Difference Minimizing Theory

Christopher J. G. Meacham

Forthcoming in Ergo

Abstract

Standard decision theory has trouble handling cases involving acts without finite expected values. This paper has two aims. First, building on earlier work by Colyvan (2008), Easwaran (2014b), and Lauwers & Vallentyne (2016), it develops a proposal for dealing with such cases, *Difference Minimizing Theory*. Difference Minimizing Theory provides satisfactory verdicts in a broader range of cases than its predecessors. And it vindicates two highly plausible principles of standard decision theory, Stochastic Equivalence and Stochastic Dominance. The second aim is to assess some recent arguments against Stochastic Equivalence and Stochastic Dominance. If successful, these arguments refute Difference Minimizing Theory. This paper contends that these arguments are not successful.

1 Introduction

One of the challenges facing standard decision theory is how to handle cases with acts without finite expected values, such as the Pasadena game, the St. Petersburg game, and the like.[1] A number of proposals have been offered for extending standard decision theory to handle some of these cases, but no proposal has been able to provide a satisfactory account of all of the cases under discussion.[2] Furthermore, many of these attempts to extend decision theory have troubling consequences, with several prominent discussions suggesting that we should reject a pair of decision-theoretic principles that lie at the heart of our conception of prudential rationality.[3]

The first of these principles is Stochastic Equivalence:

---

[1] See Nover & Hájek (2004) for a classic discussion of these issues.

[2] For some of these proposals, see Colyvan (2008), Easwaran (2008), Easwaran (2014b), Smith (2014), Bartha (2016), Colyvan & Hájek (2016), and Lauwers & Vallentyne (2016).

[3] Decision theory is sometimes taken to be an account of *prudential rationality* – an account of what acts are in the best interests of the subject, where this is understood in terms of something like desire satisfaction or the subject's well-being. Decision is also sometimes taken to be an account of *instrumental rationality* – an account of means-ends rationality, that is, an account of what acts are best at achieving some arbitrary goal. For simplicity I adopt the prudential reading in the text, but everything I say is compatible with both readings.

Stochastic Equivalence: If two acts a and a' have the same probabilities of yielding the same utilities, then either both are rationally permissible, or both are rationally impermissible.[4]

This principle seems like a Moorean fact. Two acts that have the same probabilities of yielding the same utilities might differ in a number of ways – one might pay the winner in dollars while the other pays the winner in pounds, one might be a gamble on coin tosses while the other is a gamble on die tosses, and so on. But it's hard to see how any of these differences could be relevant to prudential rationality. Nevertheless, several people have argued that we should reject Stochastic Equivalence.[5]

The second of these principles is Stochastic Dominance. Say that act a *stochastically dominates* act a' *iff* the probability of a yielding at least x utility is always at least as great as, and sometimes strictly greater than, the probability of a' yielding at least x utility.

Stochastic Dominance: If two available acts a and a' are such that a stochastically dominates a', then a' is rationally impermissible.

Again, this principle seems like a Moorean fact. Suppose, for example, that a and a' have the same probabilities of yielding the same utilities, with one exception – a assigns a probability of p to an outcome with utility u, while a' assigns a probability of p to a different outcome with a lower utility of u^-. Then it seems a must be rationally preferable to a', since a is at least as good as or strictly better than a' with respect to everything that seems relevant to prudential rationality. But again, several authors have suggested that we should abandon Stochastic Dominance.[6]

In this paper I'll develop a proposal, *Difference Minimizing Theory*, for extending decision theory to handle cases involving acts without finite expected values. Difference Minimizing Theory is a natural extension of earlier work, building on proposals by Colyvan (2008), Easwaran (2014b), and Lauwers & Vallentyne (2016). I'll argue that Difference Minimizing Theory provides satisfactory verdicts in a broader range of cases than its predecessors. And I'll show that Difference Minimizing Theory allows us to retain both Stochastic Equivalence and Stochastic Dominance.

The paper will proceed as follows. In section 2, I'll sketch some background. In section 3, I'll present two kinds of cases that cause trouble for standard decision theory. In section 4, I'll describe a natural response to these problems suggested by Colyvan (2008), and present Colyvan's (2008) Relative Expectation Theory. In section 5, I'll present three problems for Relative Expectation Theory. In section 6, I'll tackle the first two problems by employing a difference minimizing technique. In section 7, I'll draw on the work of Easwaran (2014b) and Lauwers & Vallentyne (2016) to tackle the third problem. The resulting account – Difference Minimizing Theory – avoids all three problems facing Relative Expectation Theory. In section 8, I'll show that Difference Minimizing Theory entails Stochastic Equivalence and Stochastic Dominance, and I'll address several arguments

---

[4]Note that Stochastic Equivalence is distinct from the strictly stronger claim – call it *Expected Isomorphism* – that if the expected utility contributions of a and a' are isomorphic (i.e., one can construct a value-preserving bijection between them) then either both are rationally permissible, or both are rationally impermissible. While the theory I defend in this paper vindicates Stochastic Equivalence, it does not vindicate Expected Isomorphism (cf. footnote 42).

[5]For example, see Seidenfeld et al. (2009), Smith (2014), and Lauwers & Vallentyne (2016).

[6]See Seidenfeld et al. (2009), Smith (2014), and Lauwers & Vallentyne (2016).

against these principles. This includes an argument by Seidenfeld et al. (2009), who prove that, given some *prima facie* plausible assumptions, Stochastic Equivalence leads to contradiction. In section 9, I'll discuss whether Difference Minimizing Theory is the final theory of prudential rationality, and summarize my results.

2   Background

Whenever a subject needs to make a decision, they're in a *decision problem*. I'll take a decision problem to be an ordered 4-tuple (A, S, cr, u), where:

- A (the set of *acts*) is a set of mutually exclusive and exhaustive propositions corresponding to the acts available to the subject in this decision problem,

- S (the set of *states*) is a set of mutually exclusive and exhaustive propositions such that each $s \in S$ is compatible with every $a \in A$, where these propositions correspond to the different potential states the world could be in,[7]

- $cr : \mathcal{P} \to [0, 1]$ (the *credence function*) is a probabilistic function from the minimal Boolean σ-algebra $\mathcal{P}$ that contains the elements of A and S as members, to a number in the real interval $[0, 1]$ representing the subject's degree of confidence in that proposition,[8]

- $u : A \times S \to \mathbb{R}$ (the *utility function*) is a function from act and state pairs – intuitively, the outcome that the act would bring about given that state – to real numbers representing the degree to which the subject desires or values that this outcome be actual.[9]

In what follows, I'll restrict my attention to decision problems with finitely many acts and countably many states.

Consider some decision problem (A, S, cr, u). In standard decision theory, the *expected*

---

[7]On this way of characterizing decision problems, every available act is defined over the same set of states. While I take this to be the most natural way of characterizing decision problems, one can find alternate characterizations in the literature (cf. footnote 21).

[8]While credences needn't be probabilistic, it's typically assumed that rational credences must be, and I'll restrict my attention to probabilistic credences in this paper. Note that the atomic elements of this algebra will be conjunctions of form $a \wedge s$, for some $a \in A$ and $s \in S$.

[9]By taking utilities to be represented by real numbers, I am assuming that while utilities can be unbounded, they cannot be infinite. I think we should ultimately reject this assumption, since there are compelling reasons to allow for infinite utilities. And once infinite utilities come into play, a number of interesting issues arise (for some recent discussions, see Hájek (2003), Bartha (2007), Bostrom (2011), and Chen & Rubio (forthcoming)). That said, these issues are orthogonal to the ones I'll be concerned with, so I bracket them here and assume utilities must be finite.

utility of an act $a \in A$ is:[10]

$$EU(a) = \sum_{s \in S} cr(a \wedge s \mid a) \cdot u(a \wedge s), \quad \text{if the sum unconditionally converges}[11] \text{ to some } r \in \mathbb{R},$$

$$= \text{undefined}, \qquad\qquad \text{otherwise.}$$

Thus if the sum of expected utility contributions unconditionally converges to some real number, the expected utility takes that value. But if the sum grows arbitrarily large, arbitrarily small, only conditionally converges, or oscillates forever, then the expected utility is undefined.

In what follows, it will be helpful to graphically represent the expected utility of an act using a diagram that plots the probability of a state (given the act) against the utility of performing that act given that state. Thus, for example, we could represent the expected utility of betting on a fair coin toss at even odds, with payoffs of +/- 2 utility, as in figure 1.
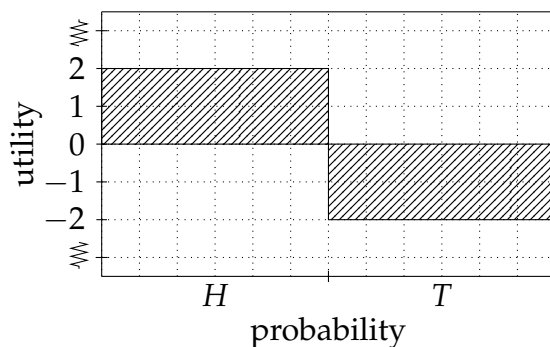

Figure 1: Coin Toss Bet

---

[10]This value is sometimes called the "evidential expected utility" of an act, in contrast to the causal expected utility of an act (see Collins (1996) and Joyce (1999)). This assumption is not entirely innocent; in particular, one of the worries that arises for Colyvan's (2008) relative expectation theory (that it can't handle cases in which acts and states are probabilistically dependent, see section 5) won't arise if we adopt causal decision theory as our starting point. But we'd presumably like a way of handling the cases under discussion which doesn't require us to commit ourselves to one side of the causal/evidential decision theory debate. So I'll take evidential decision theory to be our starting point here. (That said, tweaking this discussion to line up with causal decision theory is straightforward.)

[11]We can say that a (finite or infinite) sum *converges* to a real number $r \in \mathbb{R}$ *iff* for every $\epsilon \in \mathbb{R}$, there exists a $n \in \mathbb{N}$ such that every partial sum of n or more terms is within $\pm\epsilon$ of r. A sum that converges to $r \in \mathbb{R}$ *unconditionally converges* if every sum formed by re-ordering the terms in the original sum also converges to r, and *conditionally converges* otherwise.

Conditional convergence is usually contrasted with *absolute convergence*, where a sum absolutely converges *iff* the sum formed by taking the absolute value of each term converges to some $r \in \mathbb{R}$. If we restrict our attention to the reals, a sum absolutely converges *iff* it unconditionally converges. But once we extend the notion of convergence to accommodate the extended reals, absolute convergence and unconditional convergence can come apart (e.g., the sum $1, -1, 1, ...$ will absolutely converge to an extended real number (∞), but won't unconditionally converge to anything). And it's unconditional convergence, not absolute convergence, that captures what we're interested in.

4

In these diagrams, each term of the expected utility calculation is represented by a box. The boxes above the 0-axis represent positive contributions, the boxes below the 0-axis represent negative contributions, and the area of each box represents the magnitude of that contribution. So the expected utility of an act can be visualized as the area above the 0-axis minus the area below the 0-axis.

One unhappy feature of the standard approach is that it assesses acts whose expected utility is undefined because the sum grows arbitrarily large in the same way as it assesses acts whose expected utility is undefined because the sum grows arbitrarily small. A natural way to address this problem, endorsed by a number of authors, is to allow expected utilities to be infinite.[12] I'll do the same by extending the usual notion of convergence to the *extended real numbers*. To form the extended real numbers $\overline{\mathbb{R}}$, we take the union of the standard real numbers and two new elements, ∞ and $-\infty$. We then extend the usual ordering and arithmetic relations to apply to these new elements as well.[13] Finally, we extend the notion of convergence to an extended real number $r \in \overline{\mathbb{R}}$ in the natural way.[14]

We can then take the *extended expected utility* of an act to be:

$$EEU(a) = \sum_{s \in S} cr(a \wedge s \mid a) \cdot u(a \wedge s), \quad \text{if the sum unconditionally converges to some } r \in \overline{\mathbb{R}},$$

$$= \text{undefined}, \qquad\qquad\qquad \text{otherwise.}$$

Graphically, this definition of extended expected utility has the following consequences. The extended expected utility of an act will unconditionally converge to a finite $r \in \overline{\mathbb{R}}$ *iff* both the area above and below the 0-axis of its graph is finite. The extended expected utility of an act will unconditionally converge to ∞ *iff* the area above the 0-axis is infinite and the area below the 0-axis of its graph is finite, and will unconditionally converge to $-\infty$ *iff* the area below the 0-axis is infinite and the area above the 0-axis of its graph is finite. And the extended expected utility of an act will fail to unconditionally converge to a $r \in \overline{\mathbb{R}}$ *iff* both the areas above and below the 0-axis are infinite.

Given this extended notion of expected utility, we can make prescriptions as follows. Given a decision problem (A, S, cr, u), an act $a \in A$ is permissible *iff* there's no act $a' \in A$ that has a higher extended expected utility.

From now on, when I speak of "expected utility" I'll be referring to extended expected utility, and when I speak of "standard decision theory" I'll be referring to the theory of prescriptions based on extended expected utilities just described.

---

[12] For example, see Nathan (1984), Arntzenius (2014), and Lauwers & Vallentyne (2016).

[13] More precisely, ordering and arithmetic operations are extended over these new elements as follows: *Order:* $\forall r \in \mathbb{R}, -\infty < r < \infty$. *Addition and Subtraction:* 1. $\forall r \in \mathbb{R}, r \pm \infty = \pm\infty$. 2. $\infty + \infty = \infty$. 3. $-\infty + -\infty = -\infty$. 4. $\infty + -\infty = undefined$. *Multiplication:* 1. $\forall r \in \mathbb{R}^+, r \cdot \pm\infty = \pm\infty$. 2. $\forall r \in \mathbb{R}^-, r \cdot \pm\infty = \mp\infty$. 3. $\pm\infty \cdot \pm\infty = \infty$. 4. $\pm\infty \cdot \mp\infty = -\infty$. 5. $0 \cdot \pm\infty = 0$. *Division:* 1. $\forall r \in \mathbb{R}, \frac{r}{\pm\infty} = 0$. 2. $\forall r \in \mathbb{R}^+, \frac{\pm\infty}{r} = \pm\infty$. 3. $\forall r \in \mathbb{R}^-, \frac{\pm\infty}{r} = \mp\infty$. 4. $\frac{\pm\infty}{\pm\infty} = \frac{\mp\infty}{\mp\infty} = undefined$. 5. $\frac{\pm\infty}{0} = undefined$.

[14] That is, let a sum *converge* to $r \in \overline{\mathbb{R}}$ *iff* (i) $r \in \mathbb{R}$, and for every $\epsilon \in \mathbb{R}$, there exists a $n \in \mathbb{N}$ such that every partial sum of n or more terms is within $\pm\epsilon$ of r, or (ii) $r = \infty$, and for any $\alpha \in \mathbb{R}$, there exists a $n \in \mathbb{N}$ such that every partial sum of n or more terms is larger than α, or (iii) $r = -\infty$, and for any $\alpha \in \mathbb{R}$, there exists a $n \in \mathbb{N}$ such that every partial sum of n or more terms is smaller than α. Let a sum that converges to $r \in \overline{\mathbb{R}}$ *unconditionally converge* if every sum formed by re-ordering the terms in the original sum also converges to r, and *conditionally converge* otherwise.

3 Two Problems for Standard Decision Theory

Although standard decision theory yields plausible results in a wide range of cases, there are other cases in which it fails to yield the desired verdicts. Here are two such cases.

Petrograd vs St. Petersburg (Figure 2): A fair coin will be repeatedly tossed until it lands tails. (In this case and the cases that follow, let n be the total number of times the coin is tossed, and s_n be the name of the state in which the coin is tossed that many times.) Before the coin flipping begins, you are presented with two options. The first is to play the Petrograd game, which yields $1 + 2^{n-1}$ utility. The second is to play the St. Petersburg game, which yields 2^{n-1} utility.[15]
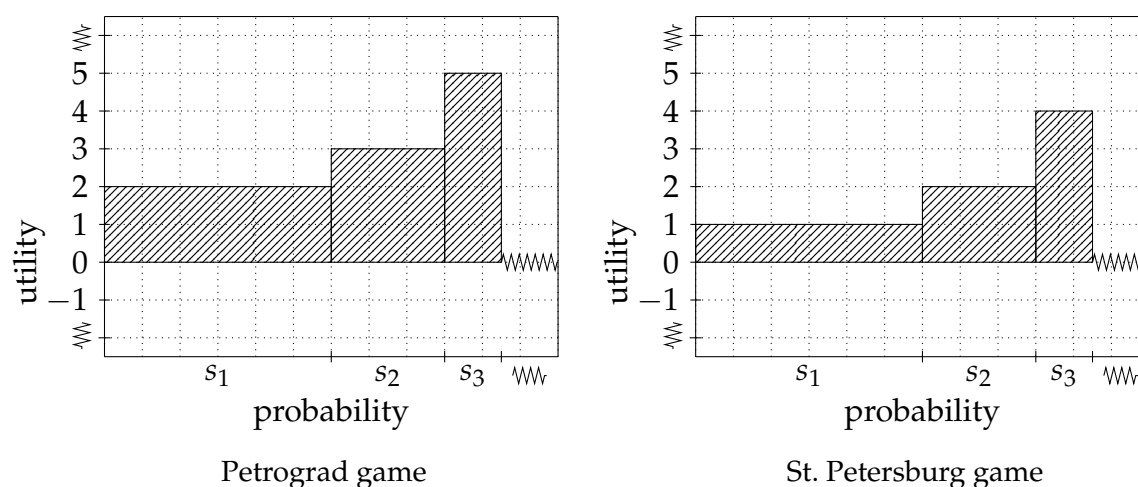

Petrograd game   St. Petersburg game

Figure 2: Petrograd vs St. Petersburg

In both the St. Petersburg game and the Petrograd game the area above the 0-axis is infinite and the area below the 0-axis is finite, so both games have an infinite expected utility. Thus standard decision theory will take both acts to be permissible. But this is implausible: the Petrograd game is strictly better, since regardless of how many coin tosses there are, you'll get an extra unit of utility if you choose the Petrograd game instead of the St. Petersburg game.

Here is another problem case for standard decision theory:

Altadena vs Pasadena (Figure 3): A fair coin will be repeatedly tossed until it lands tails. Before the coin flipping begins, you are presented with two options. The first is to play the Altadena game, which yields $1 + (-1)^{n-1} \cdot \frac{2^n}{n}$ utility. The second is to play the Pasadena game, which yields $(-1)^{n-1} \cdot \frac{2^n}{n}$ utility.[16]

The expected utility contributions of the Pasadena game correspond to the elements of the alternating harmonic series $(1 - \frac{1}{2} + \frac{1}{3} - ...)$, which converges to $\ln 2 \approx 0.69$. But since both the area above the 0-axis and the area below the 0-axis is infinite, we know from

---

[15]The Petrograd game was introduced by Colyvan (2008).
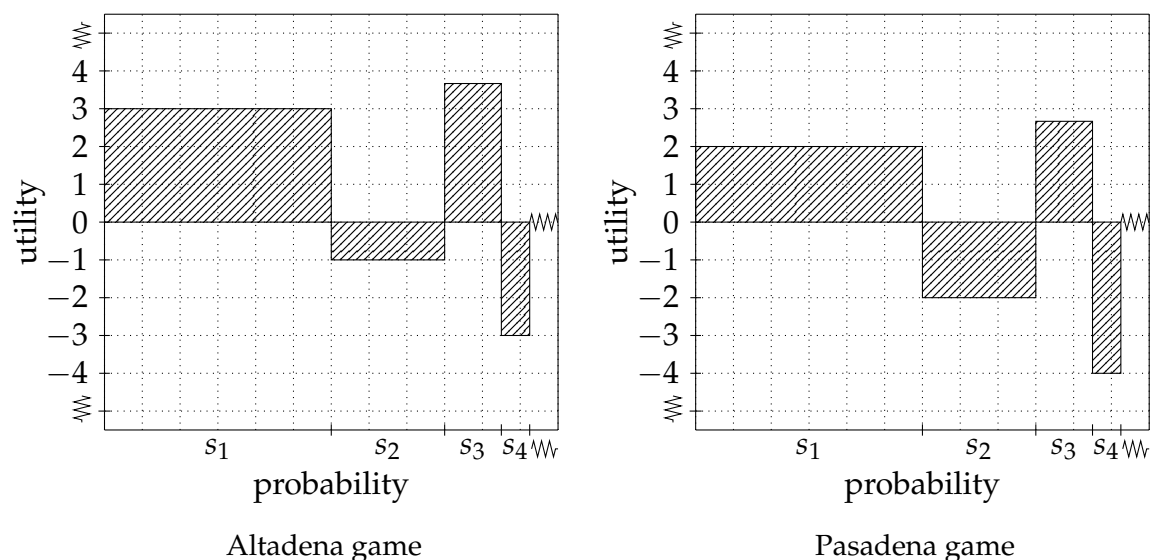[16]The Pasadena and Altadena games were introduced by Nover & Hájek (2004).

Figure 3: Altadena vs Pasadena

section 2 that it doesn't *unconditionally* converge to that value. Thus the expected utility of the Pasadena game is undefined. In a similar vein, the expected utility of the Altadena game converges to $1 + \ln 2$, but it doesn't unconditionally converge to that value. Again, it follows that the expected utility of the Altadena game is undefined. Since neither act has a higher expected utility, standard decision theory will take both acts to be permissible. Again, this is implausible, since regardless of the outcome you'll get an extra unit of utility playing the Altadena game.

4 Difference Taking and Relative Expectation Theory

Both of the cases described in section 3 reveal ways in which standard decision theory fails to respect natural dominance intuitions. In both of these cases, standard decision theory fails to favor one act over another, even though it's guaranteed to yield an extra unit of utility no matter what state obtains. These problems arise because standard decision theory has difficulty registering differences between infinite acts.

A natural way to fix this problem is to stop assessing the expected utilities of acts in isolation, and to instead assess the differences between the expected utilities of pairs of acts. For even if two acts both have infinite or undefined expected utilities, there can still be non-trivial and well-defined differences between these expected utilities.

Colyvan's (2008) *Relative Expectation Theory* is a canonical example of a theory which appeals to difference-taking in order to handle these kinds of cases.[17] In terms of the formalism presented in section 2, we can spell out (a slight extension of) Relative Expectation Theory as follows.[18] Suppose we have a decision problem (A, S, cr, u), and two

---

[17]Versions of Relative Expectation Theory have subsequently been defended by Colyvan & Hájek (2016) and Lauwers & Vallentyne (2016).

[18]It's an extension because it yields extended real relative values (instead of just real values), in order to

7

acts $a, a' \in A$ that are probabilistically independent of the states.[19] Then we can define the *relative expected utility* of a over a' as:

$$REU(a, a') = \sum_{s \in S} cr(s) \cdot \big(u(a \wedge s) - u(a' \wedge s)\big), \quad \text{if the sum unconditionally converges to some } r \in \overline{\mathbb{R}},$$

$$= \text{undefined}, \qquad\qquad\qquad \text{otherwise.}$$

Then Relative Expectation Theory says that in this decision problem, if $a \in A$ is such that for all a' $REU(a, a') \geq 0$, then a is permissible. And if $a \in A$ is such that for some a' $REU(a, a') < 0$, then a is impermissible.

Visually, we can think of Relative Expectation Theory as assessing the difference between the areas of the expected utility contributions of a pair of acts. Thus we can graph the Petrograd vs St. Petersburg case from section 3 as in figure 4, where the first diagram shows the superimposed acts, and the second shows the net difference between their areas.

Petrograd vs St. Petersburg
(superimposed)

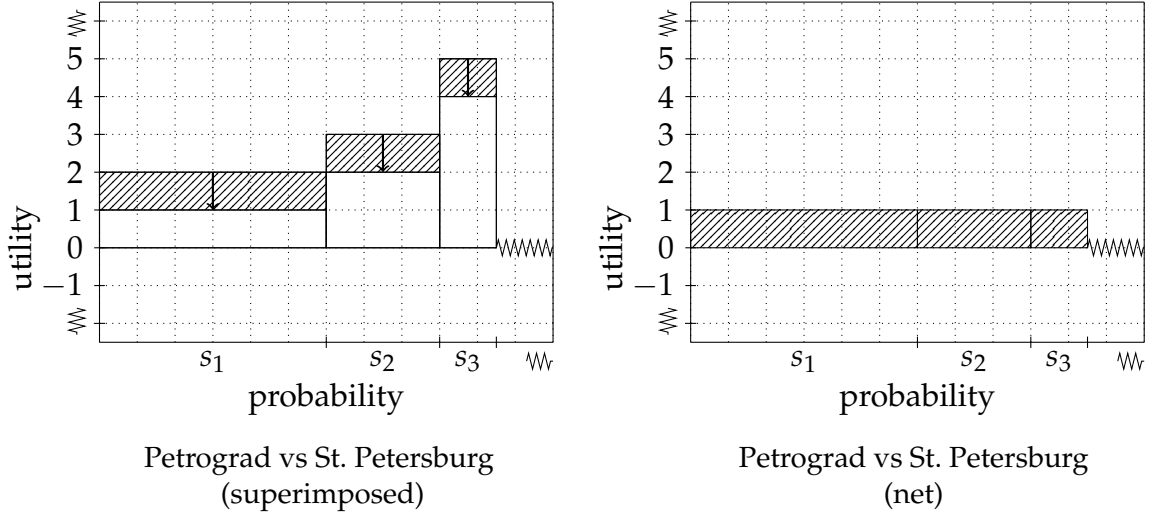Petrograd vs St. Petersburg
(net)

Figure 4: Petrograd vs St. Petersburg (relative)

The net difference graphs are what we want to focus on. These net difference graphs convey the same information about relative expected utilities as the earlier diagrams conveyed about expected utilities. Thus given a net difference graph for a pair of acts, we can deduce that their relative expected utility (i) will unconditionally converge to a finite $r \in \overline{\mathbb{R}}$ *iff* both the area above and below the 0-axis of the net difference graph is finite, (ii) will unconditionally converge to ∞ *iff* the area above the 0-axis is infinite and the area below the 0-axis of its net difference diagram is finite, and (iii) will unconditionally converge to $-\infty$ *iff* the area below the 0-axis is infinite and the area above the 0-axis of its net difference graph is finite.

---

allow the theory to yield well-defined verdicts in cases in which the difference in the expected utilities of acts is infinite.

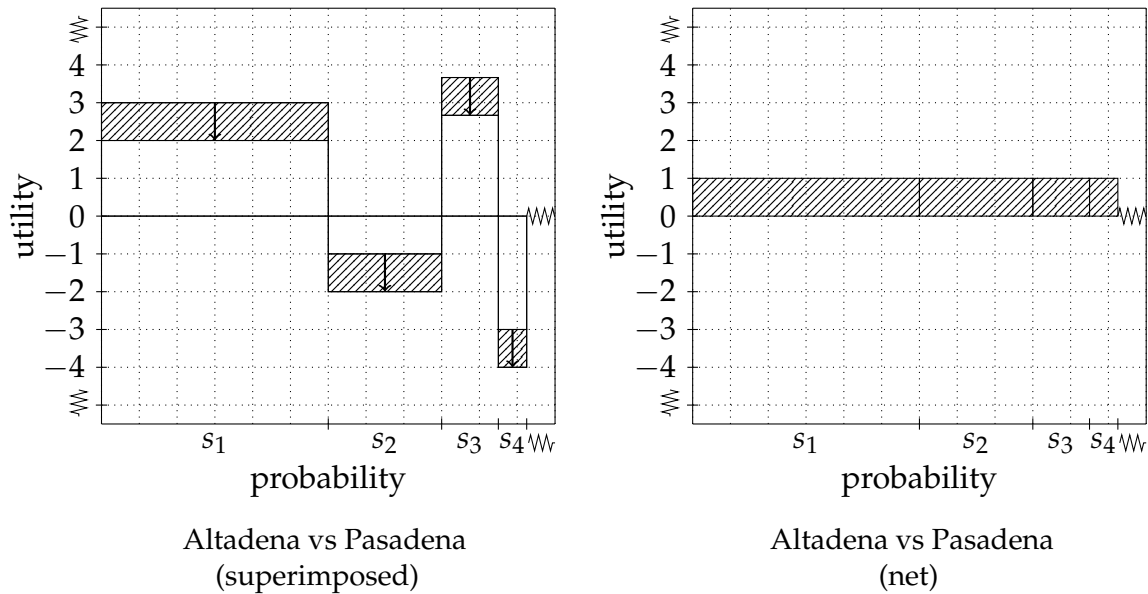[19]I.e., for all $s \in S$, $cr(s \mid a) = cr(s \mid a') = cr(s)$.

Figure 5: Altadena vs Pasadena (relative)

In the Petrograd vs St. Petersburg case, the relative expected utility of the Petrograd game over the St. Petersburg game is 1, as we can see in figure 4. Thus only playing the Petrograd game is permissible. Likewise, in the Altadena vs Pasadena case, the relative expected utility of the Altadena game over the Pasadena game is 1, as we can see in figure 5. So only playing the Altadena game is permissible.

Let's take a step back and consider the motivation for preferring Relative Expectation Theory to standard decision theory. Relative Expectation Theory and standard decision theory both appeal to the same idea – comparing the expected utilities of acts – but they go about it in different ways. Standard decision theory assesses the expected utilities of acts and then looks at the differences between them, while Relative Expectation Theory assesses these differences directly.

Now, on both standard decision theory and Relative Expectation Theory, evaluating a case boils down to assessing the relative areas above and below the 0-axis in the corresponding graphs. And while both of these theories work well when these areas are finite, neither is able to provide discriminating verdicts when comparisons between infinite areas are required.[20] But because Relative Expectation Theory directly assesses differences between acts, it's able to turn many cases that involve comparisons between infinite areas on standard decision theory into cases that only require comparisons between finite areas. Thus Relative Expectation Theory ends up providing discriminating verdicts in a range

---

[20]Thus standard decision theory fails to providing discriminating verdicts in the Petrograd vs St. Petersburg case (which requires comparing the infinite areas about the 0-axis in the Petrograd and St. Petrograd games) and the Altadena vs Pasadena case (which requires comparing infinite areas above and below the 0-axis within each game, as well as comparing those areas between games). And as we'll see in section 5, Relative Expectation Theory fails to provide discriminating verdicts in the Pasadena vs Nothing case (which requires comparing infinite net areas above and below the 0-axis) and the different coins version of the Altadena vs Pasadena case (ditto).

of cases in which standard decision theory does not.

So fans of standard decision theory seem to have a compelling reason to adopt Relative Expectation Theory. For Relative Expectation Theory is faithful to the motivation for standard decision theory – it's just another way of implementing the idea of comparing the expected utilities of acts. But Relative Expectation Theory yields discriminating verdicts in a range of cases in which standard decision theory does not.

Unfortunately, as we'll see in section 5, this rationale for Relative Expectation Theory hits a snag. For while there are cases in which Relative Expectation Theory yields discriminating verdicts while standard decision theory does not, there are also cases in which standard decision theory yields discriminating verdicts while Relative Expectation Theory fails to apply at all. Thus Relative Expectation Theory fails to be strictly better than standard decision theory.

5   Three Worries for Relative Expectation Theory

Colyvan's (2008) Relative Expectation Theory is a step in the right direction. But it falls short in at least three respects.[21]

One worry for Relative Expectation Theory is that relative expected utilities are only well-defined in decision problems in which the acts and states are probabilistically independent.[22] Thus in decision problems in which this requirement doesn't hold (e.g., Newcomb cases), there generally won't be any well-defined relative expected utilities, and Relative Expectation Theory will fall silent. In this respect, Relative Expectation Theory is worse off than standard decision theory, because standard decision theory will still apply and generally yield plausible verdicts in such cases.

A second worry for Relative Expectation Theory is that it fails to yield the right verdicts if the states in question don't line up in the right way. Consider the Altadena vs Pasadena case discussed in section 3. As described, both options rely on the same sequence of coin tosses in order to determine an outcome. This allowed us to pair the state in which you get $(-1)^{n-1} \cdot \frac{2^n}{n}$ utility in the St. Petersburg game with the state in which you get $1 + (-1)^{n-1} \cdot \frac{2^n}{n}$ utility in the Petrograd game, leading to an easy sum over their relative expected utilities. But suppose instead that the two games employed different coins:

Altadena vs Pasadena (different coins) (Figure 6): Two fair coins will be repeatedly tossed until they lands tails. Let n be the total number of times the first coin is tossed, m

---

[21]Colyvan & Hájek (2016) raise a further worry for Relative Expectation Theory that I don't discuss – the worry that this theory will have trouble dealing with acts with disjoint sets of states. As an example, they consider a decision problem in which one has two options: first, a bet that yields \$5 if a coin lands heads, and \$0 otherwise; second, a bet that yields \$6 if a die toss lands on an even number, and \$0 otherwise. In response to this worry, they suggest identifying states with the same probability (and thus identifying (say) the heads state with the even die toss state) for the purposes of making relative expected utility calculations. As I've characterized things in section 2, however, this problem can't arise – it's impossible for acts in the same decision problem to have different sets of states. Thus the states must be richer than they envision. For example, in the case they describe there must be at least four states: a heads and even die toss state, a tails and even die toss state, a heads and odds die toss state, and a tails and odds die toss state.

[22]See Colyvan (2008), Bartha (2016), and Colyvan & Hájek (2016) for discussions of this worry.

the total number of times the second coin is tossed, and s_{nm} the name of the state in which the first coin is tossed n times and the second coin is tossed m times. Before the coin flipping begins, you are presented with two options. The first is to play the Altadena game, which yields $1 + (-1)^{m-1} \cdot \frac{2^m}{m}$ utility. The second is to play the Pasadena game, which yields $(-1)^{n-1} \cdot \frac{2^n}{n}$ utility.
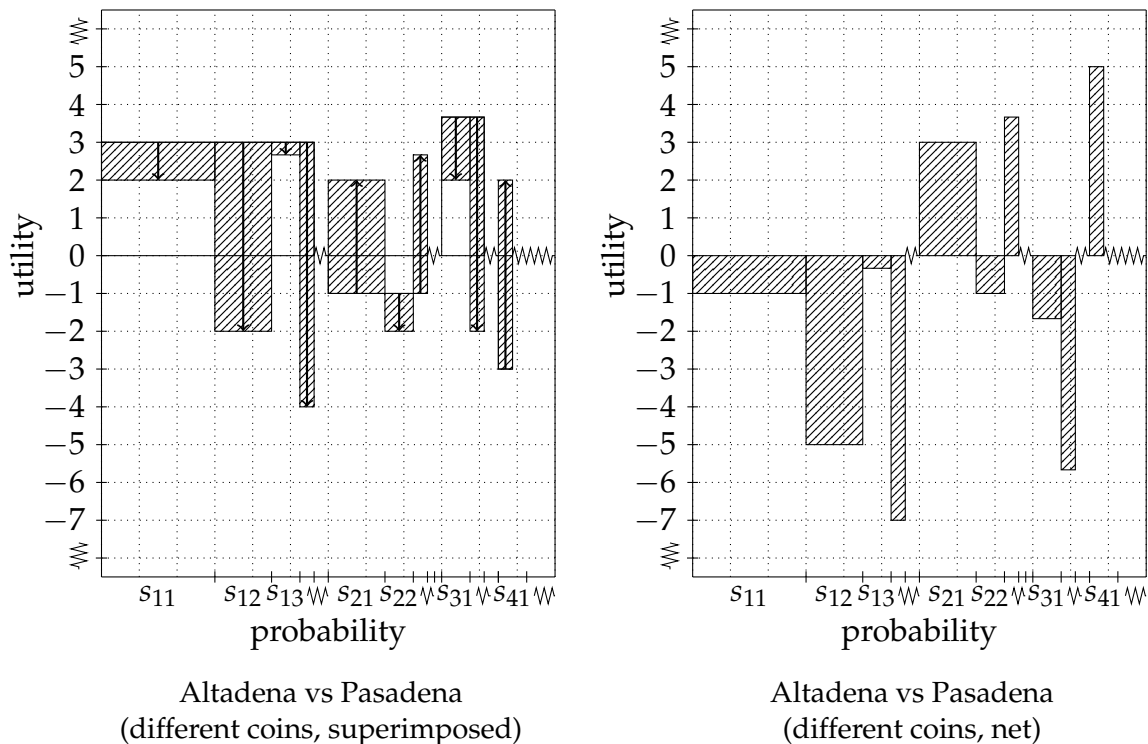

Figure 6: Altadena vs Pasadena (different coins, relative)

In this case both the areas above and below the 0-axis are infinite. So the relative expected utility of the Altadena game over the Pasadena game is undefined, and Relative Expectation Theory will fall silent. This is implausible – if the Altadena game is preferable to the Pasadena game in the same coin case, the Altadena game should be preferable to the Pasadena game in the different coin case as well.[23]

---

[23]One might wonder whether one could avoid this problem by following Colyvan & Hájek (2016) and (i) allowing for different acts to have disjoint states, and (ii) identifying states with the same probabilities (see footnote 21). Thus one might identify the different coin toss outcomes in the Altadena game with the same probability coin toss outcomes in the Pasadena game, and thus get the desired verdict that the relative expected utility of the Altadena game over the Pasadena game is 1.

This response is unsatisfying for two reasons. First, we only get the desired verdicts (that only the Altadena game is permissible) if we individuate states in a way that results in the Altadena game and Pasadena games having disjoint states. But nothing forces us to do this; we can also just form a richer space of states by permuting the possible coin toss outcomes of the different coins. Indeed, the way I've characterized decision problems requires this (see footnote 21). And once we do this, the relative expected utility will go undefined, and Relative Expectation Theory will fall silent.

A third worry one might raise for Relative Expectation Theory is that it falls silent in cases where the differences between the expected utility contributions of two acts fail to unconditionally converge. For example, consider the following case:

Pasadena vs Nothing (Figure 7): A fair coin will be repeatedly tossed until it lands tails. Before the coin flipping begins, you are presented with two options. The first is to play the Pasadena game, which yields $(-1)^{n-1} \cdot \frac{2^n}{n}$ utility. The second is to do Nothing, which yields 0 utility no matter what.

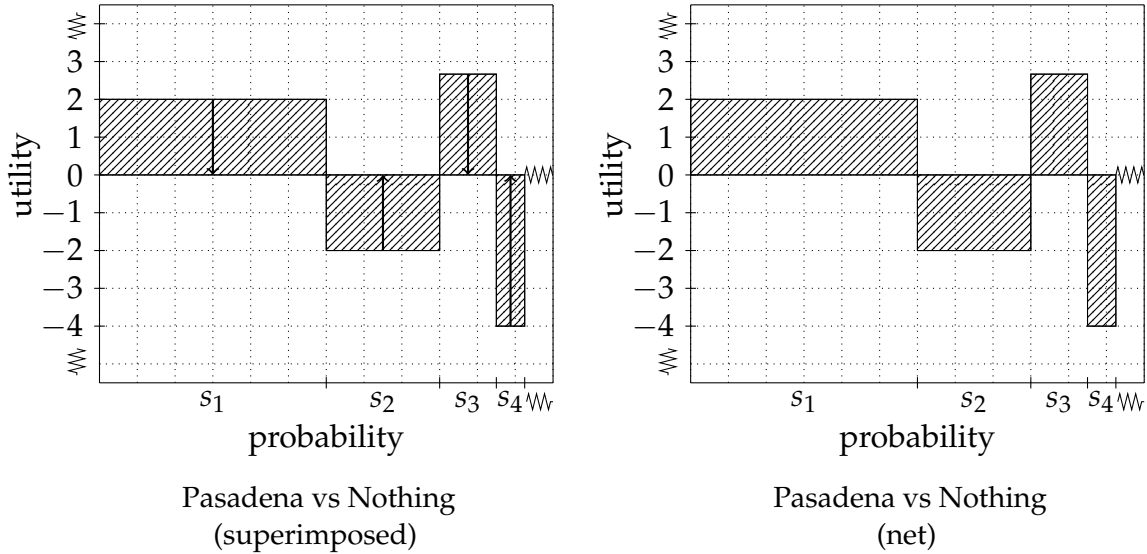Pasadena vs Nothing
(superimposed)

Pasadena vs Nothing
(net)

Figure 7: Pasadena vs Nothing (relative)

This is another case in which both the areas above and below the 0-axis are infinite. So the sum of these differences won't unconditionally converge to anything, and the relative expected utility of playing the Pasadena game over doing Nothing is undefined. Thus Relative Expectation Theory falls silent.

This worry is more mild than the first two because it's not as clear that this is a bad result. After all, it isn't obvious what the right prescriptions in this case should be. That said, if there was a principled and plausible way of providing more fine-grained prescriptions in cases like this, then Relative Expectation Theory's inability to provide fine-grained prescriptions would be a demerit of the view. And, as we'll see in section 7.1, there is a principled and plausible way to provide such prescriptions – namely, adopting the proposal offered by Easwaran (2014b).

Second, this move does little to address the underlying problem. Suppose we tweak the case so that the probabilities of the coin toss outcomes are slightly different for the two different coins. Since we can make these differences arbitrarily small, they should have no appreciable bearing on what one should do. But Colyvan and Hajek's suggestion to identify states with the same probability won't apply, and Relative Expectation Theory will again fall silent.

6 Difference Minimization

The first and second worries raised in section 5 for Relative Expectation Theory both stem from the fact that Relative Expectation Theory pairs expected utility contributions via the states that give rise to them. This leads to the first worry – that the theory won't apply to cases in which acts and states are probabilistically dependent – because the formalism requires the same probability to be assigned to each state given either act. And it leads to the second worry – that the theory falls silent in the Altadena vs Pasadena (different coins) case – because the differences between contributions paired by states is infinite with respect to both positive and negative terms.

To deal with these issues, we need to revise Relative Expectation Theory so that it doesn't pair contributions via states. There are various alternatives one might consider. But since the theory is unable to yield verdicts when the differences between contributions go infinite with respect to both positive and negative terms, we want a way of comparing the outcomes of acts that minimizes the differences between them. Or, putting the point in terms of graphs, we want a way of comparing the outcomes of acts that minimizes the area of their difference.

That is precisely what I propose we should do. I propose that we should order the contributions of each act in a way that minimizes the area of their difference – which we can do by ordering the outcomes of each act from lowest utility to highest utility – and then take the relative value of one act versus another to be equal to this difference in area. (Note that the procedure of ordering outcomes from lowest to highest utility still works in cases without highest or lowest utility outcomes, since the lack of a maximum or minimum doesn't hamper our ability to order outcomes or our ability to get well-defined verdicts.[24])

Let's look at how we might flesh out this idea, with respect to a given decision problem (A, S, cr, u). Given an act $a \in A$, let $<_a$ be a total order defined over the states in S such that if $s, s' \in S$, then $u(a \wedge s) \leq u(a \wedge s')$. Thus $<_a$ orders the states from those that yield the lowest utilities given a to those that yield the highest. There's some arbitrariness here – if there are states that yield the same utility given a, then $<_a$ will arbitrarily rank one above the other – but this arbitrariness won't affect our prescriptions, so we can ignore it.

Consider the graph of a's utility line after ordering the states from lowest to highest utility, with the height of each state determined by the state's utility given a, and the width of each state corresponding to the state's probability given a. We can describe this contour with a function $u_{<_a}(x) : [0, 1] \rightarrow \mathbb{R}$, that takes a value in the $[0, 1]$ interval, and yields the height of the contour at that point.[25] Thus, as shown in figure 8, if a is the St. Petersburg game then $u_{<_a}(0.25)$ will be 1 (since the height of the contour at the 0.25 mark is 1), and $u_{<_a}(0.66)$ will be 2 (since the height of the contour at the 0.66 mark is 2). And if

---

[24]For example, the Altadena vs Pasadena (different coins) case is a case in which neither bet has a highest or lowest utility outcome. But as we'll see below, using this procedure to derive the verdict that one should prefer the Altadena game is straightforward. (I thank an anonymous referee for encouraging me to highlight this point.)

[25]Formally, we can define $u_{<_a}(x)$ as follows. Let $s_\downarrow$ be the disjunction of state s and all of the states ranked below s by $<_a$ (i.e., $s_\downarrow = \bigvee_{(s' \in S: s' <_a s \vee s' = s)} s'$). Then $u_{<_a}(x) : [0, 1] \rightarrow \mathbb{R}$ is the function $u_{<_a}(x) = u(a \wedge s)$ for the unique s such that $cr(s_\downarrow \mid a) > x$ and $\neg \exists s' \neq s((cr(s'_\downarrow \mid a) > x) \wedge (cr(s'_\downarrow \mid a) < cr(s_\downarrow \mid a)))$.

a is the Petrograd game, then $u_{<_a}(0.25)$ will be 2, and $u_{<_a}(0.66)$ will be 3.
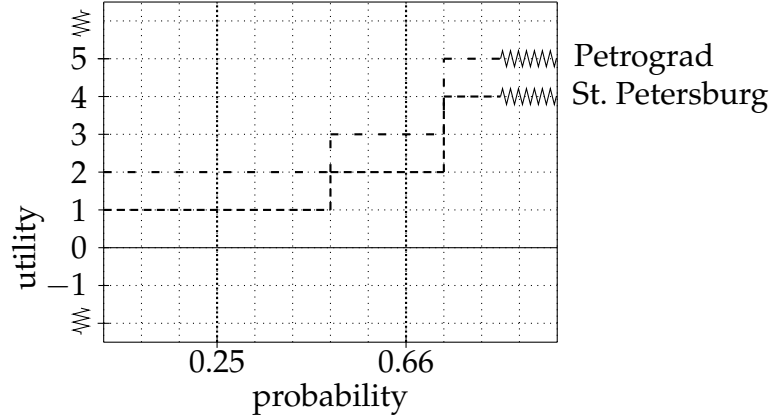

Figure 8: Petrograd vs St. Petersburg (utility profiles)

The easiest way to characterize the difference minimizing version of relative expected utilities is to use integration to determine the area between the two acts:

$$REU_{dm}(a, a') = \int_{x=0}^{1} u_{<_a}(x) - u_{<_{a'}}(x) \, dx.$$

But to neatly mesh this proposal with the earlier proposals, let's do a little more work to formulate things in terms of countable sums.

Let $P_{<_a}$ be the set of points in $[0, 1]$ that mark when each state ends according to the ordering imposed by $<_a$.[26] So if *a* is the St. Petersburg game, $P_{<_a}$ would consist of the set of points $\{\frac{1}{2}, \frac{3}{4}, \frac{7}{8}, ...\}$.

Given a set of real numbers B, let the *atomic intervals* of B ($I(B)$) be the intervals between adjacent members of B.[27] For any interval $i \in I(B)$, let b_i be the point in $[0, 1]$ marking the beginning of that interval, and e_i the point in $[0, 1]$ marking the end.

We can then characterize the difference minimizing version of the relative expected utility of *a* over *a'* ($REU_{dm}(a, a')$) as follows:

$$REU_{dm}(a, a') = \sum_{i \in I\left(P_{<_a} \cup P_{<_{a'}}\right)} |b_i - e_i| \cdot \left(u_{<_a}\left(\frac{b_i + e_i}{2}\right) - u_{<_{a'}}\left(\frac{b_i + e_i}{2}\right) \right),$$

$$\text{if the sum unconditionally converges to some } r \in \overline{\mathbb{R}},$$

$$= \text{undefined, otherwise.}$$

Here's how this works. We first order the possible outcomes of each act from lowest to highest utility. Then we cut up the $[0, 1]$ interval into sections in which the utility of both

---

[26]That is, $P_{<_a} = \bigcup_{(x: \exists s \in S \ cr(s_\downarrow) = x)} x$ (where I'm employing the definition of $s_\downarrow$ from footnote 25).

[27]That is, let the set of *atomic intervals* of B ($I(B)$) be the set of all ordered pairs (x, x') such that $x, x' \in B$, and $\neg \exists x'' \in B$ such that $x < x'' < x'$. Note that if B is dense then $I(B) = \emptyset$, since there won't be any atomic intervals. But in the cases we're concerned with, B won't be dense.

14

acts is constant. For each such section, we multiply the width of that section (its probability) by the difference in utilities between the two acts in that section. This provides the relative expected utility contribution of that section. Then we add up the contributions of each section to get the total (assuming this sum unconditionally converges to some extended real number).

With these difference-minimizing relative expected utilities in hand, we can make prescriptions as follows. Given a decision problem (A, S, cr, u), an act $a \in A$ is permissible *iff* there's no act $a' \in A$ such that $REU_{dm}(a', a) > 0$. Since this is a preliminary version of the theory I'll ultimately defend, I'll call it *Difference Minimizing Theory$^-$*.[28]

Difference Minimizing Theory$^-$ avoids the first and second worries for Relative Expectation Theory raised in section 5. Let's start with the second worry, that Relative Expectation Theory falls silent in the Altadena vs Pasadena (different coins) case. Here is how Difference Minimizing Theory$^-$ will handle this case (see figure 9). First we rank

Altadena vs Pasadena
(different coins, superimposed)
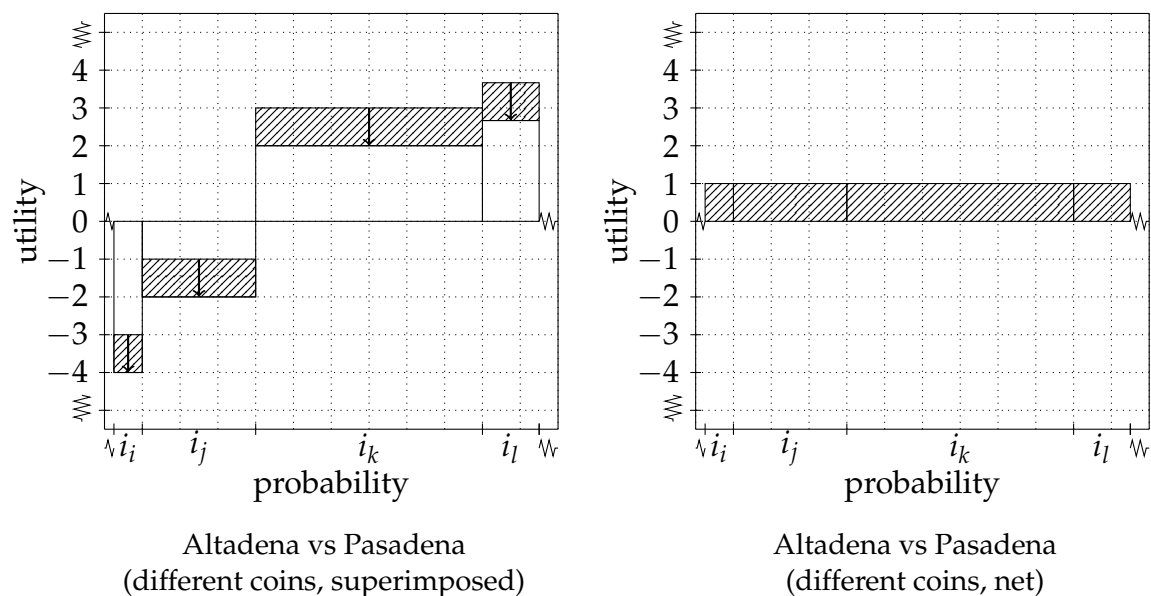
Altadena vs Pasadena
(different coins, net)

Figure 9: Altadena vs Pasadena (difference minimizing)

the states of each act in order of utility. Then we cut up the $[0, 1]$ interval into intervals in which the utility of both acts is constant. The relative expected utility contribution of each interval is its width (probability) times the difference in utilities between the Altadena and Pasadena acts in that interval. The sum of all of these terms will unconditionally converge to 1. So the difference minimizing value of playing the Altadena game over playing the Pasadena game is greater than 0. Thus playing the Altadena game is obligatory, as desired.

---

[28]Easwaran (2014a) presents an approach to constructing versions of decision theory that works by placing various normative constraints on preferences (though it departs from the standard preference-based approach in not taking preferences to be prior to credences or utilities, and in not taking these constraints on preferences to justify claims about the formal features of credences or utilities). Interestingly, one of the versions of decision theory he considers (in section 3.5.4) seems to yield verdicts that are very similar to, and perhaps identical to, those of Difference Minimizing Theory$^-$. (Thanks to Kenny Easwaran here.)

Relative Expectation Theory runs into trouble with this case because it pairs contributions by appealing to the states that gave rise to them. Since different coins are used for each game, that leads to a very different-looking graph than the one given above – a graph which pairs each potential outcome of one coin with infinitely many potential outcomes of the other. And, as we saw in section 5, this leads to differences between the contributions of the two acts that are infinite with respect to both positive and negative terms, and Relative Expectation Theory falls silent in such cases. Difference Minimizing Theory⁻ avoids these headaches because it doesn't try to pair contributions by state. Instead, it simply sets things up so as to minimize the difference in area between the two acts. And this in turn allows us to set things up in a way that yields a finite difference between the contributions of the two acts, which makes assessing the case straightforward.

Let's turn to the first worry for Relative Expectation Theory, that it requires acts and states to be probabilistically independent. Consider a case like the following:

Big Bet vs Small Bet (Figure 10): You are given the option of accepting either a big bet or a small bet on whether the next coin Smith tosses lands heads. Accepting the big bet will lose you 2 utility if the coin lands heads, and win you 2 utility if it lands tails. And if you accept the big bet, Smith will become aware of the bet, and as a friend of yours, will chose to toss a biased coin which has a two thirds chance of landing tails. Accepting the small bet will lose you 1 utility if the coin lands heads, and win you 1 utility if it lands tails. And if you accept the small bet, then Smith will be unaware of the bet, and will chose to toss a fair coin.

In this case the acts and states are probabilistically dependent, so Relative Expectation Theory won't apply. But the right answer in this case is clear: you should accept the big bet. And Difference Minimizing Theory⁻ will straightforwardly yield this result. Since

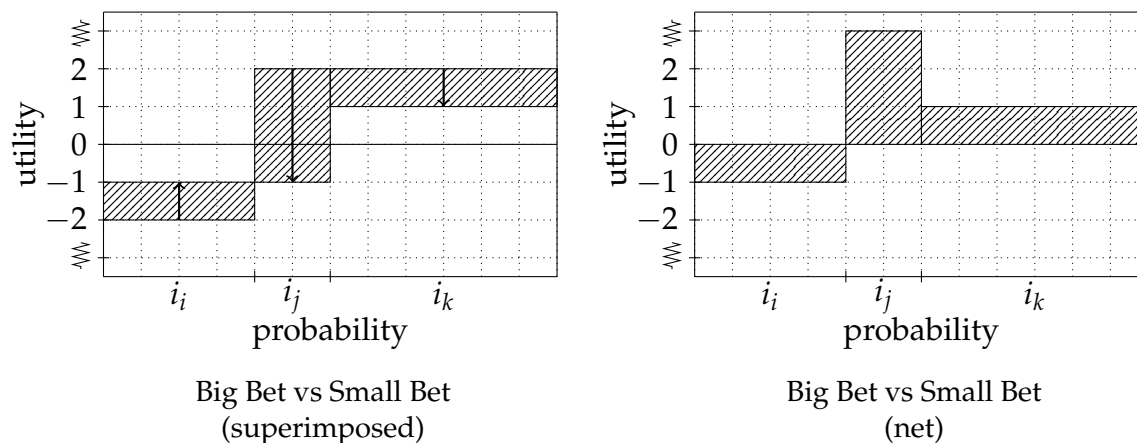Figure 10: Big Bet vs Small Bet (difference minimizing)

the net area above the 0-axis is greater than that below the 0-axis, the difference minimizing value of the big bet over the small bet is positive, and thus taking the big bet is obligatory.

Let's take a step back and consider the motivation for preferring Difference Minimizing Theory⁻ to Relative Expectation Theory. In section 4, we saw a rationale for preferring

Relative Expectation Theory to standard decision theory: both theories appeal to the same idea, but Relative Expectation Theory yields discriminating verdicts in a range of cases in which standard decision theory does not. But this rationale hit a snag – as we saw in section 5, there are also cases to which standard decision theory yields discriminating verdicts while Relative Expectation Theory fails to apply at all.

Difference Minimizing Theory$^-$ appeals to the same idea as standard decision theory and Relative Expectation Theory – evaluating acts by comparing their expected utilities. But Difference Minimizing Theory$^-$ avoids the snag that Relative Expectation Theory encountered, for Difference Minimizing Theory$^-$ applies to all the cases standard decision theory applies to. And even putting those cases aside, Difference Minimizing Theory$^-$ yields discriminating verdicts in a broader range of cases than Relative Expectation Theory. Recall that these theories only fail to provide discriminating verdicts when comparisons between infinite areas are required. Since by construction Difference Minimizing Theory$^-$ minimizes the area to be compared, it minimizes the number of cases that require comparisons between infinite areas.

So fans of standard decision theory and Relative Expectation Theory have a good reason to adopt Difference Minimizing Theory$^-$, for Difference Minimizing Theory$^-$ is faithful to the motivation for standard decision theory and Relative Expectation Theory – it's just another way of implementing the idea of comparing the expected utilities of acts. But Difference Minimizing Theory$^-$ yields discriminating verdicts in a strictly broader range of cases.

7 Alternative Aggregation Techniques

7.1 Stable Principal Values

Let's turn to the third worry for Relative Expectation Theory, that in cases like Pasadena vs Nothing, it falls silent. This in itself is not clearly a bad result, since it's not clear what the verdict in such a case should be. But if there was a principled and plausible way to extend decision theory to yield verdicts in these cases, it would be nice to do so. And Easwaran (2014b) has provided us with just such an extension.[29] Let's first see what Easwaran's extension is, and then consider how we might employ it to bolster Relative Expectation Theory.

As we saw in section 3, standard decision theory doesn't assign a well-defined expected utility to the Pasadena game. In order to evaluate such cases, Easwaran proposes to modify standard decision theory in the following way (slightly modified to incorporate the extension to the extended reals). Let the *n-truncation of the expected utility of a* ($EU^n(a)$) be an expected utility calculation that considers only the contributions of terms whose

---

[29]Easwaran (2008) also proposes a plausible extension. But as Easwaran (2014b) shows, his later proposal is strictly stronger. For example, unlike his earlier proposal, his later proposal yields the desired verdicts in Bartha's (2016) Arroyo case.

utilities have a magnitude of less than or equal to n. That is:[30]

$$EU^n(a) = \sum_{(s \in S : |u(a \wedge s)| \leq n)} cr(a \wedge s \mid a) \cdot u(a \wedge s).$$

Let the *principal value* of an act be the value of these truncated expected utility calculations in the limit as n goes to infinity. Roughly, Easwaran's proposal is to evaluate acts using their principal values instead of their expected utilities.

Visually, we can think of principal values as follows. Imagine a pair of horizontal lines n-units above and below the 0-axis of an act's graph. We can think of EU^n as the sum of the area above the 0-axis minus the area below the 0-axis, taking only contributions wholly inside these horizontal lines into account, as in figure 11. Then we can imagine
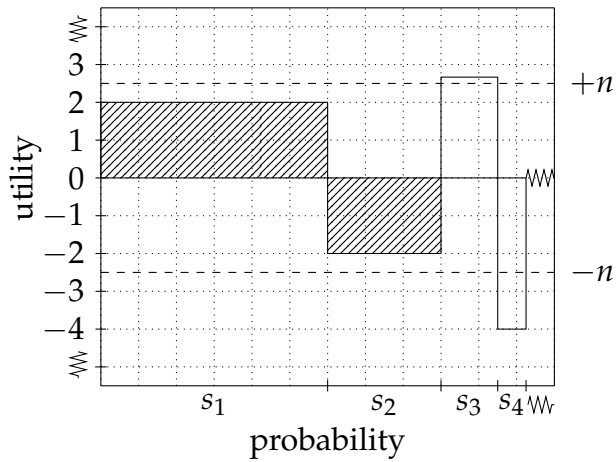

Figure 11: Pasadena vs Nothing ($n = 2.5$ truncation)

redoing this calculation as we symmetrically shift the horizontal lines farther and farther away from the 0-axis. The principal value is what these calculations yield in the limit.

This rough proposal requires two amendments. First, as Easwaran (2014b) notes, principal values are sometimes sensitive to uniform shifts in utility. That is, there are cases in which uniformly changing the utilities assigned to outcomes – such as by choosing to measure utility using a different scale – can change whether an act has a well-defined principal value. Since we want our prescriptions to be invariant to changes in scale, we need to ensure that we don't appeal to principal values in these cases.

Fortunately, Easwaran (2014b) identifies the condition that an act must satisfy in order to be insensitive to such utility transformations. Say that an act a in a decision problem

---

[30]Unlike the characterization of (extended) expected utility given in section 2, this characterization doesn't require extended reals, unconditional convergence, or an undefined clause. That's because truncating the sum to include only those terms with utility contributions at or below n entails that the magnitude of both positive and negative terms is finite, and thus that the sum unconditionally converges to a real number. (We'll reintroduce these clauses when we take the limit of these truncations.)

has *stable tails iff* there exists an $\epsilon > 0$ such that:

$$\lim_{n \to \infty} \left(\sum_{(s \in S : |u(a \wedge s)| > n - \epsilon)} cr(s) \cdot (n - \epsilon) - \sum_{(s' \in S : |u(a \wedge s')| > n + \epsilon)} cr(s') \cdot (n + \epsilon) \right) = 0.$$

Easwaran (2014b) shows that if an act has stable tails, then uniform utility shifts will yield the same uniform shift in principal value, and thus principal value-based prescriptions will be invariant to changes of scale.

Second, we'll want to slightly extend Easwaran's proposal to allow for extended real values. And we don't want to impose a stability condition in such cases, for there are acts (e.g., the St. Petersburg game) which plausibly have infinite value even though they don't have stable tails.

We can perform the desired extension to the extended reals, and incorporate Easwaran's stability condition, as follows. Let's extend the notion of the limit of a function in the same way as we extended the notion of convergence in section 2.[31] (I'll assume this extended notion of limits from now on.) Then we can introduce the notion of the *stable principal value* of an act *a* (SPV(a)) as follows:

$$
\begin{aligned}
SPV(a) &= \lim_{n \to \infty} EU^n(a), && \text{if this value is finite and } a \text{ has stable tails,} \\
&= \lim_{n \to \infty} EU^n(a), && \text{if this value is } \pm \infty, \\
&= \text{undefined}, && \text{otherwise.}
\end{aligned}
$$

And we can evaluate acts using stable principal values instead of expected utilities.

The stable principal value of an act will be the same as its expected utility whenever its expected utility is well-defined. But a number of acts without well-defined expected utilities will have well-defined stable principal values. For example, while the expected utility of the Pasadena game is undefined, the stable principal value of Pasadena game is ln 2. So in the Pasadena vs Nothing case, Easwaran's proposal will yield the verdict that playing the Pasadena game is obligatory.

7.2 Difference Minimizing Values

Now, how might we employ stable principal values to extend Difference Minimizing Theory⁻? Lauwers & Vallentyne (2016) propose to extend Relative Expectation Theory by combining it with Easwaran's (2008) proposal. A natural thought is to do something similar here, but to combine Difference Minimizing Theory⁻ with Easwaran's (2014b) stronger proposal. Let's see how one might do that.

On Easwaran's (2014b) proposal, we characterize the *n*-truncations of expected utilities, see what those values yield in the limit as *n* goes to infinity, and then use those values

---

[31]I.e., suppose we have a sequence of functions $f^n : A \times ... \to \mathbb{R}$. Let $r_{a,...}^n \in \mathbb{R}$ be the value such that $f^n(a, ...) = r_{a,...}^n$. Then $\lim_{n \to \infty} EU^n(a, ...) = r \in \overline{\mathbb{R}}$ *iff* (i) $r \in \mathbb{R}$, and for every $\epsilon \in \mathbb{R}$, there exists a $n' \in \mathbb{N}$ such that for every $n > n'$, $r_{a,...}^n$ is within $\pm \epsilon$ of r, or (ii) $r = \infty$, and for any $\alpha \in \mathbb{R}$, there exists a $n' \in \mathbb{N}$ such that for every $n \geq n'$, $r_{a,...}^n$ is larger than α, or (iiI) $r = -\infty$, and for any $\alpha \in \mathbb{R}$, there exists a $n' \in \mathbb{N}$ such that for every $n \geq n'$, $r_{a,...}^n$ is smaller than α.

to assess acts (assuming those values are stable). Here we'll want to do the same thing, but replace expected utilities with the difference-minimizing version of relative expected utilities. So let the *n-truncation of the difference-minimizing relative expected utility of a over a′* ($REU_{dm}^n(a, a')$) be a $REU_{dm}(a, a')$ calculation that considers only the contributions of terms whose utilities have a magnitude of less than or equal to n. That is:[32]

$$REU_{dm}^n(a, a') = \sum_{\left(i \in I\left(P_{<a} \cup P_{<a'} \right) : \left| \left(u_{<a}\left(\frac{b_i+e_i}{2}\right) - u_{<a'}\left(\frac{b_i+e_i}{2}\right)\right) \right| \leq n \right)} |b_i - e_i| \cdot \left(u_{<a}\left(\frac{b_i+e_i}{2}\right) - u_{<a'}\left(\frac{b_i+e_i}{2}\right)\right).$$

Let's say that the difference between a pair of acts a and a' has *stable tails iff* an act a'' such that $u(a'' \wedge s) = u(a \wedge s) - u(a' \wedge s)$ would have stable tails according to the definition given in section 7.1. We can then use these truncations to define the difference minimizing version of a stable principal value, which I'll call the *difference minimizing value* of a over a' ($DMV(a, a')$), as follows:

$$DMV(a, a') = \lim_{n \to \infty} REU_{dm}^n(a, a'), \quad \text{if this value is finite and the } a/a' \text{ difference has stable tails,}$$
$$= \lim_{n \to \infty} REU_{dm}^n(a, a'), \quad \text{if this value is } \pm\infty,$$
$$= \text{undefined}, \quad \text{otherwise.}$$

Then, given a decision problem (A, S, cr, u), we can say an act $a \in A$ is permissible *iff* there's no act $a' \in A$ such that $DMV(a', a) > 0$.

7.3 Non-Definite Values

The proposal described in section 7.2, like all of the proposals we've considered so far, presupposes that our assessment of the relevant acts yields definite values. As Lauwers & Vallentyne (2016) note, these kinds of proposals are ill-equipped to handle cases without definite values. For example, consider a relative of the Pasadena vs Nothing case – call it the Varying Pasadena vs Nothing case – in which $REU_{dm}^n(\text{Varying Pasadena, Nothing})$ never goes above 2 nor dips below 1, but never converges to a definite value either.

Varying Pasadena vs. Nothing: A fair coin will be repeatedly tossed until it lands tails. Before the coin flipping begins, you are presented with two options. The first option is to play the Varying Pasadena game. The magnitudes of the utilities for this game are the same as the Pasadena game, $2^n \left(\frac{1}{n}\right)$. But the signs of the utilities for this game are different. They start positive, and remain positive for as many terms as possible before $REU_{dm}^n(\text{Varying Pasadena, Nothing})$ would become greater than 2. Then they turn negative, and stay negative for as many terms as possible before $REU_{dm}^n(\text{Varying Pasadena, Nothing})$ would become lower than 1. Then they turn positive again until the terms would become greater than 2, and so on.[33] The second option is to do Nothing, which yields 0 utility no matter what.[34]

---

[32]Where the *n*-truncation of REU_{dm} doesn't require extended reals, unconditional convergence, or an undefined clause, for the same reason that the *n*-truncation of EU doesn't require them (cf. footnote 30).

[33]Thus the utilities of the outcomes of this game are (in order of increasing n) 2, 2, $2\frac{2}{3}$, -4, $-6\frac{2}{5}$, $-10\frac{2}{3}$, $-18\frac{2}{7}$, 32, and so on.

[34]This is a variant of a case discussed by Lauwers & Vallentyne (2016).

As n increases, the REU^n_{dm}(Varying Pasadena, Nothing) will bounce around in the $[1, 2]$ interval, but will never settle on a definite value. Since the difference minimizing value of playing the Varying Pasadena game over Nothing doesn't have a definite value, the proposal from section 7.2 will hold that both playing the Varying Pasadena game and doing Nothing are permissible.[35] But that might seem like the wrong verdict – one might think that playing the Varying Pasadena game should be better than doing Nothing.

Lauwers & Vallentyne (2016) propose to address this worry by assigning interval values to (pairs of) acts. I have no complaints about their proposal. But a simpler approach will suffice for our purposes. Let us say that $DMV(a, a')$ *floats above* 0 *iff* either (i) the a/a' difference has stable tails, and there exists a $n' \in \mathbb{N}$ such that for all $n > n'$, $REU^n_{dm}(a, a') > 0$, or (ii) $DMV(a, a') = r' > 0$.[36] Given this, we can describe my proposal as follows:

Difference Minimizing Theory: Given a decision problem (A, S, cr, u), $a \in A$ is permissible *iff* there's no $a' \in A$ such that $DMV(a', a)$ floats above 0.

This will yield the desired results in the Varying Pasadena vs Nothing case. For while there's no definite difference minimizing value for these acts, the difference minimizing value will float above 0. So we'll get the desired verdict that playing the Varying Pasadena game is obligatory.

Difference Minimizing Theory avoids all three of the worries raised for Relative Expectation Theory in section 5. And it yields satisfactory results in cases without definite values. More broadly, Difference Minimizing Theory provides satisfactory verdicts in a wider range of cases than any of the other views we've considered. In particular, none of the other views we've looked at yield plausible verdicts in all of the cases we've considered so far. But Difference Minimizing Theory does. Thus I take it to be the best proposal on offer for how to handle cases without finite expected utilities.[37,38]

---

[35]That is, playing the Varying Pasadena game (vp) will be permissible since there's no act a' such that $DMV(a', vp) > 0$ (since this value will be 0 if $a' = vp$, and undefined if a' is doing Nothing). Likewise, doing Nothing (no) will be permissible since there's no act a' such that $DMV(a', no) > 0$ (since this value will be 0 if $a' = no$, and undefined if $a' = vp$).

[36]We need the second clause in addition to the first to accommodate cases where the a/a' difference has unstable tails and $DMV(a, a') = \pm\infty$.

[37]The proposal offered by Smith (2014) yields verdicts in an equally broad range of cases. But, like Hájek (2014), I find many of these verdicts implausible. Another notable approach with broad application is the proposal offered by Bartha (2016). But I take Bartha to be engaged in a different kind of project. My goal is to provide an account that will yield concrete prescriptions given decision problems of the kind described in section 2. Bartha's proposal is not an account of this kind, for it won't yield prescriptions without further substantive assumptions about what a subject's preferences are that aren't provided by decision problems of this form. (See Colyvan & Hájek (2016) for a discussion of this feature of Bartha's proposal.)

[38]Earlier, I discussed the motivation for moving to Relative Expectation Theory and Difference Minimizing Theory$^-$. What's the motivation for incorporating Easwaran's (2014b) principal values approach and Lauwers & Vallentyne's (2016) extension to non-definite values? Providing a principled argument for the move to principal values is non-trivial; indeed, the lack of such an argument leads Easwaran (2014b) to declare himself agnostic about it. I think this extension is plausible, but those skeptical of this move can stick with an extension of Difference Minimizing Theory$^-$ that allows for non-definite values. In a similar vein, Lauwers & Vallentyne (2016) simply take their extension to non-definite values to be plausible (and I agree), but those skeptical of this move can stick with the proposal described at the end of section 7.2. Those skeptical of both moves can stick

8 Stochastic Equivalence and Stochastic Dominance

In section 1 we considered two plausible principles:

Stochastic Equivalence: If two available acts a and a' have the same probabilities of yielding the same utilities, then either both are rationally permissible, or both are rationally impermissible.

Stochastic Dominance: If two acts a and a' are such that a stochastically dominates a', then a' is rationally impermissible.

Difference Minimizing Theory entails both of these principles. If two acts a and a' have the same probabilities (conditional on that act) of yielding the same utilities, then $u_{<_a}$ and $u_{<_{a'}}$ will be identical, i.e., these acts will yield the same utility contours. Thus any comparisons involving these acts will be identical, and Difference Minimizing Theory will assign them the same deontic status. Thus Difference Minimizing Theory entails Stochastic Equivalence. Likewise, if a stochastically dominates a', then the area of the utility contour corresponding to a will include all of the area of the utility contour corresponding to a', and more besides. Thus $DMV(a, a') > 0$, and a' will be impermissible. So Difference Minimizing Theory entails Stochastic Dominance.

Given the plausibility of Stochastic Equivalence and Stochastic Dominance, this might seem like yet another reason to favor Difference Minimizing Theory. But Stochastic Equivalence and Stochastic Dominance have recently come under fire, with a number of authors raising worries for these principles.[39] In this section, I'll present and respond to these criticisms. In section 8.1, I'll present some implausible consequences of Stochastic Equivalence and Stochastic Dominance that some have taken to provide a *reductio* of these principles. In section 8.2, I'll make a case for holding on to Stochastic Equivalence and Stochastic Dominance, and then spell out in detail what a proponent of Difference Minimizing Theory should say about the implausible consequences raised in section 8.1. Of course, this defense of Stochastic Equivalence and Stochastic Dominance assumes that holding on to these principles is tenable, and Seidenfeld et al. (2009) have proved that, given certain *prima facie* plausible assumptions, holding on to Stochastic Equivalence is impossible. In section 8.3, I'll present a stripped-down version of Seidenfeld et al.'s argument. Then I'll argue that one of the key premises of the argument is false, and thus that the argument is unsound.

8.1 The Case Against

Let's look at why some have been skeptical of principles like Stochastic Equivalence and Stochastic Dominance. Consider the following case:

---

with Difference Minimizing Theory$^-$.

[39] As we'll see in section 8.3, Seidenfeld et al. (2009) argue that, given certain assumptions, a version of Stochastic Equivalence is incoherent. Lauwers & Vallentyne (2016) appeal to this result, and a related case suggested by James Joyce, to argue that we should reject both Stochastic Equivalence and Stochastic Dominance. Colyvan & Hájek (2016) also appeal to these results to justify caution regarding such principles, though they don't commit themselves to a stance on them. Smith (2014) defends a view incompatible with such principles (see Hájek (2014) for a description of a case in which Smith's view violates Stochastic Equivalence), but Smith's motivation for rejecting them is that they conflict with his positive proposal.

St. Petersburg vs Double or Nothing: A fair coin will be repeatedly tossed until it lands tails. Before the coin flipping begins, you are presented with two options. The first is to play the St. Petersburg game, which yields 2^{n-1} utility. The second is to play Double or Nothing: this is the St. Petersburg game, followed by a fair coin toss to determine whether you get double that amount (if heads) or nothing (if tails). Thus the second option yields 2^n utility if the second coin lands heads, and 0 utility if it lands tails.

Note that if you play the St. Petersburg game, the probability of getting 2^{n-1} utility (for $n \geq 1$) is $\frac{1}{2^n}$. And if you play Double or Nothing, the probability of getting 2^{n-1} utility (for $n \geq 2$) is $\frac{1}{2^n}$, and the probability of getting nothing is $\frac{1}{2}$. Thus the St. Petersburg game and Double or Nothing have the same probabilities of yielding the same utilities, except the St. Petersburg game has a $\frac{1}{2}$ probability of yielding 1 utility, while Double or Nothing has a $\frac{1}{2}$ probability of yielding 0 utility. So Stochastic Dominance entails that in this case choosing Double or Nothing is impermissible. This might strike one as bizarre: surely one should always be permitted to accept a bet for double or nothing at even odds.

In a similar vein, consider a variant of this case with a third option, Nothing or Double, which is just like Double or Nothing except that you get double if the second coin lands tails (instead of heads). The St. Petersburg game stochastically dominates both Double or Nothing and Nothing or Double, so Stochastic Dominance entails that playing the St. Petersburg game is obligatory. But note that the St. Petersburg game is stochastically equivalent to the average of Double or Nothing and Nothing or Double. This might strike one as bizarre: it suggests that the average of two gambles can have a higher value than either of the gambles it's an average of.[40]

Difference Minimizing Theory will, of course, yield the same verdicts as Stochastic Dominance (see figure 12). Let *sp* be the act of playing the St. Petersburg game, *dn* the

St. Petersburg vs Double or Nothing (superimposed)
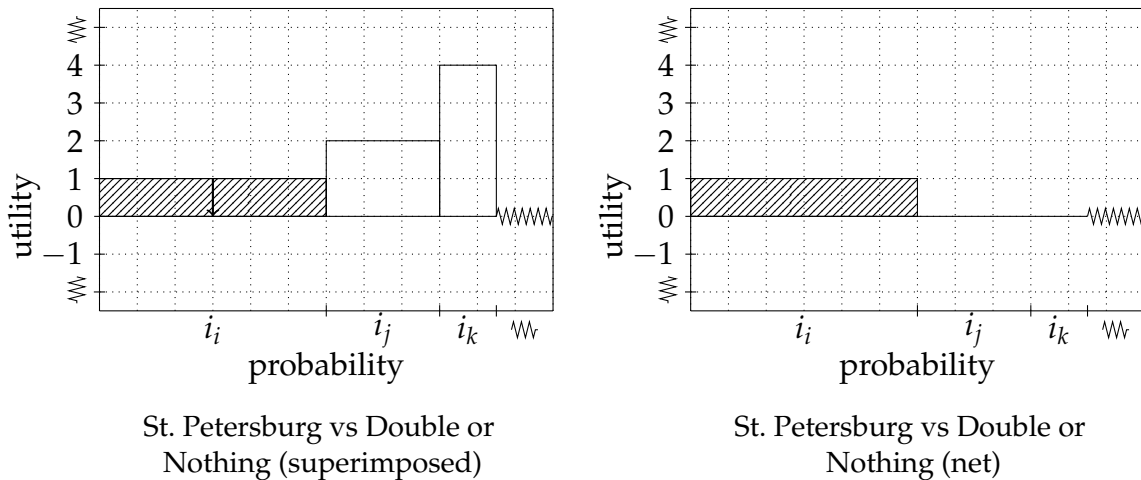
St. Petersburg vs Double or Nothing (net)

Figure 12: St. Petersburg vs Double or Nothing (difference minimizing)

act of playing Double or Nothing, and *nd* the act of playing Nothing or Double. In the St.

---

[40]This is a variant of a case from Seidenfeld et al. (2009). Colyvan & Hájek (2016) and Lauwers & Vallentyne (2016) discuss similar cases, which they credit to James Joyce.

Petersburg vs Double or Nothing case, $DMV(sp, dn) = \frac{1}{2}$, so choosing Double or Nothing is impermissible. Thus Difference Minimizing Theory denies that it's always permissible to accept a bet for double or nothing at even odds. Likewise, in the three option variant of this case, $DMV((dn + nd)/2, dn) = DMV((dn + nd)/2, nd) = \frac{1}{2} > 0$.[41] So Difference Minimizing Theory entails that the average of two gambles can effectively have a higher value than either of the gambles it's an average of.[42]

8.2 The Case in Favor

There are two ways to deal with these implausible consequences: we can reject the principles that give rise to them, or accept these principles and their consequences. So far, discussions of this issue have generally favored the first option. For example, Seidenfeld et al. (2009), Smith (2014), and Lauwers & Vallentyne (2016) all suggest we should reject Stochastic Equivalence and Stochastic Dominance, and thus reject views which entail those principles, like Difference Minimizing Theory.

I disagree. I think we should hold on to Stochastic Equivalence and Stochastic Dominance, despite these costs. Here are three reasons to hold on to these principles. First, Stochastic Equivalence and Stochastic Dominance are very plausible, akin to Moorean facts. Second, Difference Minimizing Theory entails Stochastic Equivalence and Stochastic Dominance, and we have good reason to hold on to Difference Minimizing Theory, for

---

[41]Where we're understanding "$(dn + nd)/2$" as an act which has the same probabilities of yielding every state as dn and nd, and whose utility given each state is equal to that of dn plus that of nd divided by 2.

[42]Here is a further worry one might have about Stochastic Equivalence. Consider the following *State Dominance* principle: if two (state-independent) acts a and a' are such that a yields a higher utility given every state, then a' is impermissible. (We require state-independence, because without it the principle is obviously false. If given act a states s_1/s_2 have probabilities 0.1/0.9 and yield utilities 2/-1, while given act a' states s_1/s_2 have probabilities 0.9/0.1 and yield utilities 1/-2, then a yields a higher utility than a' given every state, but a' is clearly the better choice.) It's well known that in infinite cases Pareto-style principles like State Dominance conflict with Anonymity-style principles like the Expected Isomorphism principle described in footnote 4 (e.g., see Diamond (1965)). Thus consider a case with exactly two acts a and a', which are state-independent and have the following probabilities and utilities:

Probability:	...	1/16	1/8	1/2	1/8	1/16	...
Utility given *a*:	...	-32	-8	0	8	32	...
Utility given *a'*:	...	-48	-16	-2	0	16	...
Contribution to EU of *a*:	...	-2	-1	0	1	2	...
Contribution to EU of *a'*:	...	-3	-2	-1	0	1	...

Since *a* yields a higher utility than *a'* in every state, State Dominance entails that *a'* is impermissible. But since the expected utility contributions of *a* and *a'* are isomorphic, Expected Isomorphism entails that they're on a par, and thus must both be impermissible. But that's impossible, since *a* and *a'* are the only options.

Since Stochastic Equivalence looks like an Anonymity principle, it's natural to worry that it too is incompatible with principles like State Dominance. But this worry dissipates when one realizes that Stochastic Equivalence is much weaker than the Anonymity principles required to yield these conflicts like Expected Isomorphism. For example, in the case above, *a* and *a'* are not stochastically equivalent, so Stochastic Equivalence doesn't impose any constraints on their relative permissibility. And while Difference Minimizing Theory entails Stochastic Equivalence, it does not entail – and indeed, conflicts with – Anonymity-style principles like Expected Isomorphism. To see this, note that in the case above Difference Minimizing Theory entails that only *a* is permissible, *contra* Expected Isomorphism. (I thank an anonymous referee for encouraging me to address this worry.)

Difference Minimizing Theory yields the verdicts we want in a wider range of cases than any other proposal on offer.[43] For example, to my knowledge, no other view yields the desired verdicts in all of the cases discussed in this paper.

Third, Stochastic Equivalence and Stochastic Dominance are at the heart of our understanding of what prudential rationality is. Consider what it would mean to violate Stochastic Equivalence. If Stochastic Equivalence is false, then there are cases with a pair of acts which are identical with respect to the probabilities and utilities of their outcomes, but which differ with respect to their permissibility. So if Stochastic Equivalence is false, then something other than the probabilities and utilities of outcomes is relevant to determining whether an available act is prudentially permissible.

But it's hard to conceive of how anything else *could* be relevant. Suppose, for example, that you care only about money. Then the only things that seem relevant to deciding what it's prudentially rational for you to do are how likely you think those acts are to bring about various scenarios, and how much money you'd end up with in those scenarios. One could present an account that evaluated actions in accordance with other features of the case – such as details regarding the mechanism by which you earn the money, or the shape of the decision tree leading to the act, or what plans you've made in the past, and so on – but these features all seem orthogonal to what's relevant to prudential rationality, namely, *what will best help you achieve your goal of getting as much money as possible*.[44]

Suppose we accept Difference Minimizing Theory, and thus Stochastic Equivalence and Stochastic Dominance. What should we say about the *prima facie* implausible consequences described in section 8.1? I think we should say that these consequences seem implausible because they conflict with certain principles that are plausible in finite cases, but which are revealed to be false in infinite cases.[45] Thus for proponents of Difference Minimizing Theory, the moral of section 8.1 is that certain *prima facie* plausible principles should be restricted to finite cases.

Let's work out what exactly a proponent of Difference Minimizing Theory should say. Start with the first implausible consequence discussed in section 8.1, that taking a bet for double or nothing can be impermissible. Consider some act a which has a probability p of bringing about some outcome $o = a \wedge s$ with utility u. Let the $o(0,2)$-*replacement* of a be an act which is identical to a, except that outcome o is replaced with a pair of outcomes with probabilities $\frac{p}{2}$ and utilities 0 and $2u$, respectively. It's natural to think that any act a will be on a par with any of its $o(0,2)$-replacements.[46] And after thinking through some examples, it's natural to conclude that this principle generalizes:

Replacement Parity: Any act a will be on a par with any other act that you can construct

---

[43]Proposals by Smith's (2014) and Bartha (2016) might come to mind, but these proposals either don't yield the verdicts we want (in the first case) or don't really yield verdicts at all (in the second); see footnote 37.

[44]For a contrasting view, see Smith (2014), who explicitly rejects this conception of prudential rationality. (More precisely, he rejects what he calls "evaluative compositionality", the claim that how much an agent values a bet should be determined entirely by the probabilities and utilities of its possible outcomes.)

[45]Of course, opponents of Stochastic Equivalence and Stochastic Dominance will say something similar, holding that Stochastic Equivalence and Stochastic Dominance are plausible in finite cases, but are revealed to be false in infinite cases.

[46]Assuming, as we did in section 2, that we're working with finite utilities.

from it via a sequence of $o(0,2)$-replacements.[47]

Proponents of Difference Minimizing Theory will reject this principle. They'll maintain that Replacement Parity should be restricted to finite sequences of $o(0,2)$-replacements. What they take the St. Petersburg vs Double or Nothing case to show is that Replacement Parity breaks down when we consider acts constructed from infinite sequences of $o(0,2)$-replacements. For while the St. Petersburg and Double or Nothing games are linked by an infinite sequence of $o(0,2)$-replacements, they're not on a par.[48]

Let's turn to the second implausible consequence discussed in section 8.1, that the average of a pair of bets can be better than either of the individual bets. Given two acts a and b with the same probabilities of yielding each state, let $av_{a,b}$ be the act corresponding to the average of these acts – an act which has the same probabilities of yielding each state as a and b, and whose utilities given that state are the sum of those yielded by a and b divided by 2. After thinking through some examples, it's natural to think that $av_{a,b}$ must have an effective value in-between a and b. In the context of Difference Minimizing Theory, this suggests the following principle:

Average Betweenness: If $DMV(a,c) = x$, $DMV(b,c) = y$, and $DMV(av_{a,b},c) = z$ (where $x, y, z \in \mathbb{R}$), then z must lie in the interval $[x, y]$.

Proponents of Difference Minimizing Theory will reject this principle. They'll maintain that Average Betweenness only holds in cases with finitely many states. What they take the three-option variant of the St. Petersburg vs Double or Nothing case to show is that Average Betweenness breaks down in cases with infinitely many states. In particular, we'll find that $DMV(dn, sp) = -\frac{1}{2}$ and $DMV(nd, sp) = -\frac{1}{2}$, but $DMV(av_{dn,nd}, sp) = 0$ (since $av_{dn,nd} = sp$). Thus the average of two acts ($av_{dn,nd}$) can be better than either of the acts it's an average of (dn and nd).

8.3   Seidenfeld et al.'s Theorem

Let's turn to consider a different kind of challenge to Stochastic Equivalence. Seidenfeld et al. (2009) present a theorem showing that, given certain *prima facie* plausible assumptions, Stochastic Equivalence is false. By taking these assumptions as premises, we can construct a straightforward argument against Stochastic Equivalence. If we're to defend Difference Minimizing Theory, we need to figure out where this argument goes wrong.

Seidenfeld et al. (2009) work in a framework that differs from the one employed here in several ways, so we'll need to do a little work to bring these discussions into contact with each other. A key difference between our approaches is that Seidenfeld et al. are looking at accounts of rational preferences over acts, not accounts of which acts are rationally permissible. Asking for an account of rational preference is strictly more demanding than

---

[47]I.e., in any case with a pair of acts a and b, where b can be constructed from a via a sequence of $o(0,2)$-replacements, a and b will have the same deontic status (e.g., both permissible or impermissible).

[48]After appropriate reflection, *opponents* of Difference Minimizing Theory should reject this principle as well, since one can use an infinite sequence of $o(0,2)$-replacements to, say, turn an act that yields 1 utility no matter what into an act that yields 0 utility no matter what. This gives us a further reason to believe that the intuition that double or nothing bets should always be fair (even in infinite cases) is mistaken. (Thanks to Kenny Easwaran here.)

asking for an account of what's permissible. For one can employ rational preference to fix what's permissible – an act is permissible *iff* no other act is rationally preferable to it. But one can't use what's permissible to fix rational preferences – two acts might both be impermissible, and yet one rationally preferable to the other.

In a similar vein, I've presented Stochastic Equivalence as a claim about permissibility, but Seidenfeld et al. are concerned with a stronger claim regarding rational preference:

Stochastic Equivalence$_p$: If two available acts a and a' have the same probabilities of yielding the same utilities, then it's irrational to prefer one of them over the other.

A natural way to work around this difference is to consider a relative of Difference Minimizing Theory – call it "Difference Minimizing Theory$_p$" – which determines which preferences are rational instead of which acts are permissible. Let a subject's *preferences* over a set of acts A be a binary relation $\prec$ that's irreflexive ($\forall a \in A, a \not\prec a$) and transitive ($\forall a, b, c \in A$, if $a \prec b$ and $b \prec c$, then $a \prec c$). Seidenfeld et al. impose the further condition that $\prec$ be negatively transitive – i.e., if neither a nor b is preferred to each other, and neither b nor c is preferred to each other, then neither a nor c is preferred to each other. I've left out this condition because, as we'll see below, this condition is precisely what's at issue. Following Seidenfeld et al., we can also introduce an "indifferent" preference relation $\sim$, such that if $a \not\prec b$ and $b \not\prec a$, then $a \sim b$. Given this, we can let Difference Minimizing Theory$_p$ be the view that if $DMV(a, b)$ floats above 0, then $a \succ b$.

We can gloss over the rest of the differences between Seidenfeld et al.'s approach and the one in this paper by focusing on a special case. For convenience, let's introduce the following terminology: given two acts a and b that are probabilistically independent of the states, let $[a - b]$ be an act which has the same probabilities of yielding every state as a and b, and whose utility given each state is equal to that of a minus that of b. Now consider the following case:

St. Petersburg Panoply: A fair coin will be repeatedly tossed until it lands tails. Before the coin flipping begins, you are presented with seven options. The first is to play the St. Petersburg game (sp), which yields 2^{n-1} utility. The second is to play Double or One (do): this is the St. Petersburg game, followed by a fair coin toss to determine whether you get double that amount (if heads) or 1 utility (if tails). The third is to play One or Double (od): this is the same as Double or One, but with the results of the second coin toss reversed. The fourth option is to play $[sp - od]$, which yields $2^{n-1} - 1$ utility if the second coin toss lands heads, and -2^{n-1} utility if the second coin toss lands tails. The fifth option is to play $[do - sp]$, which yields 2^{n-1} utility if the second coin toss lands heads, and $1 - 2^{n-1}$ utility if the second coin toss lands tails. The sixth option is to play $[do - sp] - [sp - od]$, which yields 1 utility no matter what. The seventh option is to play Nothing (**0**), the null act which yields 0 utility no matter what.[49]

Note that the St. Petersburg game, Double or One, and One or Double, are all stochastically equivalent. If you play the St. Petersburg game, the probability of getting 2^{n-1} utility (for $n \geq 1$) is $\frac{1}{2^n}$. And if you play Double or One or One or Double, the probability of getting 2^{n-1} utility (for $n \geq 2$) is $\frac{1}{2^n}$, and the probability of getting 1 is $\frac{1}{2}$, which is precisely

---

[49]This is a slightly simplified version of Seidenfeld et al.'s (2009) Example 3.1.

the same thing.

We can use the St. Petersburg Panoply case to construct a version of Seidenfeld et al.'s argument against Stochastic Equivalence$_p$ as follows:

The St. Petersburg Panoply Argument:

P1. A rational agent's preferences are irreflexive, transitive and negatively transitive.

P2. A rational agent's preferences satisfy:

   Coherent Indifference: If $a \sim b$, then $[a - b] \sim [b - a] \sim \mathbf{0}$.[50]

P3. A rational agent's preferences satisfy:

   Coherent Strict Preference: If a and b are probabilistically independent, and there's some positive ϵ such that, for every state, a yields at least ϵ more utility than b, then $a \succ b$.

SE$_p$. Suppose for *reductio* that Stochastic Equivalence$_p$ holds.

L1. Stochastic Equivalence$_p$ entails that a rational agent in the St. Petersburg Panoply case must have preferences such that $sp \sim do \sim od$. (SE$_p$)

L2. Coherent Indifference then entails that a rational agent in this case must have preferences such that: $[do - sp] \sim \mathbf{0}$ and $[sp - od] \sim \mathbf{0}$. (P2,L1)

L3. Since rational preferences are negatively transitive, a rational agent in this case must have preferences such that: $[do - sp] \sim [sp - od]$. (P1,L2)

L4. Applying Coherent Indifference again entails that a rational agent in this case must have preferences such that: $[[do - sp] - [sp - od]] \sim \mathbf{0}$. (P2,L3)

L5. Coherent Strict Preference entails that the rational agent in this case must have preferences such that: $[[do - sp] - [sp - od]] \succ \mathbf{0}$.[51] (P3)

C. By *reductio*, Stochastic Equivalence$_p$ is false. (L4,L5)

This argument is valid, and it's premises are *prima facie* plausible. Nevertheless, I think this argument is not convincing. This is because when we look more carefully at the first premise, and the negatively transitive requirement in particular, we can see that this requirement is implausible.

Consider two things one might mean when one says that a subject is "indifferent" between two acts. One way of being "indifferent" between two acts is to effectively take them to have the same value; call this notion of indifference "parity" ($\sim_p$). So, for example, the Coin Toss bet (*ct*) from section 2 that yields 2 utility if heads and -2 utility if tails, and the Nothing bet (*n*) that yields 0 utility no matter what, are on a par: $ct \sim_p n$. In terms of difference minimizing values, we would expect two acts a and b to be on a par whenever $DMV(a, b) = 0$.

Another way of being "indifferent" between two acts is to not prefer either of them; this is the notion of "indifferent" picked out by Seidenfeld et al.'s definition of $\sim$. Call this notion of indifference "no-victor" ($\sim_{nv}$). The no-victor relation is strictly broader than the parity relation. If parity holds between two acts then so will no-victor relation, since if

---
[50]Where again $\mathbf{0}$ is a null act which yields 0 utility no matter what.
[51]Recall that $[[do - sp] - [sp - od]]$ yield 1 utility no matter what, while $\mathbf{0}$ yields 0 utility no matter what.

two acts are on a par then neither will be preferred to the other. But there are pairs of acts that the no-victor relation holds between that aren't on a par. For example, consider the acts $[do - sp]$ and $\mathbf{0}$ from the St. Petersburg Panoply case. $[do - sp]$ looks roughly like a combination of a St. Petersburg bet with a negative St. Petersburg bet, with ever increasing positive and negative terms of ever decreasing probability. How should we weigh $[do - sp]$ against $\mathbf{0}$? It's not clear that we should prefer $[do - sp]$ to $\mathbf{0}$, or vice versa. But it's also not clear that they're on a par – it's not clear that they have the same effective value, so that if (say) we added one utility to one of these bets it would now be clearly preferable to the other. And the difference minimizing values support this impression; $DMV([do - sp], \mathbf{0}) =$ undefined. So these two acts seem to bear the no-victor relation – $[do - sp] \sim_{nv} \mathbf{0}$ – even though they're not on a par. In terms of difference minimizing values, we would expect the no-victor relation to hold between two acts a and b whenever neither $DMV(a, b)$ nor $DMV(b, a)$ floats above 0.

Now let's return to the first premise of the argument, and in particular, the assumption that rational preferences must be negatively transitive. The negatively transitive condition effectively requires both (i) that some indifference relation hold between any pair of acts that aren't preferred to each other, and (ii) that this indifference relation be transitive. But on neither understanding of "indifferent" will both of these requirements be plausible.

Suppose we understand "indifferent" as the no-victor relation, as Seidenfeld et al.'s definition of $\sim$ suggests. Then (i) holds, since any pair of acts a and b such that neither $a \prec b$ nor $b \prec a$ will bear the no-victor relation. But (ii) does not, since it's implausible that the no-victor relation is transitive; if $a \sim_{nv} b$, and $b \sim_{nv} c$, it doesn't follow that $a \sim_{nv} c$. Consider a version of the St. Petersburg Panoply case in which there's an eighth option, $[dn - sp]$, where dn (Double or Nothing) is an act just like Double or One, but where you get 0 utility instead of 1 if the second coin toss lands tails. $[dn - sp]$ bears the no-victor relation to $\mathbf{0}$ for the same reasons that $[do - sp]$ bears the no-victor relation to $\mathbf{0}$. But it doesn't follow that $[do - sp]$ and $[dn - sp]$ bear the no-victor relation to each other, for $[do - sp]$ should be preferred to $[dn - sp]$; $[do - sp]$ stochastically dominates $[dn - sp]$, and $DMV([do - sp], [dn - sp]) = 1/2$.

Suppose instead we understand "indifferent" as the parity relation. Then (ii) holds, since it's plausible that parity is transitive; if a is on a par with b, and b is on a par with c, then a should be on a par with c. But (i) does not, since it's not plausible that any pair of acts a and b such that neither $a \prec b$ nor $b \prec a$ must be on a par. In the St. Petersburg Panoply case, $[do - sp] \not\prec \mathbf{0}$ and $\mathbf{0} \not\prec [do - sp]$, but it doesn't follow that $[do - sp]$ and $\mathbf{0}$ are on a par.

In light of this, I suggest that we should reject the claim that rational preferences must be negatively transitive, and thus reject the first premise of the argument. If this is correct, then Seidenfeld et al.'s result is ultimately not a threat to Stochastic Equivalence$_p$ and views which entail it like Difference Minimizing Theory$_p$. For the argument requires a condition on rational preference – that they be negatively transitive – that we should reject.

9 Conclusion

In this paper I've developed a proposal for handling cases without finite expected values, Difference Minimizing Theory. We can see the path from standard decision theory to Difference Minimizing Theory as having five steps. The first step is to allow expected utilities to take on infinite values, and in particular, extended real values (section 2). The second step is to follow Colyvan (2008) and use relative expected utilities instead of expected utilities to assess acts (section 4). The third step is to employ a difference minimizing approach to evaluate pairs of acts (section 6). The fourth step is to follow Easwaran (2014b) and employ principal values instead of ordinary expectations when evaluating (pairs of) acts (section 7.1). The fifth step is to follow Lauwers & Vallentyne (2016) and extend the proposal to allow for non-definite values (section 7.3).

The resulting view yields satisfactory verdicts in a broader range of cases than its predecessors. And it allows us to retain Stochastic Equivalence and Stochastic Dominance – two principles that are arguably at the heart of our understanding of prudential rationality. While maintaining these principles leads to some surprising results (section 8.2), these are results we should learn to accept. For when infinities get involved, some surprising results are inevitable.

I've argued that Difference Minimizing Theory is an improvement over its predecessors. But should we think of it as the ultimate theory of prudential rationality, or as just another step toward the ultimate theory? I think it depends on the status of Difference Minimizing Theory's treatment of cases with undefined DMVs.

In standard decision theory, there are cases in which the expected utilities of various acts are undefined, such as in the Altadena vs Pasadena case. In these cases, standard decision theory defaults to permissibility – if $EU(a)$ is undefined, then there won't exist an act a' such that $EU(a') > EU(a)$, so a will be permissible. And this relatively coarse-grained way of assigning verdicts in such cases is a demerit of standard decision theory, for cases like the Altadena vs Pasadena case require more fine-grained discriminations in order to get plausible verdicts (e.g., that only playing the Altadena game is permissible). And as we've seen, it's possible to construct plausible alternatives to standard decision theory which yield these more fine-grained verdicts.

In Difference Minimizing Theory, there are likewise cases in which difference minimizing values are undefined.[52] In these cases Difference Minimizing Theory likewise defaults to permissibility – if $DMV(a', a)$ isn't well-defined, then it won't be the case that $DMV(a', a)$ floats above 0, and Difference Minimizing Theory won't take that comparison to be a reason to deem a impermissible. And in the extreme case in which there are no well-defined comparisons between acts, it will follow that all acts are permissible. Again,

---

[52]Here is an example of a case in which there is no well-defined difference-minimizing value. Consider a case in which two bets are offered. A fair coin will be repeatedly tossed until it land tails (n times), followed by a final fair coin toss. The first bet a yields 0 utility, regardless of the outcome of the final coin toss. The second bet a' yields 2^{n+1} utility if the final coin toss lands heads, and -2^{n+1} utility if the final coin toss lands tails. The difference between these payoffs will look like a sum over the mixture of the St. Petersburg game and its negative. The resulting partial sums won't unconditionally converge to any extended real number, and the move to stable principle values doesn't help since (as Easwaran (2014b) notes) a mixture of the St. Petersburg game and its negative won't have a stable principle value. Thus the difference-minimizing value of a over a' (and vice versa) will be undefined.

this is a relatively coarse-grained way of assigning verdicts in such cases. Should we take this to be a demerit of Difference Minimizing Theory?

It depends on whether there are cases in which more fine-grained discriminations are required to get plausible verdicts, and whether there are plausible theories that yield those fine-grained prescriptions. If there aren't, then I take Difference Minimizing Theory to be a plausible candidate for the ultimate theory of prudential rationality. But if there are, then it's reasonable to think of Difference Minimizing Theory as merely another step on the road to the ultimate theory, a theory which, perhaps, will build on Difference Minimizing Theory in the same way that Difference Minimizing Theory builds on the theories of Colyvan (2008), Easwaran (2014b), and Lauwers & Vallentyne (2016).[53]

References

Arntzenius, F. (2014). Utilitarianism, decision theory and eternity. *Philosophical Perspectives*, *28*(1), 31–58.

Bartha, P. (2007). Taking stock of infinite value: Pascal's wager and relative utilities. *Synthese*, *154*(1), 5–52.

Bartha, P. F. A. (2016). Making do without expectations. *Mind*, *125*(499), 799–827.

Bostrom, N. (2011). Infinite ethics. *Analysis and Metaphysics*, *10*, 9–59.

Chen, E. K., & Rubio, D. (forthcoming). Surreal decisions. *Philosophy and Phenomenological Research*.

Collins, J. (1996). Supposition and choice: Why 'causal decision theory' is a misnomer.

Colyvan, M. (2008). Relative expectation theory. *Journal of Philosophy*, *105*(1), 37–44.

Colyvan, M., & Hájek, A. (2016). Making ado without expectations. *Mind*, *125*(499), 829–857.

Diamond, P. (1965). The evaluation of infinite utility streams. *Econometra*, (33), 170–177.

Easwaran, K. (2008). Strong and weak expectations. *Mind*, *117*(467), 633–641.

Easwaran, K. (2014a). Decision theory without representation theorems. *Philosophers' Imprint*, *14*(27).

Easwaran, K. (2014b). Principal values and weak expectations. *Mind*, *123*(490), 517–531.

Hájek, A. (2003). Waging war on pascal's wager. *Philosophical Review*, *112*(1), 27–56.

Hájek, A. (2014). Unexpected expectations. *Mind*, *123*(490), 533–567.

Joyce, J. M. (1999). *The Foundations of Causal Decision Theory*. Cambridge University Press.

Lauwers, L., & Vallentyne, P. (2016). Decision theory without finite standard expected value. *Economics and Philosophy*, *32*(3), 383–407.

Nathan, A. (1984). False expectations. *Philosophy of Science*, *51*(1), 128–136.

Nover, H., & Hájek, A. (2004). Vexing expectations. *Mind*, *113*(450), 237–249.

Seidenfeld, T., Schervish, M. J., & Kadane, J. B. (2009). Preference for equivalent random variables: A price for unbounded utilities. *Journal of Mathematical Economics*, *45*, 329–340.

Smith, N. J. J. (2014). Is evaluative compositionality a requirement of rationality? *Mind*, *123*(490), 457–502.