Routledge
Taylor & Francis Group

Revisiting Risk and Rationality: a reply to Pettigrew and Briggs

Lara Buchak

Department of Philosophy, UC Berkeley, Berkeley, CA, USA

ABSTRACT
I have claimed that risk-weighted expected utility (REU) maximizers are rational, and that their preferences cannot be captured by expected utility (EU) theory. Richard Pettigrew and Rachael Briggs have recently challenged these claims. Both authors argue that only EU-maximizers are rational. In addition, Pettigrew argues that the preferences of REU-maximizers can indeed be captured by EU theory, and Briggs argues that REU-maximizers lose a valuable tool for simplifying their decision problems. I hold that their arguments do not succeed and that my original claims still stand. However, their arguments do highlight some costs of REU theory.

1. Introduction

In *Risk and Rationality*, I argued against what we might call the expected utility (*EU*) *thesis*: the claim that individuals are rational if and only if they maximize EU. I made a case for two claims. The first is that there are preferences that seem reasonable but cannot be described as cases of EU maximization, or even *re-described* as such by individuating outcomes more finely. The second is that the standard arguments for the EU thesis are not in fact strong enough to support that thesis. I proposed that we jettison the EU thesis and adopt the weaker risk-weighted expected utility (*REU*) *thesis*: individuals are rational if and only if they maximize REU.

A key difference between the EU thesis and the REU thesis is that while the former holds that there are only two internal attitudes that are combined in preferences – utilities and credences – the latter holds that there are three: utilities, credences, and risk attitudes.[1] Roughly, utilities measure how much an individual

---

CONTACT Lara Buchak ✉ buchak@berkeley.edu

values a particular outcome, credences measure how likely an individual takes a given state to be, and risk-attitudes measure how much an individual takes into account what happens in worse states of an act as opposed to better states. For example, individuals who are risk-avoidant REU-maximizers give more weight to what happens in the worst-case scenario than the best, even when these scenarios are equally likely. This contrasts with EU-maximization, according to which individuals give equal weight to scenarios that have equal probability. Whereas in EU theory, the value contribution of a particular outcome to the total value of the act is that outcome's utility times that outcome's probability, in REU theory, the value contribution of a particular outcome to the total value of the act is that outcome's utility times a factor which depends on both the outcome's probability and its position in the act's ordering.[2]

Richard Pettigrew (2015) and Rachael Briggs (2015) advance new considerations in favor of EU theory. Pettigrew advances a novel argument for the EU thesis, namely that an individual whose preferences conform to EU maximization will be better at estimating the utility values of the acts she is choosing among. He also proposes that those with REU-maximizing preferences can be redescribed as EU-maximizers. Specifically, he proposes a way of combining risk attitudes and utilities so that there are ultimately only two fundamental internal states. Consequently, he concludes that while the EU thesis holds, we can agree about which actual preferences are rational, since REU-maximization is merely a special case of EU-maximization.

Briggs has two main criticisms of REU theory. The first is a rejoinder to one of my arguments against the EU thesis. I argued that REU-maximizers are diachronically rational, and Briggs challenges this claim: in particular, while I claimed that both sophisticated and resolute choice can vindicate the diachronic choice behavior of REU-maximizers, she argues that sophisticated choice cannot in fact vindicate their diachronic choice behavior. Briggs's other criticism stems from the fact that for REU-maximizers, sub-acts lack stable utility values. Because sub-acts lack stable utility values, she claims, REU-maximizers lose a valuable tool for simplifying decision problems. This isn't an argument for the EU-thesis per se, but it foregrounds a practical cost of REU theory: REU-maximizers face more difficulties actually making ordinary decisions than EU-maximizers do.

Each of these arguments is insightful, original, and challenging. In this article, I outline responses on behalf of REU theory. In addition to allowing us to continue to deny the EU thesis, these responses will illuminate why positing three factors in preference is different from positing two potentially more complex factors; how we ought to think about the aim of decision theory; and what are the genuine costs of accepting REU theory.

Following Briggs, it will be helpful to anchor the discussion using two individuals: Eulalie, who maximizes EU, and Rhoda, who maximizes REU with $r(p) = p^2$. My claims will be that Rhoda cannot be redescribed as an EU-maximizer, that

she is both synchronically and diachronically rational, and that she is able to simplify her decision problems in natural ways.

2. Pettigrew's first challenge: re-description

Let us begin with Pettigrew's argument that Rhoda cannot be redescribed as an EU-maximizer. Redescription involves making the outcomes (the original bearers of utility) more fine-grained, and then assigning values to these fine-grained outcomes, which we can call utility* values. The upshot will be that even though there is no utility assignment to coarse-grained outcomes according to which the individual maximizes EU, there is a utility* assignment to the fine-grained outcomes according to which she maximizes EU*. Thus, Rhoda *is* an EU- maximizer, merely one with a utility function that ranges over fine-grained outcomes.

The dilemma I presented for redescription of REU-maximizers is this: there is no strategy for fine-graining outcomes and assigning utility* to fine-grained outcomes that both generates a consistent utility* function and survives what I call the proliferation problems (Buchak 2013, Ch. 4). The proliferation problems occur when the EU theorist fine-grains in such a way that too much is up for grabs: for example, if outcomes are maximally fine-grained and there are no extra-theoretical constraints on how we assign utility* to them, then we can attribute to the agent one of any number of utility functions, and any probability function whatsoever. But the only other alternative, I claimed, is to leave too little up for grabs: if outcomes are coarse-grained enough or if there are enough extra-theoretical constraints on utility* to avoid the proliferation problems, then we cannot consistently assign expectational utility* values at all.

One redescription that avoids the proliferation problems but suffers from the inconsistency problem is what I called the 'comonotonic individuation' (Buchak 2013, 141). Recall that in EU theory, the value contribution of a particular outcome is that outcome's utility times that outcome's probability, and in REU theory, the value contribution of a particular outcome is that outcome's utility times a factor which depends on both the outcome's probability and its position in the act's ordering. The comonotonic individuation first redescribes outcomes as <outcome, act> pairs. Letting 'utility' refer to an outcome's utility according to REU theory and 'utility*' refer to an <outcome, act> pair's utility according to EU theory, the comonotonic individuation assigns utility* to each pair as follows. It first sets the *utility* contribution* of the <outcome, act> pair equal to the utility contribution of the outcome to the utility of the act, and it then sets the *utility** of the pair equal to the utility* contribution divided by the probability of the outcome according to the act.

More formally, where events in act h are ordered (1 to n) from worst to best and x_i stands for the outcome that results from event E_i, the comonotonic individuation sets $u^*(x_i, h) = t_i u^*(x_i)$, where t_i is the marginal difference in r that $p(E_i)$ makes, divided by $p(E_i)$:

$$t_i = \begin{cases} \dfrac{r(p(E_i \cup ... \cup E_n)) - r(p(E_{i+1} \cup ... \cup E_n))}{p(E_i \cup ... \cup E_n) - p(E_{i+1} \cup ... \cup E_n)} & \text{if } i = 1, ..., n-1 \\ \dfrac{r(p(E_i))}{p(E_i)} & \text{if } i = n \end{cases}$$

(Pettigrew uses c to stand for the agent's credence (subjective probability) function, but I use p to match my earlier terminology.) Notice that t_i is defined almost identically to Pettigrew's s_i: the key difference is that t_i is defined as the coefficient of the difference that the utility of *outcome* x_i makes, while s_i is defined as the coefficient of the difference that *utility* u_i makes. These two values are different because there can be multiple outcomes with the same utility. The comonotonic individuation treats outcomes with the same utility separately – it takes as a starting point the value contribution that each distinct outcome makes – but Pettigrew's individuation treats them together: it takes as a starting point the value contribution that all of the outcomes with a particular utility make together. We will see that this difference is crucial.

So, for example, let us consider the three-outcome act $g = \{E_1, x_1; E_2, x_2; E_4, x_4\}$, where the ordered utilities and the probabilities are given by:

	x_1	x_2	x_4
u	3	5	6
	E_1	E_2	E_4
p	0.3	0.3	0.4

(We are examining an example that bears some similarity to Pettigrew's, so we can see clearly what problem his individuation solves.) The REU of this act for Rhoda (who has $r(p) = p^2$) is 4.14.[3]

According to the comonotonic individuation, Rhoda's utility* values are:
$u^*(x_1, g) = 1.7(3) = 5.1$
$u^*(x_2, g) = 1.1(5) = 5.5$
$u^*(x_4, g) = 0.4(6) = 2.4$

Therefore, the 'new' EU of the act agrees with the 'old' REU: the EU is $(0.3)(5.1) + (0.3)(5.5) + (0.4)(2.4) = 4.14$.

So far, so good. But this strategy doesn't work in general, because some acts can be ordered in more than one way: when an act contains two events that yield outcomes of the same utility, either event can come first in the 'ordered' act. And the comonotonic individuation, applied to these two orderings, won't yield a consistent utility* assignment to the outcomes in those acts. To see this, let us consider Pettigrew's example, the four-outcome act $h = \{E_1, x_1; E_2, x_2; E_3, x_3; E_4, x_4\}$, whose ordered utilities and probabilities are given by:

	x_1	x_2	x_3	x_4
u	3	5	5	6
	E_1	E_2	E_3	E_4
p	0.3	0.2	0.1	0.4

Applying the comonotonic individuation to this ordered act, Rhoda's utility* values are:

$$u^*(x_1, h) = 1.7(3) \quad u^*(x_2, h) = 1.2(5) \quad u^*(x_3, h) = 0.9(5) \quad u^*(x_4, h) = 0.4(6)$$

However, the following is an *equivalent* description of the act, *also* ordered from worst to best:

	x_1	x_3	x_2	x_4
u	3	5	5	6
	E_1	E_3	E_2	E_4
p	0.3	0.1	0.2	0.4

And applying the comonotonic individuation to this (equivalent) ordered act, Rhoda's utility* values are:

$$u^*(x_1, h) = 1.7(3) \quad u^*(x_2, h) = 1(5) \quad u^*(x_3, h) = 1.3(5) \quad u^*(x_4, h) = 0.4(6)$$

These two assignments are incompatible. As we can see, the comonotonic individuation is not well-defined: it does not yield a consistent assignment of utility* values to outcomes. That is because, as I pointed out (p. 143), 'the quantity that is well-defined is the total contribution that *all outcomes of a particular utility value* make to the REU of an act.'

Because an outcome can occupy more than one position in an act's ordering, I was pessimistic about whether any re-description that defined utility* in terms of an outcome's position in the act's ordering could work. However, Pettigrew ingeniously uses the quoted fact to define a different kind of redescription than the ones I considered. Specifically, he first amalgamates all outcomes with the same utility value. He then looks at how much *these outcomes together* contribute to the overall value of an act. Next, he assigns this value to these outcomes together as their utility* contribution. Finally, the utility* of the <outcome, act> pair for each of the outcomes will again be their utility* contribution divided by their total probability according to the act. This *will* give a consistent assignment. Thus, Pettigrew has successfully answered an objection I had to the redescription strategy, namely that there is no way of redescribing that will be consistent and avoid the proliferation problem.

Pettigrew's individuation manages to avoid the proliferation problem while remaining well-defined, because it directly incorporates the quantity that is well-defined. However, Pettigrew's strategy faces another problem.

To see this, let us first look at the utility* assignment we get in Pettigrew's example.

	x_1	x_2	x_3	x_4
u	3	5	5	6
u^*	5.1	5.5	5.5	2.4

Notice that the worst outcome and the middle outcomes are all ranked *higher* than the best outcome, according to utility*. Since u^* is supposed to characterize

the agent's preferences, this means that the agent prefers $<x_1, h>$, $<x_2, h>$, and $<x_3, h>$, each to $<x_4, h>$. And this is true even though she'd rather have a guaranteed x_4 than a guaranteed x_1.

Perhaps this is not too devastating a worry, and the relative value of a fine-grained outcome just doesn't reflect which outcome one would prefer to have for certain. A fine-grained outcome is, after all, a coarse-grained outcome with a particular probability, in addition to one with a particular position in an act. However, there is a more serious worry: *two utility* assignments that represent the same preference ordering will not in general order the fine-grained outcomes in the same way.*

To see this, first notice that Pettigrew's Theorem 4 (the re-interpretation of the REU Representation Theorem) includes the clause 'unique up to affine* transformation.' A utility* function is unique up to affine* transformation if the utility function from which it is defined is unique up to affine transformation. Thus, if we apply an affine transformation to u and then calculate u^* of the result, u^* will represent the original preference ordering, in Pettigrew's sense of 'represent.' With this in mind, here are the u^* values that result from our original u shifted down by three units, and down by five units, respectively:

	x_1	x_2	x_3	x_4
u	0	2	2	3
u^*	0	2.2	2.2	1.2
u	−2	0	0	1
u^*	−3.4	0	0	0.4

All three utility* assignments represent the exact same preference ordering. However, the u^*'s order the fine-grained outcomes differently. This means that for a particular family of u^*'s that purport to represent the same preferences, even the *ordering* of outcomes according to u^* isn't preserved. This is because u^* is defined as the product of two factors, u and a function of both r and p, and applying an affine transformation to just the first factor means that we will not in general replicate the same ordering for each u in a family.

The extent to which one thinks that this feature is a serious problem for EU theory will depend on how one interprets the utility* function. Although some decision theorists – realists – hold that utility values exist independently of preferences, it is more typical to hold that they are constructed from preferences, and thus that cardinal comparisons of utility are meaningless in themselves. However, we only need to be *ordinal* realists about utility* for the feature under discussion to be a problem. If there is a fact of the matter about how Rhoda *ranks* fine-grained outcomes, then utility* as Pettigrew defines it cannot track this fact. Furthermore, even constructivists about ordinal preferences will have to admit that Theorem 4 doesn't yield a value function that corresponds to any recognizable feature of preferences.

This highlights a more general difference between the REU Representation Theorem and Theorem 4. On the REU Theorem, the role of each of the three

internal attitudes is clear (Buchak 2013, Ch. 3). Even if we are not realists about these attitudes, there are three clearly distinct and recognizable features of preferences. Utilities determine, or are determined by, which pairs of outcomes constitute 'equal tradeoffs' in acts that order events in the same way.[4] Whether one assigns higher probability to one event than to another determines, or is determined by, whether one would rather bet on that event than on the other. The risk function determines, or is determined by, how much what happens in an event with a particular probability matters to the evaluation of an act, when that event is the best event: how much compensation in the rest of the act is required to make up for what happens in that event. So we have a nice story about the role each entity plays in preferences, which fits (at least somewhat) our pre-theoretical intuitions about what the role of such an entity would be. But we don't have this on Pettigrew's story: utility* is a muddle of two of these entities. Forcing two distinct attitudes to count as a single hybrid attitude results both in the hybrid attitude not being a readily recognizable contributor to preferences and in our being unable to make sense of the values we assign to it.

3. Pettigrew's second challenge: estimates

Let us now turn to Pettigrew's other challenge. Pettigrew provides a novel argument for the EU thesis: for Rhoda (as for any non-EU-maximizer), there will be some EU assignment of values to acts that *dominates* her assignment of values to acts, in the following sense. In every state, the EU assignment is closer to the true value of the act than Rhoda's assignment is.

We need to be clear about what this argument says, and in particular two things it does not say. First, it does not say that Rhoda will select a dominated *act*, an act such that there is an alternative act she prefers no matter how the world turns out. The object that is dominated here is an *estimate of act values*, not an act itself. It is dominated in the sense that there is an alternative assignment that is a better estimate no matter how the world turns out. Second, the argument does not say that any *one* act's estimate is dominated in all worlds in this sense; rather, the estimate of all the acts, taken as a whole, is dominated in the sense that the *sum* of the errors for each act is higher than it could be in each world. Let me illustrate with a simple example.

Suppose we have two states and three acts, with the following payoffs (in utility):

	HEADS	TAILS
Bet on heads	100	0
Bet on tails	0	100
Don't bet	50	50

Then Eulalie (our EU-maximizer) and Rhoda give the following assignments of act values and truth values:

	Eulalie	Rhoda
Value of act 'Bet on Heads'	50	25
Value of act 'Bet on Tails'	50	25
Value of act 'Don't Bet'	50	50
Truth value of 'Coin lands Heads'	0.5	0.5
Truth value of 'Coin lands Tails'	0.5	0.5

We now ask about the distance of each assignment from the true value of each entity in each state. Following Pettigrew, we calculate this distance using the quadratic distance measure:

Distance of Eulalie's assignments in each state

	Distance from truth, in HEADS	Distance from truth, in TAILS
Value of act 'Bet on Heads'	$(50 − 100)^2 = 2500$	$(50 − 0)^2 = 2500$
Value of act 'Bet on Tails'	$(50 − 0)^2 = 2500$	$(50 − 100)^2 = 2500$
Value of act 'Don't Bet'	$(50 − 50)^2 = 0$	$(50 − 50)^2 = 0$
Truth value of 'Coin lands Heads'	$(1 − 0.5)^2 = 0.25$	$(0 − 0.5)^2 = 0.25$
Truth value of 'Coin lands Tails'	$(0 − 0.5)^2 = 0.25$	$(1 − 0.5)^2 = 0.25$

Distance of Rhoda's assignments in each state

ACT	Distance from truth, in HEADS	Distance from truth, in TAILS
Value of act 'Bet on Heads'	$(25 − 100)^2 = 5625$	$(25 − 0)^2 = 625$
Value of act 'Bet on Tails'	$(25 − 0)^2 = 625$	$(25 − 100)^2 = 5625$
Value of act 'Don't Bet'	$(50 − 50)^2 = 0$	$(50 − 50)^2 = 0$
Truth value of 'Coin lands Heads'	$(1 − 0.5)^2 = 0.25$	$(0 − 0.5)^2 = 0.25$
Truth value of 'Coin lands Tails'	$(0 − 0.5)^2 = 0.25$	$(1 − 0.5)^2 = 0.25$

To find out the total distance from truth of the set of estimates at each world, we *sum* the distance from truth of each estimate:

	Total distance from truth in HEADS	Total distance from truth in TAILS
Eulalie's assignments	$2500 + 2500 + 0 + 0.25 + 0.25$ $= 5000.5$	$2500 + 2500 + 0 + 0.25 + 0.25$ $= 5000.5$
Rhoda's assignments	$5625 + 625 + 0 + 0.25 + 0.25$ $= 6250.5$	$625 + 5625 + 0 + 0.25 + 0.25$ $= 6250.5$

This illustrates Theorem 1's conclusion that Rhoda's assignments are further from the truth in both the HEADS state and the TAILS state than Eulalie's are. Rhoda's assignment is dominated in the sense that Pettigrew notes. It is dominated because although Rhoda's assignment is closer than Eulalie's by 25 utils to the value of 'Bet on Heads' in TAILS and 'Bet on Tails' in HEADS (true value: 0; Rhoda: 25; Eulalie: 50) and further than Eulalie's by 25 utils from the value of 'Bet on Heads' in HEADS and 'Bet on Tails' in TAILS (true value: 100; Rhoda: 25; Eulalie: 50), the distance measure implies that it is worse to be further from the truth, the further you already are. Being 75 rather than 50 utils from the true value is worse than being 50 rather than 25 utils from the true value.

Again, the conclusion is not that Rhoda will select a dominated act: she will select 'Don't Bet,' which is not dominated. It is also not that any particular estimate will be dominated by Eulalie's estimate in every state: while Rhoda's

estimate of 'Bet on Heads' does worse in the HEADS state, it does better in the TAILS state. Incidentally, it is also worth noting that although in this particular example, there is a dominating assignment that has the same probabilities as Rhoda's assignment, this won't always hold. In particular, if all of the acts are comonotonic and if Rhoda and Eulalie agree about the probabilities, then Rhoda's assignment will be closer to the truth in the worst state than Eulalie's is. So although there will be a dominating assignment (a fact which we know to be true by Theorem 1), it won't preserve Rhoda's *beliefs*. Others have raised and responded to related concerns about dominance arguments (see Easwaran and Fitelson 2012; Pettigrew 2013), so I will leave it to the reader to decide whether this is a problem for the argument.

This consideration aside, should Rhoda care about the kind of dominance she is subject to? Let us start by observing that it is not a kind of dominance that captures what she *does* care about. Why? For her, not all error is created equal. Recall that when Rhoda evaluates a bet on TAILS, the possibility of ending up with 0 utils is weighted more heavily than the possibility of ending up with 100 utils. This is to say, when she evaluates an act, she considers being 25 rather than 50 utils away from the true value when that value is 0 to be *better* than being 50 rather than 75 away from the true value when that value is 100. Is it not just distance from the true value that matters, but how good that value is: roughly speaking, it is more important to correctly estimate worse values than better ones. Rhoda will object to the method Theorem 1 employs for determining the total distance from the true value in a particular state: she will object to simply summing the errors of each estimate. No score that treats the error in worse states within an act the same as the error in better states within that act will capture what she cares about. Rhoda does not care about all errors equally.

We have just established that Rhoda doesn't care about the kind of dominance mentioned in the argument. *Should* she? I claim that she need not. Decision theory is a theory of instrumental rationality, and what we want from decision theory is a *ranking* of acts (and perhaps an assessment of their relative instrumental value) rather than an *estimate* of the actual value of an act. The instrumental value of an act – its EU or REU – is not an estimate of how valuable that act will in fact turn out to be, but an assessment of its value in potentially realizing some of one's aims, relative to other possible acts. Being close to an act's actual value (the success condition for an estimate) is only helpful for decision-making insofar as the ranking of act-estimates corresponds to which acts turn out to be best at potentially realizing some of one's aims. But the disagreement is precisely about which acts turn out to be best in this way. Therefore, to assume a distance measure that treats all estimation error the same is to prejudge the answer to this question. However, Pettigrew's argument does reveal a cost of REU theory: if we accept REU theory, we cannot see the agent as trying to accurately estimate the values of acts, where accuracy is defined without reference to a quantity's relative position.

4. Briggs's first challenge: choice over time

Like Pettigrew, Briggs claims that Rhoda will forgo something that dominates what she actually does – but where Pettigrew claimed that she would forgo a dominant estimate of act values, Briggs claims that she will forgo a dominant *strategy*, where a strategy is a series of acts over time.

To understand Briggs's argument, we must recall three methods of choice over time: naïve choice, sophisticated choice, and resolute choice. Choosing naively means choosing at each time the act that is part of your preferred strategy going forward, with no regard to what your past or future selves *do* or *prefer*. Choosing sophisticatedly means choosing at each time as if your future act is determined by your future selves' preferences, using backward induction to eliminate options that you know your future self will not choose.[5] Choosing resolutely means picking a strategy and sticking to it – the strategy could be determined by your earliest preferences, or it could represent some compromise between the preferences of all your time-slices.[6]

Keeping these methods in mind, let us examine the dialectic. Briggs interprets me as having challenged both premises in the following argument:

(1) According to REU theory, it is sometimes permissible for an agent to choose a strategy in an extended choice situation, where that strategy is dominated by some option available to the agent.
(2) Choosing a strategy that is dominated by some available option in an extended choice situation is irrational.

I've altered her statement of the first premise to make clear that the dominance must be by some strategy that is available to the agent. For example, buying baseball tickets with a dollar-off coupon dominates buying them without the coupon – it is preferable whether I sit in the cheap seats or the expensive seats – but if the option of using the coupon is not available to me, then I do nothing wrong by paying full price. (I will later consider a slightly different interpretation of Briggs's argument, one that recasts the first premise in a different way.)

I challenged the first premise, Briggs points out, by claiming that only naïve choosers pick strategies that are dominated by other strategies available to them, and that both resolute and sophisticated choosers do not do so: resolute choosers because although all of the relevant options are available to them, they do not select a dominated one; and sophisticated choosers because the only options that dominate their selected option are not available to them. Briggs objects to my challenge by arguing that the strategy that dominates the sophisticated chooser's selection is in fact available to her. She leaves unchallenged the claim that resolute choosers do not pick dominated options. So, according to Briggs, only proponents of resolute choice can deny the first premise – proponents of sophisticated choice must accept it.

According to Briggs, I also challenged the second premise by claiming that choosing a dominated strategy in an extended choice situation can be rational, on the grounds that we can explain and vindicate why an agent would make these choices. And her response is to argue that choosing a dominant strategy *constitutes* irrationality. I prefer to put things slightly differently. I agree with premise (2) that choosing a strategy dominated by an available option in an extended choice situation is irrational; therefore, I agree with Briggs that naïve Rhoda is irrational, and that if sophisticated Rhoda really does choose an option dominated by an available option, then sophisticated Rhoda is also irrational. However, I think that having preferences of a certain form – preferences that necessitate sophisticated or resolute choice – is rational. That was the intended force of my arguments that Briggs mentions in her Section 3. (These preferences *would* lead one to choose a dominated strategy *if* one chose naively, so the mis-interpretation is understandable.) Still, we can recast Briggs's argument against choosing dominated options as instead against having these preferences: she can argue that *having preferences that necessitate sophisticated or resolute choice* constitutes practical irrationality.

Let us begin with the question of whether sophisticated Rhoda indeed chooses a strategy that is dominated by a strategy available to her: Briggs claims she does, I claim she does not. Our locus of disagreement is the choice problem that Briggs labels Problem 2, in which Rhoda faces a choice between two acts, L_1 and L_2 (from Allais [1953] 1979). Rhoda prefers L_1 to L_2. But before making her choice, she first must choose whether to accept more information about the true state of the world along with a sweetener – a sweetener which is not sweet enough to make sweetened L_2 preferable to unsweetened L_1. The information will reveal which of two events obtains, and for both events, Rhoda will prefer the sub-act of sweetened L_2 under that event to that of unsweetened L_1 under that event. Since Rhoda is a sophisticated chooser, she knows she will choose according to her preferences after obtaining the information, and will therefore end up with an act logically equivalent to sweetened L_2 if she does receive the information. As a sophisticated chooser, she therefore sees her initial choice as between receiving the information and ending up with sweetened L_2, or forgoing the information and ending up with unsweetened L_1. Since she prefers L_1, she will forgo the information. However, there was a series of choices she could have made – namely, accepting the information and the sweetener, and choosing (the logical equivalent of) L_1 – which would have resulted in sweetened L_1. Since sweetened L_1 strictly dominates what she ends up with (unsweetened L_1), she chooses a dominated option.
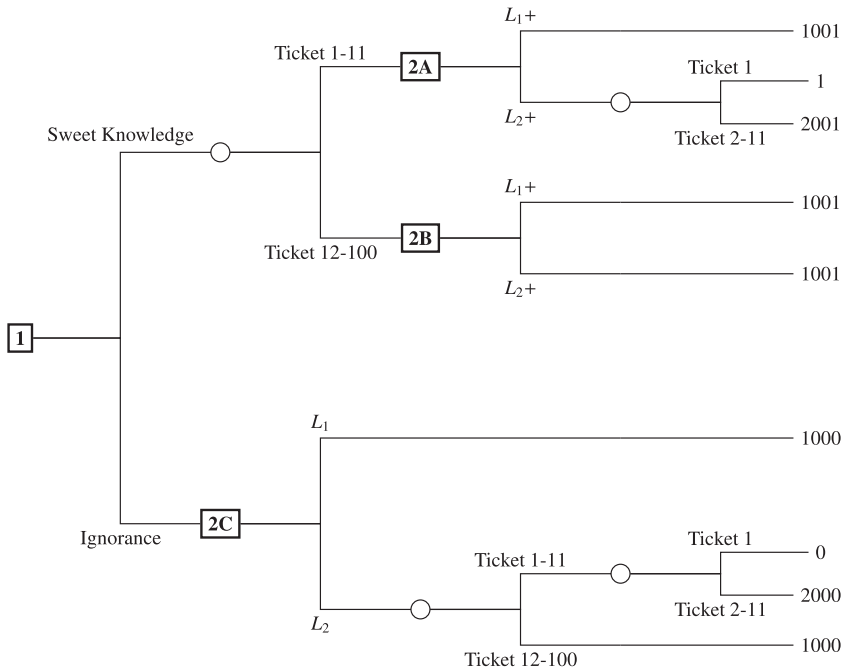
Figure 1. Problem 2 (Briggs, 2015).

What I said in *Risk and Rationality* is that sweetened L_1 is not available to sophisticated Rhoda. Sophisticated choice only makes sense if we operate with a restricted notion of an agent's options: an agent's options are her options at a particular time. And there is no time at which the acts that constitute choosing sweetened L_1 are available to Rhoda. However, Briggs challenges this point (note that 'L_{1+}' stands for sweetened L_1):

> Nonetheless, there is a sequence of actions, each of which Rhoda can perform, that will guarantee she ends up with L_{1+}. (All she has to do is choose sweet knowledge at [1], and choose L_{1+} at [2A], should she get there.) In other words, there is a strategy available to Rhoda that results in her ending up with L_{1+}. Buchak holds that on the view that motivates sophisticated choice, we are not entitled to evaluate entire strategies, since there is no time at which strategies are the object of choice. But the objector should press the point that *a strategy can be available, even if there is no one time at which it is available*. (p. [8], italics mine)

The argument says, look, Rhoda can have sweetened L_1 – it's available to her! All she has to do is select, at each time, the acts that constitute it. There is a difference between saying that an option is available to Rhoda and saying that there is a single time at which it is available to her, and sweetened L_1 is available to Rhoda in this former sense.

I want to press my point that the reasons you might have for accepting sophisticated choice are also reasons to think that availability is just availability at a time. Recall that there are two views one might have of what an agent

is. One is the time-slice view, which sees a rational agent as a series of time-slices that each forms preferences and makes decisions in isolation of what other time-slices prefer. On this view, one's choices are fully determined by what each of one's time-slices prefers, and furthermore, one stands in a similar relationship to one's other time-slices as one does to other people. I can influence others' actions indirectly – by incentivizing them in various ways (changing their payoffs) or by limiting what they can do (changing their options) or by convincing them to care about what I care about (indirectly changing their preferences) – but I cannot choose for them. So too can I influence the actions of my future self only indirectly: I can set up rewards and penalties, I can remove future options, I can cultivate a desire to honor past wishes, but I cannot now choose what to do later in a way that will bind me to that choice. An individual at a time has influence over the actions of her time-slices taken together to the same extent and in the same way (though perhaps by a different mechanism) that an individual has influence over the actions of a group of which she is a part. Carrying out a strategy is the result of several 'selves' trying to get what they most prefer, given the behavior they can expect from the other selves. As I put it in *Risk and Rationality*, there is no agent to whom temporally extended acts are available: there is only a temporal extension of time-slices to whom acts-at-a-time are available.[7]

The time-slice view provides a natural home for sophisticated choice: sophisticated choice enjoins me to see my future selves' choices as fixed, and to treat them as I would treat the choices of other individuals. Sophisticated choice looks less natural on a holistic view of agency. On this view, the primary locus of action is the temporally extended agent. And the primary entities over which such an agent has control are temporally extended acts. Given these assumptions, sophisticated choice is a puzzling strategy. Why should I treat my future actions as determined, while deliberating about my present actions? I should instead deliberate about strategies over time, and choose once for all. This is why resolute choice finds a natural home in this picture.

I won't adjudicate the debate between the time-slice view of agency, with sophisticated choice, and the holistic view of agency, with resolute choice. What I continue to claim is that we don't have a reason to accept both sophisticated choice *and* the claim that the dominating option is available. The proponent of EU theory will have to supply a picture on which this combination of commitments makes sense. Indeed, what Briggs highlights is that if we accept sophisticated choice, then the debate between REU theory and EU theory hinges on what counts as an option.

Before moving on to a different version of Briggs's argument, let us briefly return to her second point. Recall that she reads me as denying that choosing dominated strategies is irrational, on the grounds that we can explain and vindicate these choices. Her point as stated is that choosing dominated strategies is *constitutive* of irrationality, because doing so is both 'internally' irrational – it

amounts to choosing something that is worse by Rhoda's own lights – and 'externally' irrational – it frustrates the aim of choosing the act with the best consequences. I agree with this point, and this is a good way of explaining what's wrong with choosing dominated strategies. However, recall that what I actually deny is that having preferences that in some situations require sophisticated or resolute choice is irrational. So, let us look at an amended version of Briggs's argument for the claim that having these preferences is constitutive of irrationality. This argument would say that having such preferences is irrational because having them amounts to preferring something that is worse by Rhoda's own lights, or that having such preferences is irrational because having them frustrates the aim of preferring the act with the best consequences. But these claims only hold if preferences over strategies are determined *both* by the consequences of those strategies *and* by what one would do if one were naïve – and there is no reason to think this second clause holds. Therefore, the amended version of the argument has not yet revealed that Rhoda is irrational.

We can further explore the question of whether anything is wrong with sophisticated Rhoda by considering the interpersonal analogue of Problem 2. Assume that choices are made by two different agents with different knowledge states, rather than by two time-slices of the same agent before and after receiving information. These agents, $Rhoda_1$ (who doesn't know whether E) and $Rhoda_2$ (who does know whether E), have Rhoda's preferences about outcomes and lotteries. $Rhoda_1$ will choose Sweet Knowledge or Ignorance at the initial decision node, and if she chooses Ignorance she will also choose at $2C$; and $Rhoda_2$ will choose at $2A$ (we will ignore the choice at $2B$, as Briggs does, because both Rhoda's are indifferent between the two options at that node). We can represent this choice using a game matrix, where $Rhoda_1$ is the 'row' player, $Rhoda_2$ is the 'column' player, and the payoffs to each player are the lotteries that they will receive.

	2A: L_{1+}	2A: L_{2+}
Sweet knowledge	L_{1+}, (L_{1+E} or $L_{1+(-E)}$)	L_{2+}, ($L_{2+}E$ or $L_{2+(-E)}$)
Ignorance; 2C: L_1	L_1, ($L_1\bar{E}$ or $L_{1(-E)}$)	L_1, ($L_1\bar{E}$ or $L_{1(-E)}$)
Ignorance; 2C: L_2	L_2, (L_{2E} or $L_{2(-E)}$)	L_2, (L_2E or $L_{2(-E)}$)

The two highlighted acts constitute the *sub-game perfect equilibrium*, the result of backward induction: $Rhoda_2$ will choose L_{2+} at $2A$; so the payoff of Sweet Knowledge for $Rhoda_1$ will be L_{2+}; so $Rhoda_1$ chooses Ignorance and L_1.

What of the original claim that Rhoda employs a strategy that is diachronically dominated? Translating into the interpersonal framework, it is the claim that $Rhoda_1$ and $Rhoda_2$ choose an act pair that is *strictly sub-optimal*: there is an alternative act pair whose outcome is strictly preferred by both $Rhoda_1$ and $Rhoda_2$ to the selected outcome: namely, (Sweet Knowledge, L_{1+}).[8] Thus, the notion of diachronic dominance in the intrapersonal framework translates into the notion of sub-optimality in the intrapersonal framework; and the strategy

selected by sophisticated choice translates into the sub-game perfect equilibrium. The interpersonal framework also has a concept of *dominance*: an individual player's act is dominated if there is some alternative act which leads to a preferred outcome given any choice made by the other player. Dominance in this sense is a property of an individual's act, not of an act pair. Translation in hand, we can restate my claim that the (intrapersonal) strategy that selects L_1 is not dominated by any other available option: for the acts chosen by each Rhoda-at-a-time, there is no alternative act that is preferable no matter what the other Rhoda's-at-a-time choose.

If we take seriously the correspondence between the intrapersonal problem and the interpersonal problem – as it seems the time-slice view must – then what is bad about Rhoda's strategy is the bad of an act pair's being sub-optimal. Given this, we can construct the following version of Briggs's argument:[9]

> (1') According to REU theory, it is permissible for the sophisticated chooser to choose a strategy (a pair of acts-over-time) in an extended choice situation, where that strategy is intrapersonally sub-optimal.
> (2') Choosing a strategy that intrapersonally sub-optimal is irrational for the sophisticated chooser.

This argument correctly isolates what is available to an agent on the time-slice picture, so premise (1') is true. But is premise (2') true? We must consider what norms an agent is subject to, on the time-slice picture.

Let us examine the corresponding interpersonal question: is it *irrational* for two players to select a sub-optimal act pair? The standard view is that choosing a sub-optimal act pair is not irrational because there is no one for whom it is irrational. Being in a problem in which the equilibria are sub-optimal is unfortunate, but the result is not a failure of rationality on the part of any actors, since both are doing the best they can given what the other chooses. Analogously, if rational norms concern the agent at a time,[10] then choosing a sub-optimal act pair is not irrational because there is no *one time-slice* for whom it is irrational. It is unfortunate that the agent over time suffers, but each time-slice is acting rationally, and one is only answerable for what one does at a time.

There is a minority view according to which selecting a sub-optimal strategy pair in the interpersonal situation is in fact irrational. However, those who accept this view typically accept the analogue of resolute choice *for groups*, that separate individuals can jointly decide on an act pair and each do their part.[11] Thus, on this view, selecting a sub-optimal strategy pair is irrational because *both* players are acting irrationally. Analogously, a proponent of the time-slice picture could hold that both time-slices ought to choose resolutely – resolute Rhoda would be irrational to pick a sub-optimal act pair, but she won't do so. If we accept the time-slice picture, then either the time-slices ought to choose resolutely, in which case Rhoda won't end up with the sub-optimal strategy, or

they ought not to, in which case Rhoda won't be violating any norms in ending up with the sub-optimal strategy.

What Briggs and I agree on is that decision theorists face a choice. If a theorist wants to accept sophisticated choice *and* the idea that diachronic dominance by a strategy only available to the agent-over-time is irrational, she must reject REU theory. I do not think this is worrisome for REU theory, because sophisticated choice and this idea do not fit together well. In any case, the REU theorist still has many theoretical options open to her: she can be a proponent of sophisticated choice and hold that this kind of diachronic dominance is not bad. Or she can be a proponent of resolute choice, holding that Rhoda will not be subject to diachronic dominance. Insofar as one is convinced that the intrapersonal problem is not like the interpersonal problem – that there is a unity among time-slices that does not hold among people – this latter option may be the best.

5. Briggs's second challenge: simplifying choices

Briggs's other criticism of REU theory is that for Rhoda, sub-acts lack stable utility values: the contribution of a particular sub-act to an act's value can depend on what happens in the rest of the act. Put another way, the certainty equivalent of each sub-act – the sure-thing amount we can substitute for that sub-act while preserving the value of the act – can depend on what happens in the rest of the act. This follows from the fact that preferences about 'sure-thing' sub-acts are fixed – if an agent prefers a sure-thing *x* to a sure-thing *y*, she will always rather include *x* as a sub-act than include *y* as a sub-act – but preferences about 'gamble' sub-acts depend on which other sub-acts they are paired with. Since there is a stable ordering of sure-thing sub-acts, but not of gamble sub-acts, certainty equivalents for gamble sub-acts cannot stay fixed.

Is the fact that sub-acts lack stable values bad? One possible way to argue in the affirmative, though Briggs herself does not make this argument, is to claim that not being able to assign stable values to sub-acts is in itself bad. One could claim that the individual who does not stably value sub-acts has inconsistent values. But the REU response is to point out that since the instrumental value of an act depends on its global properties, the contribution of any sub-act will depend on which global properties it helps instantiate, which depends on which sub-acts it is paired with. In short, REU theory holds that while there is a fixed fact of the matter about the value of an outcome, and about the instrumental value of an entire act, there is no fixed fact of the matter about the instrumental value of a sub-act, apart from the act in which it is embedded. As I said in *Risk and Rationality*, what instrumental value is like generally – including whether sub-acts with the same instrumental value are interchangeable – is not a pre-theoretical question. So this argument would straightforwardly beg the question against REU-maximization.

The argument Briggs actually makes is that because sub-acts lack fixed values, REU-maximizers lose a key strategy for simplifying 'grand world' decisions by recasting them as 'small world' decisions. This isn't necessarily to say that Rhoda is irrational, but she will have a more difficult time actually making decisions. REU theory is unwieldy and impractical.

Briggs's example involves a decision between buying a frogurt and not buying a frogurt, where one does not know whether the frogurt is cursed, whether it comes with a free topping, and if so whether the topping contains E212. The decision matrix is as follows:

	Cursed Topping E212	Cursed Topping No E212	Cursed No topping	Not cursed Topping E212	Not cursed Topping No E212	Not cursed No topping
Buy	−50	−30	−45	10	30	15
Don't buy	0	0	0	0	0	0

Briggs points out that Eulalie can easily simplify the problem by 'collapsing' the values under 'cursed' and 'not cursed' into the expected values in those events. However, claims Briggs, while Rhoda can impute values to these events, the value of the event in which the frogurt is cursed 'will not be a function of the probabilities and utilities of states in which the frogurt is cursed,' but instead 'will also depend on both the probabilities and utilities of the states in which the frogurt is not cursed'([13]); and similarly for the value of the event in which the frogurt is not cursed. More generally, while Eulalie will be able to coarse-grain her problem by collapsing some states into a single event with a fixed utility value, Rhoda cannot coarse-grain her decision problems in this way.

As it turns out, only a weakened version of Briggs's claim is true. To see this, let us first consider an abstract example in which there are four states. Let $u(S)$ stand for the utility of ACT in state S, and let $u(A) \leq u(B) \leq u(C) \leq u(D)$:

	A	B	C	D
ACT	$u(A)$	$u(B)$	$u(C)$	$u(D)$

The REU of ACT is:

$$REU(ACT) = u(A) + r(p(B \vee C \vee D))(u(B) - u(A)) + r(p(C \vee D))(u(C) - u(B))$$
$$+ r(p(D))(u(D) - u(C))$$

Which can be rewritten as:

$$REU(ACT) = u(A)(1 - r(p(\neg A))) + u(B)(r(p(A)) - r(p(\neg(A \vee B))) + u(C)\,(r(p(C \vee D))$$
$$- r(p(D))) + u(D)(r(p(D)))$$

Notice that the terms on the first line are a function only of the probabilities and utilities of states A and B, and the terms on the second line are a function only of the probabilities and utilities of the states C and D. So we can, indeed, divide the REU of ACT into two parts, an 'A or B' part and a 'C or D' part. But notice that we cannot divide the REU of ACT into an 'A or C' part and a 'B or D' part: we

cannot separate terms that involve only probabilities and utilities of *A* and *C* from those that involve only probabilities and utilities of *B* and *D*.

We can separate ACT in one way and not the other because the successful separation involves a division into events that are *contiguous*. Define a contiguous event (set of states) for a given act as follows: *E* is contiguous iff *E* contains all states *S* such that $\min_{T \in E} u(T) < u(S) < \max_{T \in E} u(T)$. And we can note that if the states of an act are divided into two contiguous events, then the value contribution of each event to the act will depend only on the probabilities and utilities of the states in the event.[12] For the sub-act over each contiguous event, its value will be stable and will depend only on the probabilities and utilities of the states *within* that event.

It is worth noting two caveats in support of Briggs's objection, as well as an additional fact in support of Rhoda's ability to simplify decisions. First, even with a contiguous division, there will be no natural way to divide the value of each event into a 'probability' and 'utility' component, as there is for EU theory. Second, the claim that the value of each event only depends on the probabilities of the states in that event only holds true for a coarse-graining into *more than* two events if the probabilities of all the coarse events are fixed. However, in support of REU theory, the sub-acts of any contiguous division (including those into more than two events) do have stable certainty equivalents. We know this because preferences about sub-acts are stable for acts that are comonotonic, and substituting a certainty equivalent for a sub-act on a contiguous event preserves comonotonicity (because the certainty equivalent will always be between the minimum and maximum value of the sub-act).

To return to the frogurt problem: in this problem, all of the cursed states have lower utility than all of the non-cursed states. If we coarse-grain the problem into the event in which the frogurt is cursed and that in which the frogurt is not cursed, we can separately determine the value of the act in the 'cursed' states and in the 'non-cursed' states – where 'value' here means the value contributed to the REU by these states – and these two values will depend only on the probabilities and utilities within those states. This also means that the value of act in the cursed states will be stable across changes to probabilities and utilities in the non-cursed states. For example, here are some possible changes: the value of a not-cursed frogurt with no topping might change; or the agent might come to realize that this value depends on a further unknown (whether it is possible to swirl two flavors or whether one has to choose only one); or the probability that a non-cursed frogurt comes with a topping might change, holding fixed the probability that the frogurt is cursed. As long as these changes do not disrupt the contiguity of the division into cursed and non-cursed states or the probability of these two events, then these changes will only alter the value of the non-cursed event – the value of the cursed event will stay the same. So we *can* simplify the decision problem by separately considering how good the cursed event is and how good the non-cursed event is.

However, if we coarse-grain the problem into the event in which the frogurt comes with a free topping and the event in which is does not, we cannot separately consider how good each event is. So we cannot simplify the decision problem by separately considering how good the free-topping event is and how good the no-free-topping event is.

To sum up: for some coarse-grainings of the problem – coarse-grainings into contiguous sets – Rhoda can assign stable values to coarse-grained events, and these values will depend only on what happens within those events. For other coarse-grainings, this will not be possible. Coarse-graining will be an effective way of simplifying a decision problem only if it involves a division into two contiguous events (or into two or more contiguous events where probabilities of events are fixed or the value of an is allowed to depend on the probability of the other events). But this may not be very bad for REU theory. Indeed, it explains why some ways of thinking about a decision problem strike us as productive and others as unproductive. In the frogurt problem, the consideration that ought to loom largest in our minds is whether or not the frogurt is cursed, not whether or not the frogurt comes with a free topping. It is true that Rhoda has a more difficult time simplifying her decision problems than Eulalie does: only certain simplifications will do. But perhaps this just means that Rhoda is appropriately discerning in how she thinks about decision problems.

6. Conclusion

Pettigrew and Briggs both raise important challenges for REU theory. Pettigrew's first challenge tried to show that REU theory is superfluous, since Rhoda can in fact be recast as an EU-maximizer. However, his way of incorporating her risk-attitudes into the utility value assignment did not generate a stable ordering of outcomes. Pettigrew's second challenge tried to show that Rhoda's utility and probability assignments are dominated in the sense that there is an alternative set of assignments that are closer to the truth in every world. However, there are reasons to think that she will reject the measure of closeness according to which this is true. Furthermore, one could hold that what decision theory is concerned with is not estimating truth values, but with comparing acts – on this view, dominance according to this measure need not be irrational. Briggs's first challenge concerned a different kind of dominance: diachronic dominance. However, depending on our picture of agency and norms, we must either hold that Rhoda will not choose the dominated option or that choosing the "dominated" option is not irrational. Still, Briggs brings out a cost of accepting REU theory: one cannot also accept both sophisticated choice and the claim that norms of rationality apply non-derivatively to the agent over time. Finally, Briggs's second challenge pointed out that Rhoda lacks a way to simplify her decisions, a way that Eulalie possesses. Importantly, this claim is only true in a limited sense: Rhoda can simplify her decisions in some ways but not others,

and this distinction plausibly tracks ways of simplifying decisions that seem natural to us.

These challenges help us see more clearly what REU theory is committed to, as well as some of the costs of rejecting the EU thesis. For my part, I continue to hold that it is rational to take risk into account in a way that cannot be subsumed under EU-maximization, and that rational preferences involve three internal attitudes: utilities, credences, and risk-attitudes.

Notes

1. Throughout, except where it is an explicit topic of discussion, I remain neutral on whether the relevant attitudes exist independently of preferences or are constructed from them.
2. More precisely, the REU of an act $f = \{E_1, x_1; \ldots; E_n, x_n\}$, where $u(x_1) \leq \ldots \leq u(x_n)$, is
$u(x_1) + r(\sum_{i=2}^{n} p(E_i))(u(x_2) - u(x_1)) + r(\sum_{i=3}^{n} p(E_i))(u(x_3) - u(x_2)) + \ldots + r(p(E_n))(u(x_n) - u(x_{n-1}))$
. In other words, $REU(f) = \sum_{j=1}^{n} r(\sum_{i=j}^{n} p(E_i))(u(x_j) - u(x_{j-1}))$ This is equivalent
to $REU(f) = \sum_{j=1}^{n} (r(\sum_{i=j}^{n} p(E_i)) - r(\sum_{i=j+1}^{n} p(E_i)))u(x_j)$. We can see REU as a generalization of Quiggin's (1982) anticipated utility to subjective rather than objective probabilities; or as a restriction of Gilboa's (1987) (and, recast, Schmeidler's 1989) Choquet EU to weightings that are a function of probabilistic credences; or as a similar generalization of objective cumulative prospect theory or restriction of subjective cumulative prospect theory (see Kahneman and Tversky 1979; Tversky and Kahneman 1992), with the additional stipulation that gains and losses are not treated differently. See (Buchak 2013, Ch. 1 and 2) for a discussion of the relationship of REU theory to other non-EU theories.
3. $REU = u(x_1) + r(p(E_2) + p(E_4))(u(x_2) - u(x_1)) + r(p(E_4))(u(x_4) - u(x_2)) = 3 + (0.3 + 0.4)^2(5 - 3) + 0.4^2(6 - 5) = 4.14$.
4. More precisely, since the only real facts about utility functions are difference ratios, the utility difference between two pairs of outcomes is the same if these two pairs constitute equal tradeoffs in acts with the same structure.
5. See Strotz (1955), Hammond (1988), and McClennen (1990).
6. See Strotz (1955), McClennen (1990, 1997), and Gauthier (1997).
7. This coheres with a view recently put forth by Hedden (2015), a proponent of the time-slice picture, of what an option is: 'your options are all and only the *decisions* you are presently able to make' (95).
8. The usual game-theoretic terminology is 'strategy' for an individual's choice and 'strategy pair' for the pair of strategies chosen by the two agents. However, since I follow Briggs in using 'strategy' to refer to the pair of acts chosen by the agent at the earlier time and at the later time, the interpersonal analogue of which is the pair of acts chosen by the two agents, I will instead use 'act' and 'act pair.'
9. This isn't quite the right way to interpret Briggs herself, because the argument of mine that Briggs cites against premise (1) would actually concern premise (2'). Nonetheless, this argument is at least strongly suggested by her discussion.
10. A recent proponent of this view is Hedden (2015), who argues for 'time-slice rationality,' according to which 'how you rationally ought to be at a time directly depends only on your mental states at that time, not on how you (or time-slices psychologically continuous with you) were in the past or will be in the future' (7). Moss (2015) argues for a similar view in epistemology, which she dubs 'time-slice epistemology.'

11. See, for example, McClennen (1997), who explicitly makes the connection between intrapersonal decisions and the interpersonal decisions, and argues that resolute choice can resolve problems within both. Gauthier (1997) denies that the two types of decisions are analogous: other people's concerns always ought to be taken seriously but my past concerns can cease to matter when I no longer embrace them. Furthermore, he holds that sophisticated choice rather than resolute choice ought to be employed in typical intrapersonal situations; however, when sophisticated choice would be sub-optimal, resolute choice ought to be employed. Thus, insofar as he holds that landing on a sub-optimal strategy pair is irrational, he holds that resolute choice is appropriate.

12. Notice that if there are two states with the same utility, then a contiguous event with this utility as the minimum or maximum value can include one state, the other state, or both. This won't pose the same kind of problem as it does for the comonotonic individuation (discussed in Section 2), since we are not trying to come up with a utility contribution for an event independent of the ordering under consideration (i.e. independent of the particular coarse-graining we consider).

Notes on contributor

Lara Buchak is Associate Professor of Philosophy at UC Berkeley. She is the author of *Risk and Rationality* (Oxford, 2013). She has published articles in decision theory, epistemology, philosophy of religion, and philosophy of action.

References

Allais, Maurice. ([1953] 1979). "The Foundations of a Positive Theory of Choice Involving Risk and a Criticism of the Postulates and Axioms of the American School." In *Expected Utility Hypothesis and the Allais Paradox*, edited by Maurice Allais and Ole Hagen (1979), 27–145. Dordrecht: Reidel. English Translation of "Fondements d'une Théorie Positive des Choix Comportant un Risque et Critique des Postulats et Axiomes de L'Ecole Americaine," Econometrie (1953): 257–332.

Briggs, Rachael. 2015. "Costs of Abandoning the Sure-thing Principle." *Canadian Journal of Philosophy* (forthcoming).

Buchak, Lara. 2013. *Risk and Rationality*. Oxford: Oxford University Press.

Easwaran, Kenny, and Branden Fitelson. 2012. "An 'Evidentialist' Worry About Joyce's Argument for Probabilism." *Dialectica* 66 (3): 425–433.

Gauthier, David. 1997. "Resolute Choice and Rational Deliberation: A Critique and a Defense." *Noûs* 31 (1): 1–25.

Gilboa, Itzhak. 1987. "Expected Utility with Purely Subjective Non-additive Probabilities." *Journal of Mathematical Economics* 16: 65–88.

Hammond, Peter. 1988. "Consequentialist Foundations for Expected Utility." *Theory and Decision* 25: 25–78.

Hedden, Brian. 2015. *Reasons Without Persons: Rationality, Identity, and Time*. Oxford: Oxford University Press.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47: 263–291.

McClennen, Edward F. 1990. *Rationality and Dynamic Choice: Foundational Explorations*. Cambridge: Cambridge University Press.

McClennen, Edward F. 1997. "Pragmatic Rationality and Rules." *Philosophy and Public Affairs* 26 (3): 210–258.

Moss, Sarah. 2015. "Credal Dilemmas." *Noûs* 49 (4): 665–683.

Pettigrew, Richard. 2013. "Accuracy and Evidence." *Dialectica* 67 (4): 579–596.

Pettigrew, Richard. 2015. "Risk, Rationality, and Expected Utility Theory." *Canadian Journal of Philosophy* (forthcoming).

Quiggin, John. 1982. "A theory of anticipated utility." *Journal of Economic Behavior and Organization* 3: 323–343.

Schmeidler, David. 1989. "Subjective Probability and Expected Utility Without Additivity." *Econometrica* 57 (3): 571–587.

Strotz, R. H. 1955. "Myopia and Inconsistency in Dynamic Utility Maximization." *The Review of Economic Studies* 23 (3): 165–180.

Tversky, Amos, and Daniel Kahneman. 1992. "Advances in Prospect Theory: Cumulative Representation of Uncertainty." *Journal of Risk and Uncertainty* 5: 297–323.