CANADIAN JOURNAL OF PHILOSOPHY, 2016
VOL. 45, NO. 5–6, 798–826
http://dx.doi.org/10.1080/00455091.2015.1119610

Routledge
Taylor & Francis Group

Risk, rationality and expected utility theory

Richard Pettigrew

Department of Philosophy, University of Bristol, Bristol, UK

ABSTRACT

There are decision problems where the preferences that seem rational to many people cannot be accommodated within orthodox decision theory in the natural way. In response, a number of alternatives to the orthodoxy have been proposed. In this paper, I offer an argument against those alternatives and in favour of the orthodoxy. I focus on preferences that seem to encode sensitivity to risk. And I focus on the alternative to the orthodoxy proposed by Lara Buchak's risk-weighted expected utility theory. I will show that the orthodoxy can be made to accommodate all of the preferences that Buchak's theory can accommodate.

ARTICLE HISTORY Received 3 October 2015; Accepted 5 November 2015

KEYWORDS Risk; decision theory; expected utility theory; rational choice theory; accuracy; scoring rules

1. Introduction

Philosophers, psychologists and economists have known for a long time that there are a number of decision problems for which the preferences over the available options that seem rational to many people cannot be accommodated within orthodox decision theory in the natural way. In response to this observation, a number of alternatives to the orthodoxy have been proposed (Allais 1953; Quiggin 1982; Schmeidler 1989; Buchak 2013). In this paper, I offer an argument against those alternatives and in favour of the orthodoxy. This argument is very general: it is effective against any deviation from the orthodoxy. As a result, we need some account of the preferences that seem rational and yet which this

CONTACT Richard Pettigrew richard.pettigrew@bris.ac.uk

This paper began life as my contribution to the Author Meets Critics session on Lara Buchak's *Risk and Rationality* at the Pacific APA in San Diego in April 2014. I received extremely helpful feedback from Lara at that point and then later, when I came to write up the paper for publication. I would also like to thank Jason Konek, Rachael Briggs, Ralph Wedgwood, Greg Wheeler, and Ben Levinstein for further comments on this paper. The work on this paper was supported by the European Research Council Seventh Framework Program (FP7/2007-2013) Starting Researcher Grant (308961-EUT), Epistemic Utility Theory: Foundations and Applications.

orthodoxy cannot accommodate naturally: we need an error theory for our intuition that they are rational, or a way of making the orthodoxy accommodate them. I will focus here on those preferences that seem to encode sensitivity to risk. And I will focus on the alternative to the orthodoxy proposed by Lara Buchak's *risk-weighted expected utility theory*, which is intended to accommodate these preferences (Buchak 2013). I will show that, in fact, the orthodoxy can be made to accommodate the preferences in question; and I will argue that this is in fact the correct way to accommodate them. Thus, the paper has two parts: the first is a general objection to any non-expected utility theory; the second is a specific account of how to accommodate within the orthodoxy the preferences that Buchak's theory permits.

2. The argument for orthodox expected utility theory

2.1. Decision problems and the framework of decision theory

Here's a decision problem. An agent is choosing between two options: on the first, which we will call *Safe*, she is guaranteed to receive £50; on the second, which we will call *Risky*, a fair coin will be flipped and she will receive £100 if the coin lands heads and £0 if the coin lands tails. There are three components of this decision problem. First: the states of the world, namely, the state in which the coin lands heads and the state in which the coin lands tails. Second: the outcomes, namely, £0, £50, £100. Third: the acts between which our agent is choosing, namely, *Safe* and *Risky*. In general, a decision problem consists of these three components:

- $\mathcal{S}$ is the set of *states* (or *possible worlds*). Thus, in our example, $\mathcal{S} = \{Heads, Tails\}$.

Degrees of belief or *credences* will be assigned to a finite algebra $\mathcal{F}$ of subsets of the set of states $\mathcal{S}$. These are propositions represented as sets of states or possible worlds.[1]

- $\mathcal{X}$ is the set of *outcomes*.

Thus, in our example, $\mathcal{S} = \{£0, £50, £100\}$. We will take outcomes to be entire descriptions of the world, rather than merely changes in the agent's wealth. Thus, £0 is really *Status quo* + £0; £50 is really *Status quo* + £50; and so on. But we will continue to denote them just by the change in the status quo that they represent. *Utilities* will be assigned to the elements of $\mathcal{X}$.

- $\mathcal{A}$ is the set of *acts*.

Thus, in our example, $\mathcal{A} = \{Safe, Risky\}$. We represent acts as finite-valued functions from $\mathcal{S}$ to $\mathcal{X}$. That is, they take ways the world might be and return the outcome of the act if the world were that way. Thus, for our purposes, we can represent an act *f* in the set of acts $\mathcal{A}$ using the following notation:

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

$f = \{E_1, x_1; \ldots; E_n, x_n\}$, where $x_1, \ldots, x_n$ are the values that the function f might take – that is, the possible outcomes of the acts – and, for each $i = 1, \ldots, n$, the proposition E_i says that f will have outcome x_i. Thus, if we represent propositions as sets of possible worlds, as we did above, $E_i = \{s \in \mathcal{S} : f(s) = x_i\}$. So E_i is the set of states of the world in which f has outcome x_i. Thus, in our example above, $Safe = \{Heads \vee Tails, £50\}$ and $Risky = \{Heads, £100; Tails, £0\}$. We assume that all such propositions E_i are in the algebra $\mathcal{F}$. For each outcome x in $\mathcal{X}$, there is an act in $\mathcal{A}$ – which we write $\overline{x}$ – that has outcome x regardless of the state of the world. That is, representing the act $\overline{x}$ as a function from states to outcomes, $\overline{x}(s) = x$ for all states s in $\mathcal{S}$. We call $\overline{x}$ the *constant act on x*. Let $\overline{\mathcal{X}} = \{\overline{x} : x \in \mathcal{X}\} \subseteq \mathcal{A}$ be the set of constant acts. They will prove particularly important in Section 5.2 below.

2.2. The business of decision theory

With this framework in place, we can state the business of decision theory. It is concerned with the relationship between two sorts of attitudes, which I will call *external attitudes* and *internal attitudes*.[2] The external attitudes are typically taken to be represented by the agent's preference ordering $\succeq$. $\succeq$ is an ordering on the set $\mathcal{A}$ of acts. If f and g are acts in $\mathcal{A}$, we say that $f \succeq g$ if the agent weakly prefers act f to act g. The internal attitudes, on the other hand, are typically taken to be given by the agent's credences and her utilities. As mentioned above, her credences are defined on propositions in a σ-algebra $\mathcal{F}$ on the set of states $\mathcal{S}$. They measure how strongly she believes those propositions. And her utilities are defined on the outcomes in $\mathcal{X}$. They measure how strongly she desires or values those outcomes.[3] If you are a constructivist about the internal attitudes, then you will take only the external attitudes to be real: you will then take the business of decision theory to be the representation of the external attitudes by treating the agent *as if* she has internal attitudes and *as if* she combines those attitudes in a particular way to give her external attitudes. If, on the other hand, you are a realist about the internal attitudes, then you will take both sorts of attitudes to be real: you will then say that a rational agent's internal and external attitudes ought to relate in a particular way; indeed, they ought to relate *as if* she obtains her external attitudes by combining her internal attitudes in a particular way. We will have more to say about the business of decision theory later (cf. Section 5.2 below).

2.3. The EU rule of combination

Expected utility theory posits only two types of internal attitudes: these are given by the agent's credences and utilities. Her credences are given by a credence function $c : \mathcal{F} \to [0, 1]$, which we assume to be a probability function on $\mathcal{F}$. Her utilities are given by a utility function $u : \mathcal{X} \to \mathbb{R}$. As with most decision theories, expected utility theory posits one type of external attitude, namely,

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

the agent's preference ordering. Expected utility theory then employs the following *rule of combination*, which states how her internal and external attitudes ought to relate:

EU Rule of Combination Suppose $f = \{E_1, x_1; \ldots; E_n, x_n\}$ is an act in $\mathcal{A}$ – that is, if E_i is true, the outcome of f is x_i. Then define

$$EU_{c,u}(f) := \sum_{i=1}^{n} c(E_i)u(x_i)$$

Then if the agent is rational, then

$$f \succeq g \iff EU_{c,u}(f) \geq EU_{c,u}(g)$$

That is, an agent's preferences ought to order acts by their subjective expected utility.

A number of decision theorists wish to deny the EU Rule of Combination. Buchak is amongst them, but there are other so-called non-expected utility theorists (Allais 1953; Quiggin 1982; Schmeidler 1989). I disagree with them about the rule of combination; however, as we will see in the second half of this paper, I agree with them about the rationality of the preferences that they try to capture by reformulating the rule of combination. In Section 4, I try to effect a reconciliation between these two positions – the correctness of the EU Rule of Combination, on the one hand, and the rationality of risk-sensitive preferences, on the other. In this part of the paper, I wish to argue that we ought to combine our internal attitudes in exactly the way that expected utility theory suggests. That is, I want to argue for the EU Rule of Combination.

How can we tell between different rules of combination? It is commonly assumed that representation theorems help us to do this, but this is a mistake. A representation theorem *presupposes* a rule of combination. Relative to a particular rule of combination, it demonstrates that, for any agent whose preferences satisfy certain axioms, there are internal attitudes with certain properties (unique to some extent) such that these internal attitudes determine the preferences *in line with that rule of combination*. As many authors have emphasized, given a different rule of combination, there will often be different internal attitudes with different properties that determine the same preferences, but this time in line with this different rule of combination (Zynda 2000; Eriksson and Hájek 2007; Meacham and Weisberg 2011).

While both constructivists and realists must appeal to rules of combination, my argument for the EU Rule of Combination applies primarily to the realist. I attempt to show that, for an agent with a credence function of a certain sort and a utility function, they are irrational if they don't combine those two functions in a particular way and set their preferences in line with that way of combining them. It is directed at an agent whose credence function and utility function have a

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

psychological reality beyond her being represented as having them. Thus, it does not apply to the constructivist, who thinks of the credence function and utility function as merely convenient mathematical ways of representing the preference ordering.

2.4. Estimates and the EU rule of combination

Finally, we come to state our argument in favour of the EU Rule of Combination. It draws on a mathematical result due to Bruno de Finetti, which we present as Theorem 1 below.

(EU1) A rational agent will weakly prefer one option to another if, and only if, her estimate of the utility of the first is at least her estimate of the utility of the second.
(EU2) A rational agent's estimate of a quantity will be her subjective expectation of it.
(3) Therefore,
(EUC) A rational agent's preference ordering $\succeq$ will be determined by the EU Rule of Combination.

The first premise (EU1) is supposed to be intuitively plausible. Suppose I desire only chocolate – obtaining as much of it as possible is my only goal. And suppose my estimate of the quantity of chocolate in the wrapper on my left is greater than my estimate of the quantity of chocolate in the wrapper on my right. And suppose that, nonetheless, I weakly prefer choosing the chocolate in the wrapper on my right. Then I would seem irrational. Likewise, if I desire only utility – surely an analytic truth if there are any – then I would seem irrational if my estimate of the utility of an act g were higher than my estimate of the utility another act f and yet I were to weakly prefer f to g. This is the argument for premise (EU1).

The second premise (EU2) is based on a mathematical argument together with a plausible claim about the goodness of estimates. Estimates, so the plausible claim goes, are better the closer they are to the true quantity they estimate. Indeed, we might take this to be an analytic truth. That is, we might take it to be a necessary condition on something being an estimate of a quantity that it is valued for its proximity to the actual value of that quantity. Thus, if I estimate that the amount of chocolate remaining in my cupboard is 73 g and my friend estimates that it is 79 g and in fact it is 80 g, then her estimate is better than mine. The mathematical argument is a generalization of a result due to de Finetti, which says, very roughly, that if an agent has estimates that are not expectations of the quantities that they estimate, there are alternative estimates of those quantities that are guaranteed to be closer to the true values of the quantities; so estimates that aren't expectations are irrational.

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

Let's make all of this precise. Suppose X is a quantity. Mathematically, we might understand this as a random variable, which is a function that takes a state of the world s and returns $X(s)$, which is the value that this quantity takes in this state of the world. Thus, if C is the quantity of chocolate in my cupboard in grams, and @ is the actual state of the world, where there is 80 g of chocolate in my cupboard, then $C(@) = 80$. Now, given what we said above about the goodness of estimates, we can measure the *badness* or *disvalue* of an estimate $e(X)$ of a quantity X given a state of the world s by the distance between $e(X)$ and $X(s)$. Now, I will focus on just one measure of distance here, for the sake of simplicity, but the result also holds of a wide range of distance measures that mathematicians call the *Bregman divergences*.[4] Having said that, in Section 2.5, I will offer a reason to prefer the measure of distance I use here to all other measures. The measure of distance between two numbers x and y that I will use here is the square of their difference $|x - y|^2$, which is itself a Bregman divergence. For obvious reasons, we call this the *quadratic distance measure* and we write it $\mathfrak{q}(x, y) := |x - y|^2$. Thus, relative to this measure of distance, the badness of an estimate $e(X)$ of a quantity X given a state of the world s is $\mathfrak{q}(e(X), X(s)) = |e(X) - X(s)|^2$.

Now, in the argument we wish to give for (EU2), we are interested in evaluating the goodness or badness not only of a single estimate of a single quantity, but also of a set of estimates of a set of quantities. So our next job is to say how we measure this. Suppose $\mathcal{X}$ is a set of quantities for which our agent has estimates. One of these quantities might be the quantity of chocolate in my cupboard, for instance; another might be the quantity of chocolate in my hand; another still might be the distance between my house and the nearest chocolate shop; and so on. And suppose that e is her *estimate function* – that is, e takes each quantity X in $\mathcal{X}$ and returns our agent's estimate $e(X)$ of that quantity. Then we will measure the badness of an estimate function at a state of the world by adding together the badness of each of the individual estimates that comprise it. Thus, the badness of e at the state of the world s is the sum of each $\mathfrak{q}(e(X), X(s))$ for each X in $\mathcal{X}$. So the badness of an estimate function e defined on the quantities in $\mathcal{X}$ at a state of the world s is

$$\mathfrak{B}(e, s) := \sum_{X \in \mathcal{X}} \mathfrak{q}(e(X), X(s)) = \sum_{X \in \mathcal{X}} |e(X) - X(s)|^2$$

With these notions defined, we have nearly defined everything that we need for our argument for (EU2). But there is one final observation to make. Consider our credences. It seems natural to say that my credence in a true proposition is better, epistemically speaking, the closer it is to the maximal credence, which is 1. And it seems natural to say that my credence in a false proposition is better, epistemically speaking, the closer it is to 0. That is, our credence in a proposition can be seen as an estimate of a particular quantity, namely, the quantity that takes value 1 if the proposition is true and value 0 if the proposition is false

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

(Jeffrey 1986; Joyce 1998). Given a proposition A, call this the *indicator quantity for A*: thus, $A(s) = 0$ if A is false at s; $A(s) = 1$ if A is true at s.

Now, suppose that our agent has credences in all propositions in a finite algebra $\mathcal{F}$. Let's say that her *credence function* is the function that assigns to each of these propositions her credence in it. Then the observation that a credence in a proposition is, or at least should be evaluated as if it is, an estimate of the indicator quantity for that proposition suggests that the badness of a credence function c at a state of the world s should be given by

$$\mathfrak{I}(c, s) := \sum_{A \in \mathcal{F}} \mathfrak{q}(c(A), A(s)) = \sum_{A \in \mathcal{F}} |c(A) - A(s)|^2$$

That is, it is the sum of the distance between each credence, $c(A)$, and the indicator quantity, A, corresponding to the proposition to which the credence is assigned.

Finally, we can say that an agent with a credence function c defined on the propositions in the finite algebra $\mathcal{F}$ and an estimate function e defined on a finite set of quantities $\mathcal{X}$ should be evaluated at a state of the world s as follows: her badness is given by

$$\mathfrak{I}(c, s) + \mathfrak{I}(e, s) = \sum_{A \in \mathcal{F}} \mathfrak{q}(c(A), A(s)) + \sum_{X \in \mathcal{X}} \mathfrak{q}(e(X), X(s))$$

That completes our account of how badly an agent is doing who has credences in propositions in $\mathcal{F}$ and certain estimates in quantities in $\mathcal{X}$.

Next, we turn to what we might show using this account. Let's say that our agent's credence function c defined on finite algebra $\mathcal{F}$ is *probabilistic* if

(i) (Range) $0 \leq c(A) \leq 1$ for all A in $\mathcal{F}$.
(ii) (Normalization) $c(\top) = 1$, where $\top$ is the tautologous proposition; that is, it is true at all states of the world.
(iii) (Additivity) $c(A \vee B) = c(A) + c(B)$ if A and B are mutually exclusive propositions; that is, A and B are not true together at any state of the world.

Now suppose that c is probabilistic. Then we say that the estimate function e defined on $\mathcal{X}$ is *expectational with respect to c* if

(iv) (Expectation) For each X in $\mathcal{X}$,

$$e(X) = \sum_{s \in S} c(s)X(s)$$

So an estimate function is expectational with respect to a probabilistic credence function if its estimate of each quantity is the weighted sum of the possible values of that quantity where the weights are given by the credence assigned to the relevant state of the world by the credence function. We say that a pair

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

(c, e) is *probabilistic and expectational* if c is probabilistic and e is expectational with respect to c – that is, if they jointly satisfy (i)–(iv).

For instance, suppose there are just two states of the world, s_1 and s_2. And let C be the quantity of chocolate in my cupboard in grams. Let's suppose that, in state s_1, there is a meagre 80 g of chocolate in my cupboard (so $C(s_1) = 80$), whereas in state s_2, there is veritable bounty, namely, 1000 g (so $C(s_2) = 1000$). And suppose that, having resisted the temptation to indulge in wishful thinking, I have credence $c(s_1) = 0.9$ in state s_1 and $c(s_2) = 0.1$ in state s_2. Then my credences are probabilistic (since they sum to 1), and my estimate of C is expectational with respect to my credences just in case it is $e(C) = c(s_1)C(s_1) + c(s_2)C(s_2) = (0.9 \times 80) + (0.1 \times 1000) = 172$.

In order to establish (EU2), the second premise of the argument for the EU Rule of Combination, we need to show that it is a requirement of rationality that an agent have a probabilistic credence function and an estimate function that is expectational with respect to it. Our argument is based on the following mathematical theorem, which is due to de Finetti (1974, 136).

Theorem 1 (de Finetti) *Suppose c is a credence function on $\mathcal{F}$ and e is an estimate function on $\mathcal{X}$.*

(i) If (c, e) is not probabilistic and expectational, then there is another pair (c', e') that is probabilistic and expectational such that

for all states of the world s.

$$\mathfrak{I}(c', s) + \mathfrak{I}(e', s) < \mathfrak{I}(c, s) + \mathfrak{I}(e, s)$$

(ii) If (c, e) is probabilistic and expectational, then there is no other pair (c', e') such that

$$\mathfrak{I}(c', s) + \mathfrak{I}(e', s) \leq \mathfrak{I}(c, s) + \mathfrak{I}(e, s)$$

for all states of the world s.

Thus, if an agent either has a credence function that is not a probability function, or if her credence function is a probability function but her estimates of quantities are not all given by her expectations of those quantities relative to that credence function, then there are alternative credences and estimates that, taken together, will be less bad than her credences and estimates taken together; that is, the alternative credences and estimates will be closer to the quantities that they estimate however those quantities turn out to be. What's more, if her credence function is a probability function, and if her estimates are given by her expectations, then there are no alternative credences and estimates that are guaranteed to be better than hers; indeed, there are no alternative credences and estimates that are guaranteed to do at least as well as hers. I provide a proof of this result in the Appendix.

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

This gives us an argument for having credences that are probabilities and estimates that are expectations. If we fail to do this, the theorem says, there are alternative credences and estimates we might have had that are guaranteed to do better than our credences; and there is nothing that is guaranteed to do better than those alternatives; indeed, there is nothing else that is even guaranteed to do just as well as them. Compare: I am offered two gambles on a fair coin toss. On the first, if the coin lands heads, I receive £5; if it lands tails, I lose £6. On the second, if the coin lands heads, I receive £10; if it lands tails, I lose £3. Now suppose I choose the first gamble. You would charge me with irrationality. After all, the second is guaranteed to be better than the first; whether the coin comes up heads or tails, I'll end up with more money if I've taken the second gamble. We are using a similar piece of reasoning here to argue that an agent is irrational if she has credences that are not probabilities, or if she has credences that are probabilities, but her estimates are not expectations with respect to them. That is, we are appealing to the so-called *Dominance Principle*, which says that an option is irrational if there is an alternative that is guaranteed to be better than it, and if there is nothing that is guaranteed to be better than that alternative. This completes our justification of the second premise (EU2) of our argument for the EU Rule of Combination.

You might worry here that, in the preceding justification of (EU2), we appeal to one principle of rational choice in order to justify another: we are appealing to the Dominance Principle in order to establish the EU Rule of Combination. And of course we are. But that is permissible in this context. After all, the Dominance Principle is an uncontroversial principle of decision theory. It is agreed upon by all parties to the current debate. Buchak and all other non-expected utility theorists disagree with me and other expected utility theorists about how credences and utilities should combine to give preferences. But we all agree that if one option is guaranteed to be better than another, and there is nothing that is guaranteed to be better than the first, then the second is irrational. So the argument strategy is legitimate.

Having given our justification for (EU2), this completes our argument for the EU Rule of Combination. According to the first premise (EU1), our preference ordering over acts should match our estimates of the utility of those acts: that is, I should weakly prefer one act to another iff my estimate of the utility of the first is at least as great as my estimate of the utility of the second. According to the second premise (EU2), our estimate of a given quantity, whether it is the utility of an act or the mass of chocolate in my fridge, should be our expectation of that quantity; that is, it should be the weighted average of the possible values that that quantity might take where the weights are given by our credences in the relevant states of the world. Putting these together, we obtain the EU Rule of Combination.

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

2.5. Measuring the badness of estimates

Before we leave our argument for the EU Rule of Combination, it is worth noting two things about the distance measure $\mathfrak{q}$ that we used to give the badness of an estimate of a given quantity at a given state of the world, and the function $\mathfrak{T}$ that we used to give the badness of a set of estimates in a set of quantities at a given world. First, as noted above, Theorem 1 holds for a wide range of alternative measures of distance; indeed, for any of the so-called Bregman divergences. However, second, it is also true that Theorem 1 fails for a wide range of alternative measures; indeed, it fails for the so-called *absolute value measure* $\mathfrak{a}$, which takes the distance between real numbers x and y to be the difference between them, that is, $\mathfrak{a}(x, y) := |x - y|$. Thus, this argument will be compelling to the extent that we can justify using the quadratic measure $\mathfrak{q}$, or some other Bregman divergence, instead of the absolute value measure $\mathfrak{a}$. Arguments have been given for this assumption in the case where we are measuring only the badness of credences (D'Agostino and Sinigaglia 2010; Leitgeb and Pettigrew 2010; Pettigrew ta 2016). In this context, it looks most promising to extend the argument of D'Agostino and Sinigaglia (2010). The arguments of Leitgeb and Pettigrew (2010) assume too much, and the argument of Pettigrew (ta) is too closely bound to the case of credences.

Above, we assumed that we begin with a measure $\mathfrak{d}$ of the distance between a single estimate $e(X)$ of a single quantity X and the true value $X(s)$ of X at a state of the world s; and then we measure the distance between an entire estimate function e defined on a set $\mathcal{X}$ of quantities, on the one hand, and the true values of those quantities at a state of the world s, on the other hand, by summing the distances, $\mathfrak{d}(e(X), X(s))$, between each $e(X)$ and $X(s)$ for X in $\mathcal{X}$. If we adapt the argument given by D'Agostino and Sinigaglia (2010), we do not make this assumption: instead, we justify it. That is, we assume that the badness of an estimate function e defined on $\mathcal{X}$ at a state of the world s is given by some function $\mathfrak{D}(e, s)$, and we lay down conditions on this function such that, if $\mathfrak{D}$ satisfies all of the conditions, then there is a continuous and strictly increasing function $H{:}\mathbb{R} \to \mathbb{R}$ such that

$$\mathfrak{D}(e, s) = H\left(\sum_{X \in \mathcal{X}} \mathfrak{q}(e(X), X(s))\right) = H\left(\sum_{X \in \mathcal{X}} |e(X) - X(s)|^2\right)$$

Since we appealed to the Dominance Principle in order to justify the EU Rule of Combination, and since the Dominance Principle pays attention only to the *ordering* of options at a world, rather than their *cardinal utilities* at that world, it does not matter whether we use the sum of the squared differences between the values or whether we use some strictly increasing transformation of that sum. Thus, this characterization of $\mathfrak{D}$ is sufficient for our purposes.

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

Here are the conditions that D'Agostino and Sinigaglia (2010) place on our measure $\mathfrak{D}$ of the badness of an estimate function.[5]

Extensionality Let us say that the *estimate profile* of e at s is the multiset of all pairs $(e(X), X(s))$ for X in $\mathcal{X}$ – that is, $\{\{(e(X), X(s)) : X \in \mathcal{X}\}\}$.[6] Then, if e has the same estimate profile at s as e' has at s', then $\mathfrak{D}(e, s) = \mathfrak{D}(e', s')$.

That is, the badness of your estimate function is a function only of its estimate profile. It does not depend on the particular quantities to which you assign estimates. If you and I assign estimates to very different quantities, but our estimate profiles match, then our estimates are exactly as bad as each other.

Accurate Extension Invariance If e is an estimate function on $\mathcal{X}$ and $\mathcal{X}' \subseteq \mathcal{X}$, then let $e|_{\mathcal{X}'}$ be the estimate function on $\mathcal{X}'$ that agrees with e on all quantities in $\mathcal{X}'$ – $e|_{\mathcal{X}'}$ is sometimes called *the restriction of e to* $\mathcal{X}'$. Then, if the estimates that e assigns to quantities not in $\mathcal{X}'$ are equal to the true values of those quantities at s, then $\mathfrak{D}(e|_{\mathcal{X}'}, s) = \mathfrak{D}(e, s)$.

That is, adding perfectly accurate estimates to your estimate function does not affect its badness.

Difference Supervenience If e assigns an estimate to just one quantity X, then $\mathfrak{D}(e, s) = g(|e(X) - X(s)|)$ for some continuous and strictly increasing function $g{:}\mathbb{R} \to \mathbb{R}$.

That is, the badness of a single estimate is a continuous and strictly increasing function of the difference between that estimate and the true value of the quantity.

Separability If $\mathcal{X}' \subseteq \mathcal{X}$ and

(i) $\mathfrak{D}(e|_{\mathcal{X}'}, s) = \mathfrak{D}(e'|_{\mathcal{X}'}, s)$ and
(ii) $\mathfrak{D}(e|_{\mathcal{X}-\mathcal{X}'}, s) < \mathfrak{D}(e'|_{\mathcal{X}-\mathcal{X}'}, s)$,

then $\mathfrak{D}(e, s) < \mathfrak{D}(e', s)$.

That is, if two estimate functions are equally bad on some subset of the quantities to which they assign estimates, then one is worse than the other if it is worse on the remaining quantities.

Taken together, these four properties entail that there are continuous and strictly increasing functions $H{:}\mathbb{R} \to \mathbb{R}$ and $f{:}\mathbb{R} \to \mathbb{R}$ such that:

$$\mathfrak{D}(e, s) = H\left(\sum_{X \in \mathcal{X}} f(|e(X) - X(s)|)\right)$$

Indeed, these four conditions are equivalent to the existence of two such functions H and f. Thus, what we need for our conclusion is a further property that ensures that $f(x) = x^2$. That is the job of the final condition on $\mathfrak{D}$. To state it, we need the notion of an *order-reversing swap*. Suppose e is an estimate function defined on a set of quantities $\mathcal{X}$ and s is a state of the world. And suppose that the estimates that e assigns to quantities X and Y are ordered correctly at s. That is,

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

(i) $e(X) > e(Y)$ and $X(s) > Y(s)$ or
(ii) $e(X) < e(Y)$ and $X(s) < Y(s)$.

Then, if we define e_{XY} to be the estimate function that is obtained from e by swapping its estimates for X and Y – so $e_{XY}(X) = e(Y)$ and $e_{XY}(Y) = e(X)$ – then we say that e_{XY} is an *order-reversing swap of e*, since the estimates that e_{XY} assigns to quantities X and Y are ordered *in*correctly at s. Our next condition says two things: first, it says that an order-reversing swap always increases the badness of the estimates; second, it says that if you compare two order-reversing swaps on the same estimate function and if (a) the two swaps involve swapping estimates that are themselves equally far apart and (b) the two swaps involve quantities whose true values are equally far apart, then the badness of the swaps is equal. The motivation for this condition is the following: The badness of a set of estimates is supposed to be determined by the extent to which they match the truth about the quantities that they estimate. As well as matching the *quantitative* facts about those quantities – such as their values – it also seems to be a good thing to match the *qualitative* facts about them – such as their ordering. Clearly e matches this qualitative fact for X and Y, whereas e_{XY} does not. Thus, other things being equal, e_{XY} is worse than e. And other things are equal, since all that has changed in the move from e to e_{XY} is that the quantities to which the estimates $e(X)$ and $e(Y)$ are assigned have been switched. Moreover, if we consider two possible order-reversing swaps on the same estimate function where (a) and (b) hold, then the effect of each swap on the badness of the estimate function should be the same, since there is nothing to tell between them.

The Badness of Order-Reversing Swaps Suppose e is defined on $\mathcal{X}$ and suppose X, Y, X', Y' are quantities in $\mathcal{X}$. And suppose $e(X), e(Y)$ are ordered as $X(s), Y(s)$ are; and $e(X'), e(Y')$ are ordered as $X'(s), Y'(s)$ are. And suppose $|e(X) - e(Y)| = |e(X') - e(Y')|$ and $|X(s) - Y(s)| = |X'(s) - Y'(s)|$. Then $\mathfrak{D}(e, s) < \mathfrak{D}(e_{XY}, s) = \mathfrak{D}(e_{X'Y'}, s)$.

We now have the following theorem:

Theorem 2 (D'Agostino & Dardanoni) *If $\mathfrak{D}$ satisfies Extensionality, Accurate Extension Invariance, Difference Supervenience, Separability and The Badness of Order-Reversing Swaps, then there is a continuous and strictly increasing functions H:$\mathbb{R} \to \mathbb{R}$ such that:*

$$\mathfrak{D}(e, s) = H\left(\sum_{X \in \mathcal{X}} |e(X) - X(s)|^2\right)$$

This gives us what we need.

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

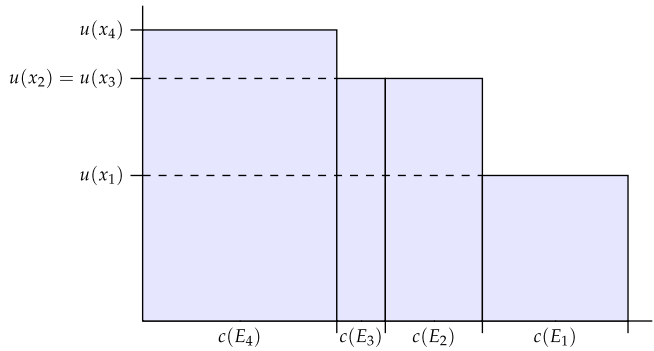

Figure 1. The expected utility $EU_{c,u}(h)$ of h is given by the grey area.Note: It is obtained by summing the areas of the four vertical rectangles: working from right to left, their areas are $c(E_1)u(x_1), \ldots, c(E_4)u(x_4)$.

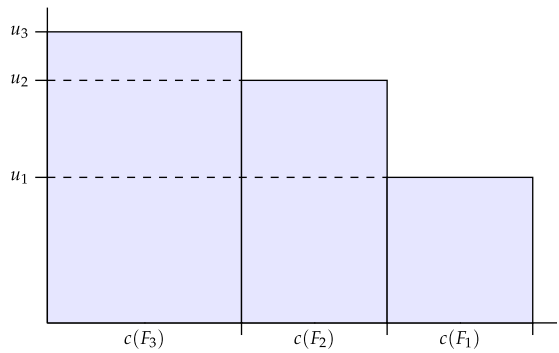

Figure 2. Again, the expected utility $EU_{c,u}(h)$ of h is given by the grey area.Note: It is obtained by summing the areas of the three vertical rectangles (the middle two vertical rectangles from Figure 1 have been merged): working from right to left, their areas are $c(F_1)u_1, \ldots, c(F_3)u_3$.

3. Expected utility and risk-weighted expected utility

In the previous section, we gave our defence of the EU Rule of Combination. In this section, we describe Lara Buchak's proposed alternative. To do this, we'll illustrate the difference between expected utility and risk-weighted expected utility using a particular act as an example. We'll first describe the expected utility of that act, and then we'll show how to define its risk-weighted expected utility. Our example is the following act: $h = \{E_1, x_1; \ldots; E_4, x_4\}$. The agent's probabilistic credences over the events $E_1, \ldots, E_4$ and her utilities for the outcomes $x_1, \ldots, x_4$ are given as follows:

	$(x_1, \overline{x_1})$	$(x_2, \overline{x_2})$	$(x_3, \overline{x_3})$	$(x_4, \overline{x_4})$
u^*	3	5	5	6

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

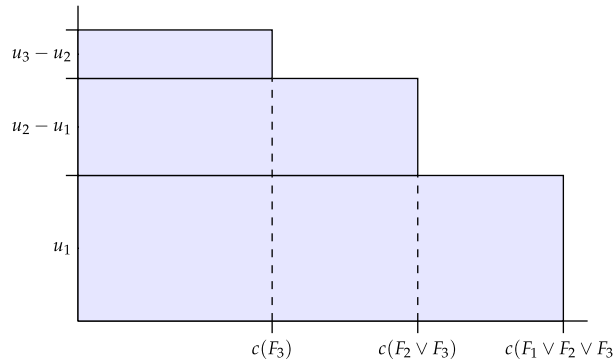

Figure 3. Once again, the expected utility $EU_{c,u}(h)$ of h is given by the grey area.Notes: It is obtained by summing the areas of the three *horizontal* rectangles. Working from bottom to top, their areas are $c(F_1 \vee F_2 \vee F_3)u_1 = u_1$, $c(F_2 \vee F_3)(u_2 - u_1)$ and $c(F_3)(u_3 - u_2)$.

In Figure 1, we exploit a useful diagrammatic way of representing the expected utility of h, which is used by Quiggin (1993), Wakker (2010) and Buchak (2013).

Figure 1 suggests two ways in which we might reformulate $EU_{c,u}(f)$. These will be very useful in understanding how expected utility theory relates to Buchak's proposal.

- First, it is clear that $EU_{c,u}(f)$ depends only on the utilities of the outcomes to which the act f may give rise and the probabilities that f will produce outcomes with those utilities. Thus, given an act $f = \{E_1, x_1; \ldots; E_n, x_n\}$ and a utility function u, we might redescribe f as $\{F_1, u_1; \ldots; F_k, u_k\}$ where

- $u_1, \ldots, u_k$ are the utilities to which f might give rise ordered from least to greatest – that is, $u_1 < \ldots < u_k$. For instance, in our example act h: $u_1 = 3$, $u_2 = 5, u_3 = 6$.
- F_j is the proposition that f will give rise to u_j. Thus, $F_j = \{s \in \mathcal{S}: u(h(s)) = u_j\}$. For instance, in our example act h: $F_1 \equiv E_1, F_2 \equiv E_2 \vee E_3, F_3 \equiv E_4$. We call this the *ordered utility-based description of f relative to u*. Then

$$EU_{c,u}(f) = \sum_{j=1}^{k} c(F_j)u_j$$

Figure 2 illustrates this reformulation of expected utility for our example act h.

- The second reformulation of $EU_{c,u}(f)$ builds on this first and is illustrated in Figure 3. Suppose $f = \{F_1, u_1; \ldots; F_k, u_k\}$ is the ordered utility-based description of f relative to u. Then

$$EU_{c,u}(f) = u_1 + \sum_{j=2}^{k} c(F_j \vee \ldots \vee F_k)(u_j - u_{j-1})$$

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

> Again, the expected utility of an act is given by a weighted sum: but this time the quantities to be weighted are the differences between one possible utility and the possible utility immediately below it; and the weight assigned to that difference is the probability that the act will give rise to at least that much utility.

With this in hand, we're ready to formulate Buchak's alternative to expected utility theory. Buchak is motivated by the apparent rationality of risk-sensitive behaviour. Notoriously, some seemingly rational risk-sensitive behaviour cannot be captured by expected utility theory at all: for instance, Allais described seemingly rational preferences that cannot be generated by any rational credence function and utility function in the way prescribed by expected utility theory (Allais 1953). Moreover, there are other seemingly rationally preferences that can be generated by a credence function and utility function in line with expected utility theory, but which seem to be rational even for agents who do not have credences and utilities that would generate them in this way. Thus, for instance, consider the two acts described in the introduction to this article: *Safe* = { Heads $\vee$ Tails, £50 }and *Risky* = { Heads, £100; Tails, £0 } Suppose that our agent strictly prefers *Safe* to *Risky*: that is, *Safe* $\succ$ *Risky*. Can expected utility theory capture the rationality of this preference? Suppose that, since the coin is known to be fair, rationality requires that the agent assigns credences to the two states of the world as follows: $c(\text{Heads}) = 0.5 = c(\text{Tails})$. Then it is still possible to describe a utility function on the outcomes £0, £50, £100 that generates these preferences in the way expected utility theory requires. Let $u(£0) = 0$ and $u(£100 = £50 + £50) < u(£50) + u(£50)$. That is, suppose the agent treats money as a *dependent good*: how much utility it gives depends on how much of it she has already; so, money has diminishing marginal utility for this agent. Then, for an agent with this credence function and utility function, $EU_{c,u}(Safe) > EU_{c,u}(Risky)$, as required. So expected utility theory can capture the rationality of these preferences. However, as Buchak rightly observes, those preferences – that is, *Safe* $\succ$ *Risky* – seem rational not only for an agent for whom money has diminishing marginal utility; they seem rational even for an agent whose utility is linear in money. And this is something that expected utility cannot capture. Thus, Buchak is interested not only in saving the Allais preferences, but also in saving other risk-sensitive behaviour without attributing the risk-sensitive behaviour to the shape of the utility function (Buchak 2013, Chapter 1).

How does Buchak hope to capture these risk-sensitive preferences? Where expected utility theory countenances only two types of internal attitude as relevant to preferences, Buchak countenances a third as well: this third component is supposed to capture the agent's attitude to risk, and it is given by a function $r:[0, 1] \to [0, 1]$, which Buchak assumes to be strictly increasing, continuous, and taking the following values, $r(0) = 0$ and $r(1) = 1$ (Buchak 2013, Section 2.2). Buchak's *risk-weighted expected utility theory* then employs the following

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

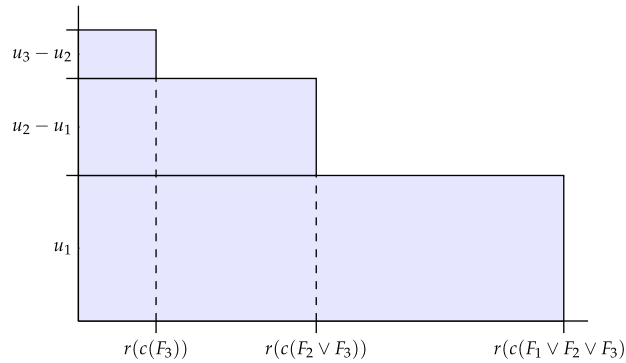

Figure 4. The risk-weighted expected utility $REU_{r_2,c,u}(h)$ of h is given by the grey area, where $r_2(x) := x^2$.

rule of combination, which states how an agent's internal and external attitudes ought to relate, where the agent has credence function c, utility function u and risk function r:

> *REU Rule of Combination* (Buchak 2013, 53) Suppose $f = \{F_1, u_1; \ldots; F_k, u_k\}$ is the ordered utility-based description of act f relative to utility function u. Then let
>
> $$REU_{r,c,u}(f) := u_1 + \sum_{j=2}^{k} r(c(F_j \vee \ldots \vee F_k))(u_j - u_{j-1})$$
>
> Then if the agent is rational, then
>
> $$f \succeq g \iff REU_{r,c,u}(f) \geq REU_{r,c,u}(g)$$

In Figure 4, we illustrate the risk-weighted expected utility of our example act h when the agent has the risk function $r_2(x) := x^2$. Notice that the formulation of $REU_{r,c,u}(f)$ is exactly like the formulation of $EU_{c,u}(f)$ that we gave above except that each probability weight is transformed by the agent's risk function. Thus, if $r(x) < x$ (for all $0 < x < 1$), then, as Figure 4 illustrates, the lowest utility to which the act can give rise – namely, u_1 – contributes just as much to $REU_{r,c,u}(f)$ as it does to $EU_{c,u}(f)$ – it contributes u_1 to both. But further increases in utility – such as the increase from getting at least utility u_1 to getting at least u_2 – make less of a contribution since their probability – $c(F_2 \vee F_3)$ – is acted on by the risk function, and it is this reduced value – $r(c(F_2 \vee F_3))$ – that weights the possible increases in utility. Thus, such an agent displays risk-averse behaviour. r_2 is such a risk function.

Similarly, if $r(x) > x$ (for all $0 < x < 1$), then the lowest utility to which the act can give rise contributes just as much to $REU_{r,c,u}(f)$ as it does to $EU_{c,u}(f)$, but further increases in utility make more of a contribution since their probability is acted on by the risk function and it is this increased value that weights the

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

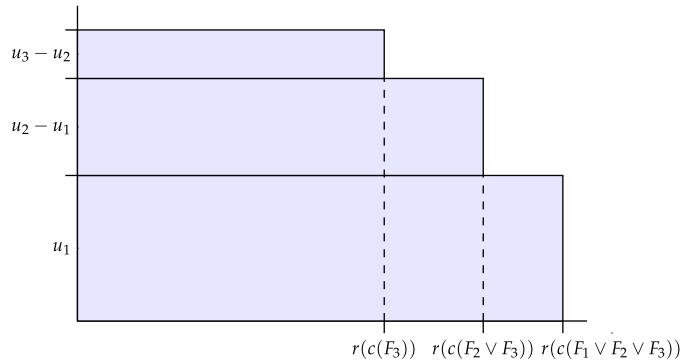

Figure 5. The risk-weighted expected utility $REU_{r_{0.5},c,u}(h)$ of h is given by the grey area, where $r_{0.5}(x) := \sqrt{x}$.

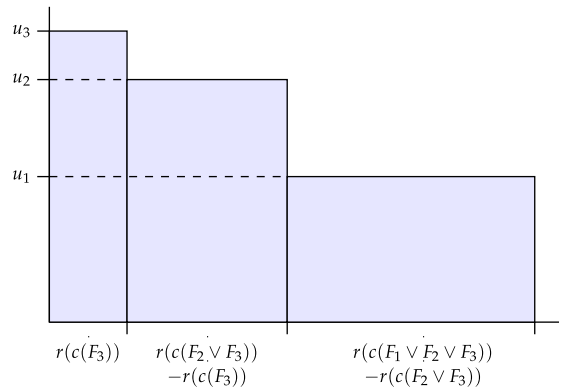

Figure 6. Again, the risk-weighted expected utility $REU_{r_2,c,u}(f)$ of f is given by the grey area, where $r_2(x) = x^2$.

possible increases in utility. This is illustrated in Figure 5. Such an agent displays risk-seeking behaviour. $r_{0.5}(x) := \sqrt{x}$ is such a risk function.

It's also easy to see that, if $r_1(x) := x$ (for $0 \leq x \leq 1$), then $REU_{r_1,c,u}(f) = EU_{c,u}(f)$. Thus, expected utility theory is the special case of risk-weighted expected utility theory given by a linear risk function. In such a situation, we say that the agent is *risk-neutral*. This means that Buchak's theory permits any preferences that expected utility theory permits. But it also permits a whole lot more. For instance, one can easily recover the Allais preferences or the preference *Safe* $\succ$ *Risky* described above by attributing to an agent a certain sort of risk function – in both cases, a risk-averse risk function.

This, then, is Buchak's proposal.

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

4. Redescribing the outcomes

Moving from expected utility theory to risk-weighted expected utility theory involves an agent evaluating an act in the way illustrated in Figure 3 to evaluating it in the way illustrated in Figure 4. In order to begin to see how we can redescribe the REU rule of combination as an instance of the EU rule of combination, we reformulate the REU rule in the way illustrated in Figure 6.[7] Thus, we can reformulate $REU_{r,c,u}(f)$ as follows:

$$REU_{r,c,u}(f) = \sum_{j=1}^{k-1}(r(c(F_j \vee \ldots \vee F_k)) - r(c(F_{j+1} \vee \ldots \vee F_k)))u_j + r(c(F_k))u_n$$

And we can reformulate this as follows:

$$REU_{r,c,u}(f) = \sum_{j=1}^{k-1} c(F_j)\frac{r(c(F_j \vee \ldots \vee F_k)) - r(c(F_{j+1} \vee \ldots \vee F_k))}{c(F_j \vee \ldots \vee F_k) - c(F_{j+1} \vee \ldots \vee F_k)}u_j + c(F_k)\frac{r(c(F_k))}{c(F_k)}u_n$$

since $c(F_j \vee \ldots \vee F_k) - c(F_{j+1} \vee \ldots \vee F_k) = c(F_j)$.

Now, suppose we let

$$s_j = \begin{cases} \frac{r(c(F_j \vee \ldots \vee F_k)) - r(c(F_{j+1} \vee \ldots \vee F_k))}{c(F_j \vee \ldots \vee F_k) - c(F_{j+1} \vee \ldots \vee F_k)} & \text{if } j = 1, \ldots, k-1 \\ \frac{r(c(F_j))}{c(F_j)} & \text{if } j = k \end{cases}$$

Then we have:

$$REU_{r,c,u}(f) = \sum_{j=1}^{k} c(F_j)s_j u_j$$

Reformulating Buchak's rule of combination in this way suggests two accounts of it. On the first, utilities attach ultimately to outcomes x_i, and they are weighted not by an agent's credences but rather by a function of those credences that encodes the agent's attitude to risk (given by a risk function). On this account, we group $c(F_j)s_j$ together to give this weighting. Thus, we assume that this weighting has a particular form: it is obtained from a credence function c and a risk function r to give $c(F_j)s_j$; this weighting then attaches to u_j to give $(c(F_j)s_j)u_j$. This is the account that Buchak favours.

On the second account, credences do provide the weightings for utility, as in the EU rule of combination, but utilities attach ultimately to outcome-act pairs (x_i, f). On this account, we group $s_j u_j$ together to give this utility; this utility is then weighted by $c(F_j)$ to give $c(F_j)(s_j u_j)$. That is, we say that an agent's

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

utility function is defined on a new outcome space: it is not defined on a set of outcomes $\mathcal{X}$, but on a particular subset of $\mathcal{X} \times \mathcal{A}$, which we will call $\mathcal{X}^*$. $\mathcal{X}^*$ is the set of outcome-act pairs (x_i, f) such that x_i is a possible outcome of f: that is, $\mathcal{X}^* = \{(x, f) \in \mathcal{X} \times \mathcal{A} : \exists s \in \mathcal{S}(f(s) = x)\}$. Now, just as the first account assumed that the weightings of the utilities have a certain form – namely, they are generated by a risk function and probability function in a certain way – so this account assumes something about the form of the new utility function u^* on $\mathcal{X}^*$: we assume that a certain relation holds between the utility that u^* assigns to outcome-act pairs in which the act is the constant act over the outcome and the utility u^* to outcome-act pairs in which this is not the case. We assume that the following holds:

$$u^*(x, f) = s_j u^*(x, \overline{x}) \tag{1}$$

If a utility function on $\mathcal{X}^*$ satisfies this property, we say that it *encodes attitudes to risk relative to risk function r*. Thus, on this account an agent evaluates an act as follows:

- She begins with a risk function r and a probability function c.
- She then assigns utilities to all constant outcome-act pairs $(x, \overline{x})$, defining u^* on $\overline{\mathcal{X}}^*$, where $\overline{\mathcal{X}}^* = \{(x, \overline{x}) : x \in \mathcal{X}\} \subseteq \mathcal{X}^*$.
- Finally, she extends u^* to cover all outcome-act pairs in $\mathcal{X}^*$ in the unique way required in order to make u^* a utility function that encodes attitudes to risk relative to r. That is, she obtains $u^*(x, f)$ by weighting $u^*(x, \overline{x})$ in a certain way that is determined by her probability function and her attitudes to risk.

Let's see this in action in our example act h; we'll consider h from the point of view of two risk functions, $r_2(x) = x^2$ and $r_{0.5}(x) = \sqrt{x}$. Recall: r_2 is a risk-averse risk function; $r_{0.5}$ is risk seeking. We begin by assigning utility to all constant outcome-act pairs $(x, \overline{x})$:

Then we do the same trick as above and amalgamate the outcome-act pairs with the same utility: thus, again, F_1 is the event in which the act gives outcome-act pair (x_1, h), F_2 is the event in which it gives (x_2, h) or (x_3, h), and F_3 the event in which it gives (x_4, h). Next, we assign utilities to (x_1, h), (x_2, h), (x_3, h) and (x_4, h) in such a way as to make u^* encode attitudes to risk relative to the risk function r.

Let's start by considering the utility of (x_1, h), the lowest outcome of h. Suppose our risk function is r_2; then

$$\begin{aligned} u^*(x_1, h) :&= \frac{r_2(c(F_1 \vee F_2 \vee F_3)) - r_2(c(F_2 \vee F_3))}{c(F_1 \vee F_2 \vee F_3) - c(F_2 \vee F_3)} u^*(x_1, \overline{x_1}) \\ &= \frac{r_2(1) - r_2(0.7)}{1 - 0.7} = 1.7 u^*(x_1, \overline{x_1}) \end{aligned}$$

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

And now suppose our risk function is $r_{0.5}$; then

$$u^*(x_1, h) := \frac{r_{0.5}(c(F_1 \vee F_2 \vee F_3)) - r_{0.5}(c(F_2 \vee F_3))}{c(F_1 \vee F_2 \vee F_3) - c(F_2 \vee F_3)} u^*(x_1, \overline{x_1})$$
$$= \frac{r_{0.5}(1) - r_{0.5}(0.7)}{1 - 0.7} \approx 0.54 u^*(x_1, \overline{x_1})$$

Thus, the risk-averse agent – that is, the agent with risk function r_2 – values this lowest outcome x_1 as the result of h more than she values the same outcome as the result of a certain gift of x_1, whereas the risk-seeking agent – with risk function $r_{0.5}$ – values it less. And this is true in general: if $r(x) < x$ for all x, the utility of the lowest outcome as a result of h will be more valuable than the same outcome as a result of the constant act on that outcome; if $r(x) < x$ it will be less valuable.

Next, let us consider the utility of (x_4, h), the highest outcome of h. Suppose her risk function is r_2; then

$$u^*(x_4, h) := \frac{r_2(c(F_3))}{c(F_3)} u^*(x_4, \overline{x_4}) = \frac{r_2(0.4)}{0.4} u^*(x_4, \overline{x_4}) = 0.4 u^*(x_4, \overline{x_4})$$

And now suppose her risk function is $r_{0.5}$; then

$$u^*(x_4, h) := \frac{r_{0.5}(c(F_3))}{c(F_3)} u^*(x_4, \overline{x_4}) = \frac{r_{0.5}(0.4)}{0.4} u^*(x_4, \overline{x_4}) = 2.5 u^*(x_4, \overline{x_4})$$

Thus, the risk-averse agent – that is, the agent with risk function r_2 – values this *highest* outcome x_4 as the result of h *less* than she values the same outcome as the result of a certain gift of x_4, whereas the risk-seeking agent – with risk function $r_{0.5}$ – values it *more*. And, again, this is true in general: if $r(x) < x$ for all x, the utility of the highest outcome as a result of h will be less valuable than the same outcome as a result of the constant act on that outcome; if $r(x) < x$ it will be more valuable.

This seems right. The risk-averse agent wants the highest utility, but also cares about how sure she was to obtain it. Thus, if she obtains x_1 from h, she knows she was guaranteed to obtain at least this much utility from h or from $\overline{x_1}$ (since x_1 is the lowest possible outcome of each act). But she also knows that h gave her some chance of getting more utility. So she values (x_1, h) more than $(x_1, \overline{x_1})$. But if she obtains x_4 from h, she knows she was pretty lucky to get this much utility, while she knows that she would have been guaranteed that much if she had obtained x_4 from $\overline{x_4}$. So she values (x_4, h) less than $(x_4, \overline{x_4})$. And similarly, but in reverse, for the risk-seeking agent.

Finally, let's consider the utilities of (x_2, h) and (x_3, h), the middle outcomes of h. They will have the same value, so we need only consider the utility of (x_2, h). Suppose her risk function is r_2; then

$$u^*(x_2, h) := \frac{r_2(c(F_2 \vee F_3)) - r_2(c(F_3))}{c(F_2 \vee F_3) - c(F_3)} u^*(x_2, \overline{x_2})$$
$$= \frac{r_2(0.7) - r_2(0.4)}{0.7 - 0.4} u^*(x_2, \overline{x_2}) = 1.1 u^*(x_2, \overline{x_2})$$

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

Thus, again, the agent with risk function r_2 assigns higher utility to obtaining x_2 as a result of h than to obtaining x_2 as the result of $\overline{x_2}$. But this is not generally true of risk-averse agents. Consider, for instance, a more risk-averse agent, who has a risk function $r_3(x) := x^3$. Then

$$\begin{aligned} u^*(x_2, h) &:= \frac{r_3(c(F_2 \vee F_3)) - r_3(c(F_3))}{c(F_2 \vee F_3) - c(F_3)} u^*(x_2, \overline{x_2}) \\ &= \frac{r_3(0.7) - r_3(0.4)}{0.7 - 0.4} u^*(x_2, \overline{x_2}) = 0.93 u^*(x_2, \overline{x_2}) \end{aligned}$$

Again, this seems right. As we said above, the risk-averse agent wants the highest utility, but she also cares about how sure she was to obtain it. The less risk-averse agent – whose risk function is r_2 – is sufficiently sure that h would obtain for her at least the utility of x_2 *and possibly more* that she assigns higher value to getting x_2 as a result of h than to getting it as a result of $\overline{x_2}$. For the more risk-averse agent – whose risk function is r_3 – she is not sufficiently sure. And reversed versions of these points can be made for risk-seeking agents with risk functions $r_{0.5}$ and $r_{0.333}$, for instance. Thus, we can see why it makes sense to demand of an agent that her utility function u^* on $\mathcal{X}^*$ encodes attitudes to risk relative to a risk function in the sense that was made precise above – see Equation 1.

Since what we have just provided is a genuine redescription of Buchak's REU Rule of Combination, we can see that Buchak's representation theorem is agnostic between a version of REU in which utilities attach to elements of $\mathcal{X}$, and a version of EU in which utilities attach to elements of $\mathcal{X}^*$.

Theorem 3 (Buchak) *If $\succeq$ satisfies the Buchak axioms, there is a unique probability function c, unique risk function, and unique-up-to-affine-transformation utility function u on $\mathcal{X}$ such that $\succeq$ is determined by r, c and u in line with the REU rule of combination.*

And we have the following straightforward corollary:

Theorem 4 *If $\succeq$ satisfies the Buchak axioms, there is a unique probability function c and unique-up-to-affine*-transformation utility function u^* on $\mathcal{X}^*$ that encodes attitudes to risk relative to a risk function such that $\succeq$ is determined by c and u^* in line with the EU rule of combination (where u^* is unique-up-to-affine*-transformation if $u^*|_{\overline{\mathcal{X}}^*}$ is unique-up-to-affine-transformation).*

Thus, by redescribing the set of outcomes to which our agent assigns utilities, we can see how her preferences in fact line up with her estimates of the utility of her acts, as required by the de Finetti-inspired argument for the EU Rule of Combination given in the previous section.

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

5. What's wrong with redescription?

Although Buchak does not address precisely this particular version of the redescription strategy, she does consider others nearby. Against those, she raises what amount to two objections (Buchak 2013, Chapter 4). (Buchak raises a further objection against versions of the redescription strategy that attempt to identify certain outcome-act pairs to give a more coarse-grained outcome space; but these do not affect my proposal.)

5.1. The problem of proliferation

One potential problem that arises when one moves from assigning utilities to $\mathcal{X}$ to assigning them to $\mathcal{X}^*$ is that an element in the new outcome space is never the outcome of more than one act: (x, f) is a possible outcome of act f but not of any act g other than f. Thus, this outcome never appears in the expected utility (or indeed risk-weighted expected utility) calculation of more than one act. The result is that very few constraints are placed on the utilities that must be assigned to these new outcomes and the probabilities that must be assigned to the propositions in order to recover a given preference ordering on the acts via the EU (or REU) rule of combination. Suppose $\succeq$ is a preference ordering on $\mathcal{A}$. Then, for each act f in $\mathcal{A}$, pick a real number r_f such that $f \succeq g$ iff $r_f \geq r_g$. Now there are many ways to do this, and they are not all affine transformations of one another – indeed, any strictly increasing $\tau:\mathbb{R} \to \mathbb{R}$ will take one such assignment to another. Now pick any probability function c on $\mathcal{F}$. Now, given an act $f = \{E_1, x_1; \ldots; E_n, x_n\}$, the only constraint on the values $u^*(x_1, f), \ldots, u^*(x_n, f)$ is that $\sum_i c(E_i)u^*(x_i, f) = r_f$. And this of course permits many different values.[8] Buchak dubs this phenomenon *belief and desire proliferation* (Buchak 2013, 140).

Why is this a problem? There are a number of reasons to worry about belief and desire proliferation. There is the epistemological worry that, if utilities and probabilities are as loosely constrained as this, it is not possible to use an agent's observed behaviour to predict her unobserved behaviour. Divining her preferences between two acts will teach us nothing about the utilities she assigns to the outcomes of any other acts since those outcomes are unique to those acts. Also, those who wish to use representation theorems for the purpose of radical interpretation will be concerned by the complete failure of the uniqueness of the rationalization of preferences that such a decision theory provides.

Neither of these objections seems fatal to me. But in any case, the version of the redescription strategy presented here avoids them altogether. The reason is that I placed constraints on the sort of utility function u^* an agent can have over $\mathcal{X}^*$: I demanded that u^* encode attitudes to risk; that is, $u^*(x, f)$ is defined in terms of $u^*(x, \overline{x})$ in a particular way (given by Equation 1). And we saw in Theorem 4 above that, for any agent whose preferences satisfy the Buchak axioms, there is a unique probability function c and a unique utility function u^*

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

on $\mathcal{X}^*$ that encodes attitudes to risk relative to a unique risk function such that together c and u^* generate the agent's preferences in accordance with the EU Rule of Combination.

5.2. Ultimate ends and the locus of utility

Buchak's second objection initially seems more worrying (Buchak 2013, 137–138). A theme running through *Risk and Rationality* is that decision theory is the formalization of *instrumental* or *means-end reasoning*. One consequence of this is that an account of decision theory that analyses an agent as engaged in something other than means-end reasoning is thereby excluded.

Buchak objects to the redescription strategy on these grounds. According to Buchak, to understand an agent as engaged in means-end reasoning, one must carefully distinguish the means and the ends: in Buchak's framework, the means are the acts and the ends are the outcomes. One must then assign utilities to the ends only. Of course, in terms of these utilities and the agent's probabilities and possibly other representations of internal attitude such as the risk function, one can then assign value or utility to the means. But the important point is that this value or utility that attaches to the means is assigned on the basis of the assignment of utility to the ultimate ends. Thus, while there is a sense in which we assign a value or utility to means – i.e. acts – in expected utility theory, this assignment must depend ultimately on the utility we attach to ends – i.e. outcomes.

Thus, a first pass at Buchak's second complaint against the redescription strategy is this: the redescription strategy assigns utilities to something other than ends – it assigns utilities to outcome-act pairs, and these are fusions of means and ends. Thus, an agent analysed in accordance with the redescription strategy is not understood as engaged in means-end reasoning.

However, this seems problematic in two ways. Whether they constitute ultimate ends or not, there are at least two reasons why an agent *must* assign utilities to outcome-act pairs rather than outcomes on their own. That is, there are two reasons why at least this part of the redescription strategy – namely, the move from $\mathcal{X}$ to $\mathcal{X}^*$ – is necessary irrespective of the need to accommodate risk in expected utility theory.

Firstly, utilities must attach to the true outcomes of an act. But these true outcomes aren't the sort of thing we've been calling an outcome here. When I choose *Safe* over *Risky* and receive £50, the outcome of that act is not merely £50; it is £50 *as the result of Safe*. Thus, the true outcomes of an act are in fact the elements of $\mathcal{X}^*$ – they are what we have been calling the outcome-act pairs.

Of course, at this point, Buchak might accept that utilities attach to outcome-act pairs, but insist that it is nonetheless a requirement of rationality that an agent assign the same utility to two outcome-act pairs with the same act component; that is, $u^*(x, f) = u^*(x, g)$; that is, while utilities attach to fusions

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

of means and ends, they must be a function only of the ends. But the second reason for attaching utilities to outcome-act pairs tells against this claim in general. The reason is this: As Bernard Williams urges, it is neither irrational nor even immoral to assign higher utility to a person's death as a result of something other than my agency than to that same person's death as a result of my agency (Williams and Smart 1973). This, one might hold, is what explains my hesitation in a Williams-style example in which I must choose whether or not to shoot a particular individual when I know that, if I don't shoot him, someone else will. I assign higher utility to the death of that person at the hands of someone else than to the death of that person at my hands. Thus, it is permissible in at least some situations to care about the act that gives rise to the outcome and let one's utility in an outcome-act pair be a function also of that act.

Nonetheless, this is not definitive. After all, Buchak could reply that this is peculiar to acts that have morally relevant consequences. Acts such as those in the Allais paradox do not have morally relevant consequences; but the redescription strategy still requires us to make utilities depend on acts as well as outcomes in those cases. Thus, for non-moral acts f and g, Buchak might say, it is a requirement of rationality that $u^*(x, f) = u^*(x, g)$, even if it is not such a requirement for moral cases. And this would be enough to scupper the redescription strategy.

However, it is not clear why the moral and non-moral cases should differ in this way. Consider again the Williams-style example from above: I must choose whether to shoot an individual or not; I know that, if I do not shoot him, someone else will. I strictly prefer not shooting him to shooting him. My reasoning might be reconstructed as follows: I begin by assigning a certain utility to this person's death as the result of something other than my agency – natural causes, for instance, or murder by a third party. Then, to give my utility for his death at my hand, I weight this original utility in a certain way, reducing it on the basis of the action that gave rise to the death. Thus, the badness of the outcome-act pair (*X's death*, *My agency*) is calculated by starting with the utility of another outcome-act pair with the same outcome component – namely, (*X's death*, *Not my agency*) – and then weighting that utility based on the act component. We might call (*X's death*, *Not my agency*) the *reference pair attached to the outcome X's death*. The idea is that the utility we assign to the reference pair attached to an outcome comes closest to what we might think of as the utility that attaches solely to the outcome; the reference pair attached to an outcome x is the outcome-act pair (x, f) for which the act f contributes least to the utility of the pair.

Now this is exactly analogous to what the redescription strategy proposes as an analysis of risk-sensitive behaviour. In that case, when you wish to calculate the utility of an outcome-act pair (x, f), you begin with the utility you attach to $(x, \overline{x})$. Then you weight that utility in a certain way that depends on the riskiness of the act. This gives the utility of (x, f). Thus, if we take $(x, \overline{x})$ to be the reference pair attached to the outcome x, then this is analogous to the moral case above.

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

In both cases, we can recover something close to the notion of utility for ultimate ends or pure outcomes (i.e. elements of $\mathcal{X}$): the utility of the pure outcome x – to the extent that such a utility can be meaningfully said to exist – is $u^*(x, \overline{x})$, the utility of the reference pair attached to x. That seems right. Strictly speaking, there is little sense to asking an agent for the utility they assign to a particular person's death; one must specify whether or not the death is the result of that agent's agency. But we often do give a utility to that sort of outcome; and when we do, I submit, we give the utility of the reference pair. Similarly, we often speak as if we assign a utility to receiving £50, even though the request makes little sense without specifying the act that gives rise to that pure outcome: again, when we do so, what we really do is give the utility of £50 *for sure*, that is, the utility of $(£50, \overline{£50})$.

Understood in this way, the analysis of a decision given by the redescription strategy still portrays the agent as engaged in means-end reasoning. Of course, there are no pure ultimate ends to which we assign utilities. But there is something that plays that role, namely, reference pairs. An agent's utility for an outcome-act pair (x, f) is calculated in terms of her utility for the relevant reference pair, namely, $(x, \overline{x})$; and the agent's value for an act f is calculated in terms of her utilities for each outcome-act pair (x, f) where x is a possible outcome of f. Thus, though the value of an act on this account is not ultimately grounded in the utilities of pure, ultimate outcomes of that act, it is grounded in the closest thing that makes sense, namely, the utilities of the reference pairs attached to the pure, ultimate outcomes of the act.

6. Conclusion

Buchak proposes a novel decision theory. It is formulated in terms of an agent's probability function on $\mathcal{F}$, utility function on $\mathcal{X}$ and risk function. It permits a great many more preference orderings than orthodox expected utility theory. On Buchak's theory, the utility that is assigned to an act is not the expectation of the utility of its outcome; rather it is the risk-weighted expectation. But the argument of Section 2 of this paper suggests that the value of an act for an agent should be her estimate of the utility of its outcome; and her estimate of a quantity should be her expectation of that quantity. And these, together, give the EU Rule of Combination. In this paper, we have tried to reconcile the preferences that Buchak endorses with the EU Rule of Combination. To do this, we redescribed the outcome space so that utilities were attached ultimately to outcome-act pairs rather than to outcomes themselves. This allowed us to capture precisely the preferences that Buchak permits, whilst letting the utility of an act be the expectation of the utility it will produce. The redescription strategy raises some questions: Does it prevent us from using decision theory for certain epistemological purposes? Does it fail to portray agents as engaged in means-end reasoning? In Section 5, we tried to answer these questions.

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

Notes

1. A finite set X of subsets of a set S is an algebra if (i) S is in X; (ii) if Z is in X, then its complement $S - Z$ is in X; (iii) if Z_1, Z_2 are in X, then their union $Z_1 \cup Z_2$ is in X.
2. The names should be considered labels only. I do not take them to imply that one sort of attitude can be observed directly, while the other sort is knowable only by inference.
3. As we will see below, one of Buchak's central contentions is that there is a third type of internal attitude with which decision theory deals, namely, attitudes to risk. In my alternative to Buchak's theory, I will incorporate such attitudes into the utilities on the outcomes. So, while these internal attitudes to risk will be present in my account, they will be a component of the utilities, not separate attitudes.
4. A technical note on the definition of Bregman divergences; what follows is not essential to the rest of the argument. Suppose C is a closed, convex subset of the real numbers. And suppose $\varphi: C \to \mathbb{R}$ is a continuously differentiable and strictly convex function. Then the Bregman divergence generated by φ is defined as follows: $\mathfrak{d}_\varphi(x, y) := \varphi(x) - \varphi(y) - \varphi'(y)(x - y)$. That is, $\mathfrak{d}_\varphi(x, y)$ is the difference between the value of φ at x and the value at x of the tangent to φ taken at y. $\mathfrak{q}$ is the Bregman divergence generated by $\varphi(x) = x^2$.
5. See also (D'Agostino and Dardanoni 2009), where the original mathematical result is stated and proved.
6. Recall: like a set, a multiset is unordered, so that $\{\{1, 2\}\} = \{\{2, 1\}\}$. Unlike a set, it allows repetitions, so that $\{\{1, 1, 2\}\} \neq \{\{1, 2, 2\}\}$.
7. Note that Buchak (2013, Section 4.4) considers a redescription strategy that is very close to the one I describe in this section. However, she notes that it is ill-defined. The strategy that I describe here does not suffer from this problem.
8. In general, for $\alpha_1, \ldots, \alpha_n, r \in \mathbb{R}$, there are many sequences $0 \leq \lambda_1, \ldots, \lambda_n$ with $\sum_i \lambda_i = 1$ such that $\sum_i \lambda_i \alpha_i = r$, if there are any.
9. If C is a finite set of vectors in a vector space V over the real numbers, the *convex hull* of C is written C^+ and defined as follows: C^+ is the smallest convex set that includes C, where a set is convex if it contains every mixture of two vectors whenever it contains those vectors; alternatively,

$$C^+ := \left\{ \sum_{c \in C} \lambda_c c : 0 \leq \lambda \leq 1 \;\&\; \sum_{c \in C} \lambda_c = 1 \right\}$$

Notes on contributor

Richard Pettigrew is a professor of Philosophy at University of Bristol, UK. Over the years, his research interests have migrated from mathematical logic through philosophy of mathematics and logic to rational choice theory and formal epistemology. He has published articles in Philosophical Review, Noûs, Philosophy and Phenomenological Research and Philosophy of Science, among others. His first book, Accuracy and the Laws of Credence, will appear in April 2016 with Oxford University Press. He is on the editorial board of Ergo, Philosophia Mathematica, and the Stanford Encyclopedia of Philosophy.

Funding

This work was supported by European Research Council (ERC) [grant number 308961-EUT].References

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

Allais, M. 1953. "Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école Américaine." *Econometrica* 21 (4): 503–546.
Buchak, L. 2013.*Risk and Rationality*. Oxford: Oxford University Press.
D'Agostino, M., and V. Dardanoni. 2009. "What's so Special about Euclidean Distance? A Characterization with Applications to Mobility and Spatial Voting." *Social Choice and Welfare* 33 (2): 211–233.
D'Agostino, M., and C. Sinigaglia. 2010. "Epistemic Accuracy and Subjective Probability." In *EPSA Epistemology and Methodology of Science: Launch of the European Philosophy of Science Association*, edited by M. Suárez, M. Dorato, and M. Rédei, 95–105. Dordrecht: Springer.
Eriksson, L., and A. Hájek. 2007. "What are Degrees of Belief?" *Studia Logica* 86 (2): 183–213.
de Finetti, B. 1974. *Theory of Probability*. vol. I. New York: Wiley.
Jeffrey, R. 1986. "Probabilism and Induction." *Topoi* 5: 51–58.
Joyce, J. M. 1998. "A Nonpragmatic Vindication of Probabilism." *Philosophy of Science* 65 (4): 575–603.
Leitgeb, H., and R. Pettigrew. 2010. "An Objective Justification of Bayesianism I: Measuring Inaccuracy." *Philosophy of Science* 77: 201–235.
Meacham, C. J. G., and J. Weisberg. 2011. "Representation Theorems and the Foundations of Decision Theory." *Australasian Journal of Philosophy* 89 (4): 641–63.
Pettigrew, R. (ta). 2016. *Accuracy and the Laws of Credence*. Oxford: Oxford University Press.
Quiggin, J. 1982. "A Theory of Anticipated Utility." *Journal of Economic Behavior and Organization* 3: 323–343.
Quiggin, J. 1993. *Generalized Expected Utility Theory: The Rank-Dependent Model*. Dordrecht: Kluwer Academic Publishers.
Schmeidler, D. 1989. "Subjective Probability and Expected Utility without Additivity." *Econometrica* 57 (3): 571–587.
Wakker, P. P. 2010. *Prospect Theory: For Risk and Ambiguity*. Cambridge: Cambridge University Press.
Williams, B., and J. J. C. Smart. 1973. *Utilitarianism: For and Against*. Cambridge: Cambridge University Press.
Zynda, L. 2000. "Representation Theorems and Realism about Degrees of Belief." *Philosophy of Science* 67 (1): 45–69.

Appendix 1

Proof of theorem 1

In this appendix, we prove Theorem 1. We begin by giving a geometric characterization of the pairs (c, e), where c is a credence function and e is an estimate function, such that c is probabilistic and e is expectational relative to c.

Lemma 5 *Suppose c is a credence function defined on $\mathcal{F}$ and e is an estimate function defined on $\mathcal{X}$. Then the following two propositions are equivalent:*

(i) $\mathbf{c}$ *is probabilistic and* $\mathbf{e}$ *is expectational with respect to* $\mathbf{c}$.

(ii) For each state s, *there is* $0 \leq \lambda_s \leq 1$ *such that* $\sum_{s \in S} \lambda_s = 1$ *and*

(a) $c(A) = \sum_{s \in S} \lambda_s A(s)$, *for each proposition* A *in* $\mathcal{F}$;

(b) $e(X) = \sum_{s \in S} \lambda_s X(s)$, *for each quantity* X *in* $\mathcal{X}$.

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

Proof 1 First, we prove (ii) ⇒ (i). Suppose (ii). First, we show that c is probabilistic. Recall that there are three conditions on being probabilistic: Range, Normalization, Additivity. We take them each in turn.

- Range: Suppose A is in $\mathcal{F}$. Then, note that: (1) each λ_s lies between 0 and 1 inclusive; (2) all of the λ_ss summed together give 1; (3) $A(s) = 0$ or 1 for each s in $\mathcal{S}$. Thus, it is certainly true that $\sum_{s\in\mathcal{S}} \lambda_s A(s)$ lies between 0 and 1 inclusive.
- Normalization: Since $\top$ is true at all states of the world, $\top(s) = 1$ for all s in $\mathcal{S}$, so $c(\top) = \sum_{s\in\mathcal{S}} \lambda_s \top(s) = \sum_{s\in\mathcal{S}} \lambda_s = 1$.
- Additivity: If there are no states s at which both A and B are true, then $c(A \vee B) = \sum_{s\in\mathcal{S}} \lambda_s (A \vee B)(s) = \sum_{s\in A\vee B} \lambda_s = \sum_{s\in A} \lambda_s + \sum_{s\in B} \lambda_s = \sum_{s\in\mathcal{S}} \lambda_s A(s) + \sum_{s\in\mathcal{S}} \lambda_s B(s) = c(A) + c(B)$ .Next, we show that e is expectational with respect to c. Suppose s' is a state of the world. Then note that $c(s) = \sum_{s\in\mathcal{S}} \lambda_s s'(s)$. But of course, since the states of the world form a partition, $s'(s) = 0$ if $s' \neq s$ and $s'(s) = 1$ if $s = s'$. Thus, $c(s) = \lambda_s$. Thus, $e(X) = \sum_{s\in\mathcal{S}} \lambda_s X(s) = \sum_{s\in\mathcal{S}} c(s)X(s)$, as required. This gives Expectation.

Second, we prove (i) ⇒ (ii). Let $\lambda_s = c(s)$ and the result follows easily from Additivity and Expectation. □

The upshot of this result is the following: Suppose $\mathcal{F} = \{A_1, \ldots, A_m\}$ and $\mathcal{X} = \{X_1, \ldots, X_n\}$. And, if c is a credence function on $\mathcal{F}$ and e is an estimate function on $\mathcal{X}$, represent the pair (c, e) by the following vector in $\mathbb{R}^{m+n}$:

$$\vec{ce} := (c(A_1), \ldots, c(A_m), e(X_1), \ldots, e(X_n))$$

And represent a state of the world s by the following vector in $\mathbb{R}^{m+n}$:

$$\vec{s} := (A_1(s), \ldots, A_m(s), X_1(s), \ldots, X_n(s))$$

Then Lemma 5 says that (c, e) is probabilistic and expectational iff $\vec{ce}$ lies in the convex hull of the vectors $\vec{s}$ for s in $\mathcal{S}$ – that is, $\vec{ce} \in \{\vec{s} : s \in \mathcal{S}\}^+$.[9]

The second lemma that we require to prove Theorem 1 is a geometric fact about the following measure of distance between two vectors in a real-valued vector space. If $\mathbf{x} = (x_1, \ldots, x_k)$ and $\mathbf{y} = (y_1, \ldots, y_k)$ are vectors in $\mathbb{k}^n$, then let

$$\mathfrak{Q}(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{k} |x_i - y_i|^2$$

Thus, clearly,

$$\mathfrak{F}(c, s) + \mathfrak{F}(e, s) = \mathfrak{Q}(\vec{ce}, \vec{s})$$

for any credence function c, estimate function e and state of the world s.

Lemma 6 *Suppose $D \subseteq \mathbb{R}^k$. Then*

(i) If $\mathbf{x} \notin D^+$, then there is $\mathbf{y} \in D^+$ such that $\mathfrak{Q}(\mathbf{d}, \mathbf{y}) < \mathfrak{Q}(\mathbf{d}, \mathbf{x})$, for all $\mathbf{d} \in D$.
(ii) If $\mathbf{x} \in D^+$, then there is no $\mathbf{x} \neq \mathbf{y} \in \mathbb{R}^k$ such that $\mathfrak{Q}(\mathbf{d}, \mathbf{y}) \leq \mathfrak{Q}(\mathbf{d}, \mathbf{x})$, for all $\mathbf{d} \in D$.

I won't provide a full proof of these geometric facts – proofs can be found in any geometry textbook. But here is a brief sketch. (i) is an easy consequence of the Hilbert Projection

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press

Theorem, since $\mathfrak{Q}$ is the square of the Euclidean metric. (ii) is a consequence of the fact that, if we measure distance between vectors using the Euclidean metric or its square, $\mathfrak{Q}$, then, for any two vectors, the set of vectors that are closer to the first than to the second is a convex set.

Putting these two results together and, in Lemma 6, letting $D = \{\vec{s} : s \in \mathcal{S}\}$, Theorem 1 follows.□

https://doi.org/10.1080/00455091.2015.1119610 Published online by Cambridge University Press