensitivity, respect Betweenness, and allow for stakes-sensitive risk-attitudes. We have argued that Betweenness is indeed a compelling normative constraint and that Betweenness-satisfying theories, our version of WLU in particular, meet each of our desiderata in a natural way.

This opens up a range of avenues for further work. Some questions involve representation (e.g., providing explicit Jeffrey- or Savage-style representations for WLU). Some questions involve the connection between risk and distributive ethics (e.g., the implications of Betweenness behind the veil of ignorance). Others involve applied policy-making (e.g., the implications of stakes-sensitivity for handling catastrophic risk). We leave these weighty questions for another time.<sup>61</sup>

## ORCID

Christopher Bottomley  <https://orcid.org/0000-0002-0812-3705>

Timothy Luke Williamson  <https://orcid.org/0000-0002-9920-602X>

<sup>59</sup> Again, however, to re-iterate the point we made in Footnote 4: EU coupled with some story of outcome-individuation compatible with Allais would satisfy Betweenness while accommodating reasonable risk-sensitive preferences. Such a theory may respect the three conditions we started with, and so we need not necessarily reject it given our stated goals here.

<sup>60</sup> Stefánsson and Bradley (2019) would likely push back here and argue that risk-aversion really is a kind of desire. Again, while we defer a full-length treatment of Stefánsson and Bradley's theory for future work, another comment is warranted. They claim that risk-attitudes are entirely explained by agents' attaching desirability to chance propositions. Now, while it is possible that we sometimes care about chances (say in PARENT TRAP), we do not think such desires explain all reasonable risk-sensitivity. Stefánsson and Bradley's key claim (pp. 89-91) is that because risk-averse agents act to minimize risk, they thereby count as desiring to minimize risk, and such a desire should be reflected in their utility function. But note that acting to minimize risk only shows that risk-averse agents *instrumentally* desire to minimize risk. It is far less clear that such agents exhibit an *intrinsic* desire to minimize risk. And instrumental desire is measured by the various functions  $EU(\cdot)$ ,  $REU(\cdot)$ ,  $WLU(\cdot)$ , etc., rather than  $u$ . So, we are unconvinced that for risk-averse agents the measure of intrinsic desirability,  $u$ , must be sensitive to chances.

<sup>61</sup> Sincere thanks go to a number of friends and colleagues for helpful comments and discussions. Special thanks go to Arif Ahmed, Eleanor Boxall, Geoffrey Brennan, Colin Burns, Ray Briggs, Simon Grant, Hilary Greaves, Alan Hájek, Brian Hedden, Chad Lee-Stronach, Christoph Lernpass, Kirsten Mann, Andreas Mogensen, Philip Pettit, Theron Pummer, Pamela Robinson, Toby Solomon, Katie Steele, Jeremy Strasser, Teruji Thomas, Joshua Thong, Philip Trammell, Hayden Wilkinson, and James Willoughby. Thanks also to various anonymous referees who provided incisive and detailed feedback on previous versions of this paper.

## REFERENCES

- Ahmed, A. (2016). Review of risk and rationality. *British Journal for the Philosophy of Science*.
- Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école américaine. *Econometrica* 21, 503–546. <https://doi.org/10.2307/1907921>
- Armendt, B. (2014). On risk and rationality. *Erkenntnis*, 79, 1119–1127. <https://doi.org/10.1007/s10670-013-9543-3>
- Bradley, R. (2017). *Decision Theory with a Human Face*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9780511760105>
- Briggs, R. (2015). Costs of abandoning the sure-thing principle. *Canadian Journal of Philosophy*, 45, 827–840. <https://doi.org/10.1080/00455091.2015.1122387>
- Broome, J. (1991). *Weighing Goods*. Oxford: Oxford University Press. <https://doi.org/10.1002/9781119451266>
- Broome, J. (2009). Can a humean be moderate? In R. Frey, & C. Morris (Eds.) *Value, Welfare, and Morality*, (pp. 279–286). Cambridge: Cambridge University Press.
- Buchak, L. (2013). *Risk and Rationality*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199672165.001.0001>
- Buchak, L. (2017). Taking risks behind the veil of ignorance. *Ethics*, 127, 610–644. <https://doi.org/10.1086/690070>
- Buchak, L. (2019). Weighing the risks of climate change. *The Monist*, 102, 66–83. <https://doi.org/10.1093/monist/ony022>
- Burghart, D. (2020). The two faces of independence: Betweenness and homotheticity. *Theory and Decision*, 88, 567–593. <https://doi.org/10.1007/s11238-019-09735-2>
- Camerer, C. (1989). An experimental test of several generalized utility theories. *Journal of Risk and Uncertainty*, 2, 61–104. <https://doi.org/10.1007/BF00055711>
- Camerer, C., & Ho, T.-H. (1994). Violations of the betweenness axiom and nonlinearity in probability. *Journal of Risk and Uncertainty*, 8, 167–196. <https://doi.org/10.1007/BF01065371>
- 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, 1065–1092. <https://doi.org/10.2307/1912052>
- Chew, S. H. (1989). Axiomatic utility theories with the betweenness property. *Annals of Operations Research*, 19, 273–298. <https://doi.org/10.1007/BF02283525>
- Dekel, E., & Lipman, B. (2010). How (not) to do decision theory. *Annual Review of Economics*, 2, 257–282. <https://doi.org/10.1146/annurev.economics.102308.124328>
- Diamond, P. (1967). Cardinal welfare, individualistic ethics, and interpersonal comparison of utility: A comment. *Journal of Political Economy*, 75, 765–766. <https://doi.org/10.1086/259353>
- Diecidue, E., Schmidt, U., & Zank, H. (2009). Parametric weighting functions. *Journal of Economic Theory*, 144, 1102–1118. <https://doi.org/10.1016/j.jet.2008.10.004>
- Diecidue, E., & Wakker, P. (2001). On the intuition behind rank-dependent utility. *Journal of Risk and Uncertainty*, 23, 281–298. <https://doi.org/10.1023/A:1011877808366>
- Fishburn, P. (1983). Transitive measurable utility. *Journal of Economic Theory*, 31, 293–317. [https://doi.org/10.1016/0022-0531\(83\)90079-0](https://doi.org/10.1016/0022-0531(83)90079-0)
- Fishburn, P. (1988). *Nonlinear Preference and Utility Theory*. Baltimore: The Johns Hopkins University Press. <https://doi.org/10.1137/1032026>
- Grant, S., Kajii, A., & Polak, B. (2000). Decomposable choice under uncertainty. *Journal of Economic Theory*, 92, 169–197. <https://doi.org/10.1006/jeth.2000.2644>
- Grant, S., Özsoy, H., & Polak, B. (2008). Probabilistic sophistication and stochastic monotonicity in the savage framework. *Mathematical Social Sciences*, 55, 371–380. <https://doi.org/10.1016/j.mathsocsci.2007.10.002>
- Gul, F. (1991). A theory of disappointment aversion. *Econometrica*, 59, 667–686. <https://doi.org/10.2307/2938223>
- Gul, F., & Lantto, O. (1990). Betweenness satisfying preferences and dynamic choice. *Journal of Economic Theory*, 52, 162–177. [https://doi.org/10.1016/0022-0531\(90\)90072-R](https://doi.org/10.1016/0022-0531(90)90072-R)
- Hájek, A. (2003). Waging war on pascal's wager. *Philosophical Review*, 112, 27–56. <https://doi.org/10.1215/00318108-112-1-27>
- Hájek, A. (2021). Risky business. *Philosophical Perspectives*, 35, 189–205. <https://doi.org/10.1111/phpe.12147>
- Hammond, P. (1988). Consequentialist foundations of expected utility. *Theory and Decision*, 25, 25–78. <https://doi.org/10.1007/BF00129168>

- Hare, C. (2010). Take the sugar. *Analysis*, 70, 237–247. <https://doi.org/10.1093/analysis/anp174>
- Hill, B. (2013). Confidence and decision. *Games and Economic Behavior*, 82, 675–692. <https://doi.org/10.1016/j.geb.2013.09.009>
- Icard, T. (2021). Why be random? *Mind*, 130, 111–139. <https://doi.org/10.1093/mind/fzz065>
- Jeffrey, R. (1982). The sure thing principle. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 2, 719–730.
- Jeffrey, R. (1983). *The Logic of Decision*, 2nd Edition. Chicago: University of Chicago Press. <https://doi.org/10.2307/2183328>
- Joyce, J. (1999). *The Foundations of Causal Decision Theory*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511498497>
- Marschak, J. (1950). Rational behaviour, uncertain prospects, and measurable utility. *Econometrica*, 18, 111–141. <https://doi.org/10.2307/1907264>
- Pettigrew, R. (2015). Risk, rationality and expected utility theory. *Canadian Journal of Philosophy*, 45, 798–826. <https://doi.org/10.1080/00455091.2015.1119610>
- Pettit, P. (1991). Decision Theory and Folk Psychology, In M. Bacharach & S. Hurley (Eds.), *Essays in the foundations of decision theory* (pp. 147–175). Oxford: Blackwell Publishers.
- Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behaviour and Organisation*, 3, 323–343. [https://doi.org/10.1016/0167-2681\(82\)90008-7](https://doi.org/10.1016/0167-2681(82)90008-7)
- Rinard, S. (2015). A decision theory for imprecise probabilities. *Philosophers' Imprint*, 15, 1–16. <http://hdl.handle.net/2027/spo.3521354.0015.007>
- Savage, L. J. (1954). *The Foundations of Statistics*. New York: John Wiley & Sons.
- Schoenfield, M. (2014). Decision making in the face of parity. *Philosophical Perspectives*, 21, 263–277. <https://doi.org/10.1111/phpe.12044>
- Skyrms, B. (1990). *The Dynamics of Rational Deliberation*. Cambridge, Massachusetts: Harvard University Press.
- Steele, K. (2020). What are the minimal requirements of rational choice: Arguments from the sequential-decision setting? *Theory and Decision*, 68, 1–11. <https://doi.org/10.1007/s11238-009-9145-3>
- Stefánsson, H. O., & Bradley, R. (2019). What is risk aversion? *The British Journal for the Philosophy of Science*, 70, 77–102. <https://doi.org/10.1093/bjps/axx035>
- Stiglitz, J., & Rothschild, M. (1970). Increasing risk: I. a definition. *Journal of Economic Theory*, 2, 225–243. [https://doi.org/10.1016/0022-0531\(70\)90038-4](https://doi.org/10.1016/0022-0531(70)90038-4)
- Thoma, J. (2019). Risk aversion and the long run. *Ethics*, 129, 230–253. <https://doi.org/10.1086/699256>
- von Neumann, J., & Morgenstern, O. (1953). *The theory of games and economic behavior* (3rd ed.). Princeton: Princeton University Press.
- Weirich, P. (2020). Risk as a consequence. *Topoi*, 39, 293–303. <https://doi.org/10.1007/s11245-018-9570-4>
- Wilkinson, H. (2022). In defence of fanaticism. *Ethics*, 132, 445–477. <https://doi.org/10.1086/716869>
- Williamson, T. L. (2023). A risky challenge for intransitive preferences. *Noûs*. <https://doi.org/10.1111/nous.12455>

**How to cite this article:** Bottomley, C., & Williamson, T. L. (2024). Rational risk-aversion: Good things come to those who weight. *Philosophy and Phenomenological Research*, 108, 697–725. <https://doi.org/10.1111/phpr.13006>

## APPENDIX A

We here establish sufficient conditions on  $u$  and  $w$  for the WLU-maximizer to respect FOSD and SOSD. The intuition is that  $uw$  in WLU plays a similar role to  $u$  in EU, so  $uw$  being increasing is tied to FOSD, while concavity is tied to SOSD.<sup>62</sup>

<sup>62</sup> Chew (1983) also provides sufficient conditions for FOSD and SOSD, though he assumes smooth preferences, whereas we take lotteries to be finite but do not assume smoothness. Note also that Chew's proofs rely on fairly sophisticated

*Claim 1.* If  $u$  and  $uw$  are increasing and  $w$  is decreasing, then WLU satisfies FOSD.

*Proof.* For  $A, B \in \mathcal{L}$ , if  $A$  First-Order Stochastically Dominates  $B$ , we can obtain  $B$  from  $A$  by a sequence of probability-shifts from each outcome to the next-worst. Without loss of generality, let  $A_\delta$  be a  $\delta$ -probability shift from the best to the second-best outcome. It suffices to show that  $WLU(A) \geq WLU(A_\delta)$ :

$$WLU(A) = \frac{p_A(x_1)w(x_1)u(x_1) + p_A(x_2)w(x_2)u(x_2) + \cdots + p_A(x_n)w(x_n)u(x_n)}{p_A(x_1)w(x_1) + p_A(x_2)w(x_2) + \cdots + p_A(x_n)w(x_n)}$$

$$WLU(A_\delta) = \frac{(p_A(x_1) - \delta)w(x_1)u(x_1) + (p_A(x_2) + \delta)w(x_2)u(x_2) + \cdots + p_A(x_n)w(x_n)u(x_n)}{(p_A(x_1) - \delta)w(x_1) + (p_A(x_2) + \delta)w(x_2) + \cdots + p_A(x_n)w(x_n)}$$

Assume  $u$  is increasing and  $w$  is decreasing. Call the denominators  $D_A$  and  $D_{A_\delta}$  respectively. Then  $D_A - D_{A_\delta} = -\delta(w(x_2) - w(x_1)) = -\delta K$ , where  $K = (w(x_2) - w(x_1)) \geq 0$ , since  $w$  is decreasing. Then:

$$WLU(A) - WLU(A_\delta) = \frac{(D_A + \delta K) \sum_i p_A(x_i)w(x_i)u(x_i) - D_A \sum_i p_{A_\delta}(x_i)w(x_i)u(x_i)}{D_A(D_A + \delta K)}$$

Since  $\delta K, D_A > 0$ , it suffices that the numerator is positive, which holds if  $u(x_1)w(x_1) \geq u(x_2)w(x_2)$ ; i.e if  $uw$  is non-decreasing.  $\square$

*Claim 2.* If  $u$  is increasing,  $w$  is decreasing, and  $f(x, y) = \frac{u(x)-u(y)}{w(x)-w(y)}$  is decreasing in both arguments, WLU satisfies SOSD.

*Proof.* For  $A, B \in \mathcal{L}$ ,  $A$  second-order stochastically dominates  $B$  if  $B$  results from a sequence of transformations which shift pairs of probabilities either side of the mean farther away, leaving the mean unchanged. Consider such a shift:

$$A = p_A(x_1)x_1 + \cdots + p_A(x_i)x_i + \cdots + p_A(x_j)x_j + \cdots + p_A(x_n)x_n$$

$$A_{\delta, \epsilon} = A + (\delta x_i - \delta x_{i+1} - \epsilon x_j + \epsilon x_{j+1})$$

Since  $A_{\delta, \epsilon}$  is mean-preserving:

$$\delta u(x_i) - \delta u(x_{i+1}) - \epsilon u(x_j) + \epsilon u(x_{j+1}) = 0 \quad (4)$$

It suffices to show  $WLU(A) - WLU(A_{\delta, \epsilon}) > 0$ . That is:

$$\sum_k \frac{w(x_k)}{\sum_m p_A(x_m)w(x_m)} u(x_k) p_A(x_k) - \sum_k \frac{w(x_k)}{\sum_m p_{A_{\delta, \epsilon}}(x_m)w(x_m)} u(x_k) p_{A_{\delta, \epsilon}}(x_k) > 0$$

$$\sum_k \frac{w(x_k) \left( (\sum_m p_{A_{\delta, \epsilon}}(x_m)w(x_m)) u(x_k) p_A(x_k) - (\sum_m p_A(x_m)w(x_m)) u(x_k) p_{A_{\delta, \epsilon}}(x_k) \right)}{R} > 0$$

machinery, whereas the proofs here only presuppose basic algebra, so hopefully give the general reader an idea of the structure of WLU.

Where  $R$  is simply obtained from cross-multiplying the denominators in the previous line. Since  $R$ ,  $w(x_k)$  are positive, it suffices that:

$$(\sum_m p_{A_{\delta,\epsilon}}(x_m)w(x_m))u(x_k)p_A(x_k) - (\sum_m p_A(x_m)w(x_m))u(x_k)p_{A_{\delta,\epsilon}}(x_k) > 0$$

Expanding this gives:

$$\begin{aligned} & \sum_k (\sum_m p_A(x_m)w(x_m))p_A(x_k)u(x_k) + p_A(x_k)u(x_k)(\delta w(x_i) - \delta w(x_{i+1}) - \epsilon w(x_j) + \epsilon w(x_{j+1})) \\ & - \sum_k (\sum_m p_{A_{\delta,\epsilon}}(x_m)w(x_m))p_A(x_k)u(x_k) + \sum_m p_A(x_m)w(x_m) \\ & \times (\delta u(x_i) - \delta u(x_{i+1}) - \epsilon u(x_j) + \epsilon u(x_{j+1})) > 0 \end{aligned}$$

By (4) and that  $p_A(x_k)u(x_k)$  is always positive, it suffices that:

$$\delta w(x_i) - \delta w(x_{i+1}) - \epsilon w(x_j) + \epsilon w(x_{j+1}) > 0$$

Again by (4) and that  $u$  is increasing,  $w$  is decreasing, this amounts to:

$$\frac{u(x_{i+1}) - u(x_i)}{w(x_{i+1}) - w(x_i)} > \frac{u(x_{j+1}) - u(x_j)}{w(x_{j+1}) - w(x_j)}$$

And this holds whenever  $f(x, y) = \frac{u(x)-u(y)}{w(x)-w(y)}$  is decreasing in both arguments. □