

What Is Risk Aversion?

H. Orri Stefánsson and Richard Bradley

ABSTRACT

According to the orthodox treatment of risk preferences in decision theory, they are to be explained in terms of the agent's desires about concrete outcomes. The orthodoxy has been criticized both for conflating two types of attitudes and for committing agents to attitudes that do not seem rationally required. To avoid these problems, it has been suggested that an agent's attitudes to risk should be captured by a risk function that is independent of her utility and probability functions. The main problem with that approach is that it suggests that attitudes to risk are wholly distinct from people's (non-instrumental) desires. To overcome this problem, we develop a framework where an agent's utility function is defined over chance propositions (that is, propositions describing objective probability distributions) as well as ordinary (non-chance) ones, and argue that one should explain different risk attitudes in terms of different forms of the utility function over such propositions.

- 1 *Introduction*
 - 2 *Risk Attitudes in the von Neumann–Morgenstern Framework*
 - 2.1 *Conceptual challenges*
 - 2.2 *Empirical challenges*
 - 3 *Risk-Weighted Expected Utility Theory*
 - 3.1 *Risk-weighted expected utility versus expected utility*
 - 3.2 *Problems with risk-weighted expected utility theory*
 - 4 *Risk Attitudes in the Jeffrey Framework*
 - 4.1 *Linearity, chance neutrality, and risk aversion*
 - 4.2 *Distinguishing risk attitudes*
 - 4.3 *Ambiguity and the four-fold pattern*
 - 5 *Conclusion*
-

1 Introduction

In colloquial talk, someone is said to be risk averse if they are disinclined to pursue actions that have a non-negligible chance of resulting in a loss or whose benefits are not guaranteed. This disinclination can be spelled out in a number

of different ways. Our starting point will be the one that prevails in much of the literature in economics and decision theory, namely that to be risk averse is to prefer any action A to another with the same expected objective—for instance, monetary—benefit but with greater variance (canonically called ‘a mean-preserving spread’ of A). More precisely, let G be any real-valued class of good (for example, money) and consider lotteries that yield quantities g of G with different probabilities. Then an agent is said to be risk averse with respect to G just in case, for all quantities g , she prefers g for sure to a (non-trivial) lottery with expectation g . For instance, someone who is risk averse with respect to money will disprefer a gamble yielding either \$0 or \$100 with equal probability to getting \$50 for sure.

In the orthodox treatment of risk preferences that prevails in both economics and decision theory, this idea is typically formalized using the expected utility (EU) framework of John von Neumann and Oskar Morgenstern ([2004]) (hereafter vN–M), though a similar treatment can be given in other frameworks. Within this framework someone who is risk averse with respect to money, say, must, in virtue of the way in which utility is cardinalized, have a concave utility function for money—that is, she must assign diminishing marginal utility to quantities of money—and *vice versa*. So the orthodoxy identifies risk aversion with respect to some good G with a particular property of the agent’s desires about quantities of G , as captured by the shape of her utility function on such quantities.

This treatment of risk attitudes has been challenged on two different, if related, grounds. First, it has been extensively criticized for failing to distinguish desire attitudes to concrete goods from attitudes to risk itself. Many people feel that because of this failure, the orthodoxy just does not capture the phenomenology of risk attitudes (see, for example, Watkins [1977]; Hansson [1988]). Second, there is a now a large body of empirical evidence suggesting that people exhibit attitudes to risk that cannot be explained within this framework but which are not obviously irrational; most famously in the paradoxes of Allais ([1953]) and Ellsberg ([1961]).

In response to these problems with the orthodox treatment of risk, there has been a recent trend towards introducing a special function—a risk function—that, in addition to a probability and a utility function, is used to represent attitudes to risky prospects: most notably in cumulative prospect theory (see Tversky and Wakker [1995]; Wakker [2010]) and the rank-dependent utility theory that it draws on (for instance, Quiggin [1982]), and in the recent risk-weighted expected utility (REU) theory (Buchak [2013]). Since these theories account for risk attitudes in terms of the form of the risk function, they can accommodate the intuition that attitudes to risk itself should be distinguished from desire attitudes to concrete goods. Moreover, these theories are more permissive than the orthodoxy as to what counts as rational, and thus allow

for many of the intuitively rational preferences that the orthodox theory deems irrational. The introduction of a risk function to account for risk attitudes however raises the question of whether risk attitudes really are a special type of attitude, wholly distinct from non-instrumental desires, as these theories suggest. We shall argue that they are not.

The approach of this article will differ from this recent trend. We show that it is possible to cardinalize utility without making any assumptions about risk preferences, by extending Jeffrey's ([1965]) decision theory to domains containing chance propositions; that is, propositions about objective probability distributions over outcomes of one kind or another. This allows us to model intrinsic attitudes to risk in terms of the form of the agent's desirability function for chances of goods—thereby respecting the intuition that risk attitudes are a special kind of desire—and to show how such attitudes co-determine, with the agent's attitudes to the goods themselves, her preferences for risky prospects. In addition to better capturing the phenomenology of risk attitudes than either orthodox EU theory or its contemporary rivals, our framework differs from these theories in providing a unified explanation of the empirical evidence regarding risk attitudes and so-called 'ambiguity attitudes'.

It might be worth emphasizing from the start that the decision theories we discuss, and the one we offer, are all normative. That is, they are meant to formalize the preferences and choices of a rational agent, and to characterize how these preferences and choices are based on the agent's more basic attitudes (in particular, her desires and beliefs). However, our theory about what risk attitudes are—namely, that they are desires about how chances are distributed—is meant to apply to irrational as well as rational agents.

2 Risk Attitudes in the von Neumann–Morgenstern Framework

In the vN–M theory, the utilities of outcomes are determined by the lotteries that the agent is willing to accept. Suppose a person prefers A to B which is preferred to C , and take any set of lotteries with A and C as their possible outcomes (or 'prizes'). Then according to the vN–M theory, we can find the outcomes' relative utilities by figuring out what chance such a lottery L must confer on A for the agent to be indifferent between L and B . The basic idea is that your judgement about B , relative to A on the one hand and C on the other, can be measured by the riskiness of the lottery L involving A and C that you deem equally desirable as B . For instance, if you are indifferent between L and B when the chance that L confers on A is $3/4$, then B is three quarters of the way up the utility scale that has C at the bottom and A at the top. This information can be used to determine the expected utility of the lottery. If we, say, stipulate that $u(A) = 1$ and $u(C) = 0$, then $u(B) = u(L) = 3/4$. This corresponds to the expected utility of the lottery, since $1/4 \cdot 0 + 3/4 \cdot 1 = 3/4$.

It follows immediately from this way of constructing (cardinal) utilities that the agent must have utilities for lotteries that are linear in the probabilities of the prizes. So, if an agent strictly prefers \$5 to a gamble that will either result in a prize of \$0 or a prize \$10, each prize having a 0.5 chance, then we must, on the vN–M approach, account for this in terms of the agent’s attitudes towards the monetary amounts in question, as represented by her utility function. For, as we have seen, agents are assumed to evaluate lotteries by the expected utilities of their prizes, so the only value that we can adjust in order to account for this person’s risk aversion is her utility function over money (assuming that her evaluation of the probability of the prizes is in line with the chances). In particular, we account for this attitude by postulating that money (in the \$0–\$10 range) has decreasing marginal utility to the person, as represented by a concave utility function over money in this range. It is this feature—the assumption that utilities of lotteries are linear in their probabilities—that is at the centre of debate about the adequacy of the orthodox theory.

2.1 Conceptual challenges

A long-standing complaint against the vN–M approach is that it mischaracterizes attitudes to risk. Such attitudes, the complaint goes, cannot be explained in terms of attitudes to concrete outcomes. For it seems that two rational people might evaluate the possible outcomes of a bet in the same way and agree about their probabilities, but nevertheless differ in whether they accept the bet or not, since they have different views about what levels of risk are acceptable (Buchak [2013]). Similarly, Watkins ([1977]) insists that it is possible for individuals to evaluate monetary outcomes linearly but nevertheless, due to their gambling temperament, turn down bets with a positive monetary expectation. To return to the previous example, it seems a conceptual possibility that a person who has a very strong dislike for gambling would turn down the offer to trade *B* for the lottery *L* which has *A* and *C* as possible prizes, except in the special case when *L* is almost certain to result in *A*, even though she considers *B* much less desirable than *A*. In sum, it seems that people’s dislike for gambling does not, by itself, tell us much about how they value the gambles’ prizes.

The above type of criticism takes as its starting point agents who dislike risk and criticizes the vN–M approach for modelling their risk aversion in terms of their attitudes to concrete outcomes. But since the vN–M approach equates decreasing marginal utility with risk aversion, it can also be criticized for falsely implying that anyone with a concave utility function over some good is risk averse with respect to that good. Hansson ([1988]) for instance tells the story of a professional gambler who turns down an offer to trade a single copy of a book that he is fond of for an even chance gamble between receiving no

copy of the book and three copies. Having been schooled in the vN–M approach, a decision analyst concludes that the gambler must be risk averse. The gambler retorts that this is nonsense; being a professional gambler, he has habituated himself to being risk neutral. The reason he turns down the gamble, he says, is simply that the second and third copy is of almost no worth to him.¹ Since one copy of the book is of great value to him, the gamble he is being offered has an equal chance of resulting in a real loss (losing his single copy) and no real gain (getting two extra copies). And that is the reason he turns down the bet. It is of course possible that someone else might display the same preference between the single book and the gamble due to dislike of risk *per se*. But Hansson's gambler turns down the gamble because quantities of the book have decreasing marginal worth to him. These are psychologically distinct reasons that might give rise to the same pattern of preference, and should thus be kept distinct in formal models of practical reasoning.

Decision theorists often respond in one of two ways when confronted with objections like these. The first is to question whether people really can say how they evaluate concrete outcomes without consulting their preferences for risky prospects involving those outcomes; and, correspondingly, whether people really can judge their own risk aversion with respect to some good independently of their preferences between risky prospects involving those goods. But this response does not mitigate the worry that the vN–M approach conflates two distinct attitudes. For we do not need to determine by introspection the precise extent to which we desire outcomes and are willing to accept risk, to be able to justify replies like the gambler's to the decision analyst. Moreover, Hansson's story illustrates that, whether or not people can make these judgments by introspection, the vN–M approach equivocates two phenomena that conceptually and psychologically are very different: on one hand, the decreasing marginal worth of quantities of goods, on the other hand dislike for risk as such. In sum, whether or not people are able to introspect precisely how they evaluate concrete outcomes or what risk they are willing to accept, is orthogonal to the seemingly obvious point that these are two different types of attitude.

The second type of response that decision theorists and economists typically offer when faced with criticism like that raised above, is to resort to a formalistic interpretation of utility, according to which the role of EU theory is neither to explain nor to guide rational action, but simply to mathematically represent rational preferences or choices. If that is the aim, then as long as we can represent, say, risk-averse choice behaviour by postulating a concave

¹ To make the example particularly plausible, we can assume that the book is of great sentimental value to the gambler but does not have much market value, or, more generally, that the gambler knows that he won't get a price for the second and third copy that matches his evaluation of the first copy.

utility function, it does not matter that we are equivocating two conceptually distinct psychological attitudes. In other words, as long as, say, decreasing marginal utility is behaviourally indistinct from what we might call aversion to risk *per se*, it does not matter whether or not these are psychologically distinct, since the aim is simply to represent the choice behaviour (see, for instance, Harsanyi [1977]).

The formalistic interpretation has been extensively (and critically) discussed by several philosophers, and we will not add much to that discussion. Instead, we will simply sketch three problems with this second response.² First, the fact is that we often do want to be able to explain, rather than simply describe, behaviour in terms of the maximization of a utility function. In other words, we want to be able to say that a person chose an alternative because it was the alternative with highest expected utility according to her. Second, when using decision theory to make policy recommendations, as is commonly done—or, more generally, when using the theory for decision-making purposes—we need to assume that the utilities on which we base the recommendations exist prior to (and independently of) the choices that the theory recommends. Third, if we are trying to construct a formal theory of practical reasoning—that is, if we are trying to characterize how rational agents make choices rather than just mathematically representing choices they have already made—then our theory needs to distinguish the decreasing marginal utility of quantities of goods from aversion to risk as such.

These objections to the orthodox treatment of risk will not get much traction without a demonstration of how utility can be measured without presupposing the vN–M framework. Indeed unless utility can be determined independently of assumptions about the nature of risk preferences the objections of Watkins and Hansson can be regarded as literally meaningless. But there is a perfectly straightforward response to this worry: adopt one of the other frameworks for cardinalizing utility and test the claims of the vN–M theory within it. Various such frameworks are already to be found in the decision-theoretic literature, including the aforementioned cumulative prospect theory and risk-weighted EU theory. But, for reasons that we will explain later on, we consider these frameworks to have their own problems, both conceptual and empirical. So instead we will make use of Bayesian decision theory to cardinalize utility; in particular, an extended version of the variant developed by Jeffrey ([1965]) that has the virtue of not implying that the value of a lottery is linear in its probabilities.³

² See, for instance, (Broome [1991]; Dreier [1996]; Bermúdez [2009]; Buchak [2013]; List and Dietrich [2016]; Bradley [2017]; Okasha [2016]).

³ We call a decision theory ‘Bayesian’ if the probabilities that go into the expectation whose value rational agents maximize are the agent’s own subjective probabilities. Thus Savage’s ([1954])

Table 1. Four-fold pattern of risk preferences

Probability of outcome	Gains	Losses
Low probability	Risk seeking	Risk aversion
High probability	Risk aversion	Risk seeking

2.2 Empirical challenges

The orthodox treatment of risk attitudes faces two distinct types of challenges from the growing body of evidence regarding people's actual choices in situations of risk and uncertainty.⁴ First, there is much evidence that people exhibit risk attitudes in their choices that cannot be reconciled with EU theory and in particular that their preferences between risky prospects are not linear in the chances of the outcomes. In Kahneman and Tversky's ([1979]) famous study, for instance, they report the four-fold pattern of attitudes for simple lotteries of the form 'x chance of \$y' displayed in Table 1, obtained by determining the agents' dollar prices for these lotteries. The pattern of risk-averse behaviour when it comes to lotteries with high probability of monetary gains or low probability of losses, together with risk-seeking behaviour for lotteries with low probability of monetary gain or high probability of losses, cannot be reconciled with EU theory no matter what utility function is attributed to subjects. This has led most decision theorists to conclude that EU theory is not descriptively adequate as a theory of choice between risky prospects.

It is possible that the deviation from the predictions of the orthodox theory exhibited by this pattern of choice can typically be attributed to irrationality on the part of the deviating subjects. And indeed the focus of interest in the decision-theoretic literature has been on the implications of these results for descriptive decision theory, with little in the way of a consensus emerging on their normative implications. But some instances of these risk preferences do not seem irrational. The most famous example of this is the so-called Allais paradox, originally introduced by Allais ([1953]). The 'paradox' is generated by comparing people's preferences over two pairs of lotteries similar to those given in Table 2. The lotteries consist in tickets being randomly drawn, determining the prize of each lottery (for instance, lottery L_1 results in a prize of \$5 million if one of the tickets numbered 2–11 is drawn).

In this situation, many people strictly prefer L_2 over L_1 but also L_3 over L_4 , a pair of preferences we will call the Allais preference. According to the

and Jeffrey's ([1965]) decision theories are Bayesian, but von Neumann and Morgenstern's ([2004]) is not.

⁴ We will follow the convention of calling decision situations where the relevant outcomes have objective probabilities known to the decision-maker situations of risk, and we will call decision situations where the decision maker lacks such knowledge situations of uncertainty.

Table 2. Allais's paradox

Lottery	1	2–11	12–100
L_1	\$0	\$5m	\$1m
L_2	\$1m	\$1m	\$1m
L_3	\$0	\$5m	\$0
L_4	\$1m	\$1m	\$0

orthodox theory, this pattern of preference is irrational, since there is no way to assign utilities to the prizes on offer such that L_2 gets a higher EU than L_1 and L_3 gets a higher expected utility than L_4 . In other words, the Allais preference cannot be represented as maximizing expected utility (which according to the orthodox picture implies that it is irrational).⁵

Here is an explanation of why the reasoning underlying the Allais preference is inconsistent with EU theory. People with this preference find that the value of decreasing the risk of \$0 (from a baseline of \$1 million) from 0.01 to 0 exceeds the value of a 0.1 chance at \$5 million rather than \$1 million. And that is why they prefer L_2 to L_1 . However, the same marginal decrease in the risk of \$0 does not exceed the value of a 0.1 chance at \$5 million rather than \$1 million when the decrease is from 0.9 to 0.89. And that is why they prefer L_3 to L_4 . In other words, people typically consider a reduction in the risk of winning nothing from very unlikely to impossible to be more important than the same absolute decrease in the risk of winning nothing from quite likely to only slightly less likely. But in both cases, people are comparing a 0.01 chance of getting \$0 rather than \$1 million with a 0.1 chance of getting \$5 million rather than \$1 million. So what they are comparing on both occasions, according to vN–M's theory, is $0.01[u(\$1m) - u(\$0)]$ with $0.1[u(\$5m) - u(\$1m)]$ (where $u(x)$ denotes the utility of x). Hence, if this reduction in risk of ending up with \$0 is worth foregoing a 0.1 chance at \$5 million when comparing L_1 and L_2 , it should also, on the orthodox story, be worth it when comparing L_3 and L_4 .

The second type of empirical challenge that orthodox decision theory is faced with relates to choices that agents make in contexts characterized by both (subjective) uncertainty and (objective) risk. In particular, the orthodoxy is unable to account for the phenomenon of ambiguity aversion, a pattern of preference that is typically exhibited in the famous Ellsberg paradox, but which can be much more simply explained with the following example.⁶

⁵ Another way to see that the Allais preference violates vN–M's EU theory is to notice that it violates their independence axiom, which intuitively says that when comparing risky gambles, one should ignore what the gambles have in common.

⁶ The version of the paradox that we present assumes that questions about people's confidence can be distinguished from questions about their preferences for risky prospects. As a referee for *BJPS* points out, this assumption may seem question begging, since those who adhere to the formalistic interpretation of Bayesian decision theory—and behaviourists more generally—are

Suppose you have in front of you a coin, C_1 , that you know to be perfectly symmetric, and that you know to have been tossed a great number of times and has come up heads exactly as many times as it has come up tails. More generally, suppose that you possess the best possible evidence for the coin being unbiased. Here are two questions: (i) How confident are you that C_1 will come up heads on its next toss?⁷ (ii) How much would you be willing to pay for a bet that pays you \$10 if C_1 lands heads on its next toss but pays nothing otherwise?

Now suppose instead that you have in front of you a coin, C_2 , that you know to be either double headed or double tailed, but you don't know which. Here are again two analogous questions: (i') How confident are you that C_2 will come up heads on its next toss? (ii') How much would you be willing to pay for a bet that pays you \$10 if C_2 lands heads on its next toss but pays nothing otherwise?

Most people seem to use something like the principle of insufficient reason when answering questions like (i') (see, for instance, Voorhoeve *et al.* [2012]). Since they have no more reason for thinking that the coin will come up heads than tails, they are equally confident in these two possibilities. But since these possibilities exhaust the possibility space, they should believe to degree 0.5 that the second coin comes up heads. But that is, of course, the same degree of belief as they should have in the proposition that the first coin comes up heads (assuming something like Lewis's [1980] principal principle). There is, of course, an important difference between their judgements about the two coins, as we will discuss in more detail later on. In the first case, they are pretty certain that the coin has an (objective) chance of one-half of coming up heads on the next toss; in the latter case they are not. But in both cases, they are equally confident that the coin will come up heads on the next toss as that it will come up tails.

What about questions (ii) and (ii')? A number of experimental results on Ellsberg-type decision problems show that people tend to be what is called ambiguity averse, meaning that they prefer prospects with known chances of outcomes to ones with unknown chances (see, for example, Wakker [2010]). In the example under discussion, ambiguity aversion translates into a preference for a bet on C_1 over a bet on C_2 and hence a willingness to pay more for the first bet than the second.

Now the above answers may seem to create a problem for Bayesian decision theory. Since the possible prizes are the same for the two bets, standard applications of the theory imply that people should be willing to pay more for the

unlikely to accept it. So it is worth noting that the original choice problem described by Ellsberg ([1961]) does not require the assumption in question.

⁷ We will use 'confidence' and 'degree of belief' interchangeably.

first bet than the second only if they are more confident that they will get the prize (the \$10) if they accept the first bet than if they accept the second bet. But they are not: they are equally confident of getting the \$10 in both cases. In response, orthodox Bayesians might try to argue that people who are willing to pay more for the first bet than the second have made some mistake in their instrumental reasoning.⁸ But that seems implausible. Ambiguity aversion, even in these very simple set-ups, is a robust phenomenon, making it unlikely that people are simply making a mistake.

In fact however the above analysis ignores the important difference between the two cases. In the first case, a bet on heads amounts to accepting a lottery which confers a chance of one-half on the prize. In the second case a bet on heads is an action which yields the prize with chance one or with chance zero depending on whether the coin is two-headed or two-tailed. But this difference between the two cases is irrelevant if the vN–M theory is correct. The upshot is that ambiguity aversion, being a phenomenon that arises when both subjective and objective uncertainty is present, raises a challenge to the combination of the Bayesian theory of rational preference under uncertainty and the vN–M theory. Most of the literature on ambiguity aversion is based on the assumption that the vN–M theory is correct and hence draws the conclusion that ambiguity aversion is inconsistent with Bayesian rationality. We will take the contrary view, arguing that ambiguity aversion is a permissible attitude to spreads of chances that is perfectly consistent with the kind of Bayesian framework in which we will work, but inconsistent with the vN–M theory.

Although the challenges presented by the empirical evidence concerning attitudes to risk seem quite different from those presented by attitudes to ambiguity, we will see that the reason why these two types of attitudes generate trouble for orthodox EU theory is much the same: the theory's narrow conception of the (dis)value of risks and chances. As we show in Section 4, it is possible to account for both types of attitudes in a decision theory whose value function is defined over a set of chance propositions. Moreover, this makes it possible to represent an Allais-type preference and an ambiguity averse preference as maximizing the value of the same desirability function, which means that unlike previous treatments of these preferences, ours offers a unified explanation of the two (types of) empirical observations that have posed the

⁸ Alternatively, some might argue that the agents in question have made a mistake in their epistemic reasoning, by employing the principle of insufficient reason. (We thank a referee for *BJPS* for pointing out the need to respond to this objection.) In response, we contend (without having the space to really argue for our claim) that one is rationally permitted to apply the principle in this particular case. We do not claim, however, that one is rationally required to apply the principle, neither in this case, nor, of course, more generally. Moreover, it is worth keeping in mind that people might arrive at this confidence judgement without applying the principle in question. And that is all that is required to generate trouble for the orthodox application of Bayesian decision theory, as long as the individuals in question are willing to pay more for the first bet than the second.

greatest challenge to orthodox decision theory. But first, let us discuss a recent and quite influential alternative to EU theory, and explain why we think that this alternative does not do justice to ordinary risk attitudes either.

3 Risk-Weighted Expected Utility Theory

Recently several authors have constructed non-EU theories that introduce a risk function to represent people's risk attitudes, with the aim, first, to formally capture the intuition that risk attitudes with respect to some good need not be determined by how people evaluate quantities of that good; and, second, to make it possible to represent preferences like Allais's as maximizing agents' value functions. We will focus on a particularly well worked out and influential version of these theories, namely, Buchak's ([2013]) recent REU theory, but our argument equally applies to (normative versions of) the theories on which Buchak's theory is based (such as rank-dependent utility theory). The simplest way to explain REU theory is by comparing it to classical (Bayesian) EU theory. We do so in the next subsection, and raise two objections to REU theory in the subsection after that.

3.1 Risk-weighted expected utility versus expected utility

One of the main differences between REU theory and more traditional Bayesian decision theory concerns how many variables are determined by the agent we are trying to model, and, correspondingly, how many functions we use to represent her mental attitudes. The orthodox theory leaves it up to the agent to decide two things: first, the values of the possible consequences of the acts at the agent's disposal, as represented by her utility function; second, the probabilities of the different contingencies that determine which of these consequences are realized when each act is performed, as represented by her subjective probability function. In addition, REU theory leaves it up to the agents to decide how to aggregate the values of different possible outcomes of an alternative in order to evaluate the overall value of the alternative, and represents this by a risk function. The idea is that the form of an agent's aggregation will depend on how she trades-off chances of good outcomes against risks of bad outcomes. So whereas EU theory models rational agents as maximizing EU relative to a pair of utility and probability functions, REU theory models rational agents as maximizing REU relative to a triple of utility, probability, and risk functions.

To make the discussion that follows more precise, let r be a (non-decreasing) risk function on $[0, 1]$, satisfying the constraint that $0 \leq r(p) \leq 1$ and $r(0) = 0$, $r(1) = 1$. The function is intuitively to be understood as a weighing function on

probabilities, whose purpose is to discount or inflate, in accordance with the agent's attitudes to risk, the probability of attaining more than the minimum that an alternative guarantees. Let s_i denote a state of the world, and $f(s_i)$ the outcome of act f when state s_i is actual; u is a utility function on outcomes and P a probability function on states. Now the value of f , according to Buchak, is given by its risk-weighted expected utility, which, she argues, rational preferences maximize:

$$\text{REU}(f) = u(f(s_1)) + \sum_{j=2}^n [r(\sum_{i=j}^n P(s_i))(u(f(s_j)) - u(f(s_{j-1})))]. \quad (1)$$

REU theory can allow for the possibility that two individuals with the same beliefs and the same desires over risk-free outcomes differ in their evaluation of risky prospects. Suppose both Ann's and Bob's beliefs can be represented by the same probability function and that they evaluate monetary outcomes in the same way. Nevertheless, Ann is willing to pay up to and including \$5 for an even chance gamble that either results in her winning \$10 or nothing, whereas Bob is willing to pay at most \$3 for the same gamble. As we have seen, two (rational) people cannot differ in this way, according to orthodox EU theory: given the difference between Ann's and Bob's attitudes to these gambles, they must either have different beliefs or disagree about the relative values of the prizes on offer. In contrast, REU theory can account for the above difference between Ann and Bob's attitudes to gambles, without postulating different probability or utility functions, by assuming that Ann's risk function is linear ($r(p) = p$) while Bob's is convex (in particular, $r(0.5) < 0.5$).

REU theory can potentially account for the Allais preference in a similar way. Given a linearity assumption which Buchak ([2013], Footnote 39) implicitly makes—but which we will not—the Allais preference can be represented as maximizing REU whenever the risk function is convex, in particular, when the difference between $r(1)$ and $r(0.99)$ is greater than the difference between $r(0.11)$ and $r(0.10)$. The intuitive explanation for this, you may recall, is that a 0.01 probability difference counts more heavily in the agent's decision-making when it means that a prize becomes certain as opposed to almost certain, than when it means that a prize becomes only slightly more probable but still quite unlikely.

3.2 Problems with risk-weighted expected utility theory

In this section we raise two problems for REU theory. The first is that it cannot (nor is it meant to) account for the Ellsberg preferences. Our simplified version of the Ellsberg paradox suffices to illustrate this. Recall that we are assuming that a person is equally confident that C_1 and C_2 will come up heads on their next tosses; she is offered a bet on each coin that pays her \$10 if it

comes up heads on the next toss but nothing otherwise; and she is willing to pay a higher price for the bet on the first coin than the second, since only in the former case is she confident that the coin is unbiased. The REU of both bets is equal to $u(\$0) + r(0.5)(u(\$10) - u(\$0))$. Hence, REU theory cannot make sense of the willingness to pay more for one of these bets than the other.^{9,10}

The second problem with REU theory is that even in those cases where it is consistent with risk-averse preferences, such as Allais's, we think it mischaracterizes the psychology of risk attitudes. This is an especially grave problem for a theory like Buchak's, whose benefit compared to orthodox EU theory is partly meant to be that it better fits the phenomenology of risk attitudes. The problem consists in the fact that risk attitudes are, according to REU theory, primitive mental attitudes, distinct from both desires and beliefs. The risk function, r , is logically independent of both the probability function, P , and, more importantly for the present argument, also independent of the utility function, u . This is not an accident. As Buchak ([2013], pp. 53–4) explains:

The utility function is supposed to represent desire [...] and the probability function is supposed to represent belief [...] We try to make beliefs 'fit the world', and we try to make the world fit our desires. But the risk function is neither of these things: it does not quantify how we see the world—it does not, for example, measure the strength of an agent's belief that things will go well or poorly for him—and it does not describe how we would like the world to be. It is not a belief about how much risk one should tolerate, nor is it a desire for more or less risk. The risk function corresponds neither to beliefs nor desires. Instead, it measures how an agent structures the potential realizations of some of his aims.

We are happy to grant Buchak the claim that the attitudes that the risk function is meant to represent are not beliefs. But we find it hard to understand her

⁹ Another problem with EU theory that REU theory cannot solve, unlike the theory developed in the next section, is the Diamond ([1967]) 'paradox', which is based on EU theory's inability to account for the intuition that sometimes it is valuable to give people a chance at a good even if they do not end up receiving the good. See (Stefánsson and Bradley [2015]) for an explanation of how the framework discussed in the next section is partly motivated by the problem raised by Diamond, and Stefánsson ([2015], Section 4) for a demonstration of REU theory's inability to solve it.

¹⁰ As a referee for *BJPS* points out, some might find it to be a strength rather than weakness of Buchak's theory that it cannot account for the Ellsberg preference. In particular, some may find it to be a benefit of her theory that it does not provide a unified account of the paradoxes of Allais and Ellsberg. This might be either because people have pre-theoretical intuitions about the Ellsberg preference being irrational and different in nature from the rationally permissible Allais preference, or because Buchak's theory has convinced them of the need to treat these preferences differently. We hope to undermine both reasons for treating these two preferences differently, by making the case that they can both be rationalized, as features of people's attitudes to risks and chances, by a theory that has greater plausibility than Buchak's independently of how the two theories treat these preferences (that is, due to the second problem with REU theory discussed above).

view that these risk attitudes are not a special kind of desire, especially if we accept her (quite standard) characterization of desire as the type of attitude to which we try to make the world fit.¹¹ Recall Bob, who values money linearly but is nevertheless risk averse; for instance, prefers \$5 for sure to a gamble whose monetary expectation is \$5. Surely, any risk function that reflects this fact about Bob partly describes how he would like the world to be. In particular, he would rather like the world to be such that he holds \$5 than a bet with the same monetary expectation, and his risk function reflects this wish. Moreover, Bob will, if instrumentally rational, try what he can to make the world fit this attitude, for instance, by not accepting certain bets and by hedging the risks he exposed to. So, risk attitudes are attitudes to which we try to fit the world.

More generally, people who are risk averse have different views than the risk neutral (or risk loving) about how outcomes should be distributed across the possibility space, as Buchak ([2013], p. 29) herself points out. Informally put, risk-averse people prefer goods (including chances) to be spread evenly over the possibility space, such that they are guaranteed to get something that is not too bad no matter what the world happens to be like. Risk-loving people, on the other hand, prefer goods to be more concentrated, such that if a state favourable to them turns out to be actual, they get lots of the good in question. Again, these different attitudes will manifest themselves in different ways of trying to arrange the world. For instance, someone who is rational and risk averse with respect to some good will try to arrange the world such that quantities of that good are evenly spread across the possible states of the world (for instance, by hedging their bets), but the rational and risk loving with respect to some good will try to have quantities of that good more concentrated in fewer states (for instance, by accepting risky bets).

Buchak ([2013], p. 29) emphasizes in various places that she takes the risk function to represent how an agent ‘structures the potential realizations of some of his aims’, which she takes to be incompatible with seeing the risk function as representing part of an agent’s desires. But it is unclear why these are incompatible interpretations. For it seems that the risk function represents how an agent would want to (and will if she can) structure the realization of her aims. Other things being equal, risk-averse people will want to, and will try to, realize their aims as safely as they can, even if that reduces the expected realization of their aims. For instance, they will, other things being equal, want to (and try to) structure the gambles they hold in such a way to spread goods equally rather than unequally across the possibility space. Doesn’t this mean that they desire their gambles and other prospects to be structured in this way rather than in a more risky way?

¹¹ Some of these arguments against Buchak’s view can be found in (Stefánsson [2014]).

In sum, it seems to us that attitudes to risk are simply a special kind of desire, rather than a primitive mental attitude on a par with beliefs and desires. So, we should not account for such attitudes by introducing a function that is logically independent of the (utility) function that represents a person's desires. However, risk attitudes are not desires about concrete outcomes, as already discussed. So in that respect we agree with Buchak's criticism of orthodox EU theory. Instead, they are desires about chance distributions. That is, risk-averse people want chances to be distributed one way, risk-neutral and risk-loving people in other ways. In the next section we will make this suggestion more precise, by presenting a decision-theoretic framework where people's value functions are partly defined over propositions about chances. We will show how this framework, first, respects the intuition that risk attitudes are not about concrete outcomes but are still a special kind of desire, and, second, makes it possible to represent both Ellsberg- and Allais-type preferences as maximizing the expectation of agents' value functions.

4 Risk Attitudes in the Jeffrey Framework

Our aim in this section is to present a framework in which preferences for risky prospects can be cardinalized with only minimal assumptions about the properties of such preferences and in particular without assuming that they are linear in chances. To do so we build on Jeffrey's ([1965]) version of Bayesian decision theory. His theory has two great advantages in this context, compared to the rival Bayesian theory of Savage ([1954]). First, since the objects of desire in Jeffrey's theory are propositions, it is considerably more natural to extend the theory to allow for conative attitudes to chance distributions than to similarly extend Savage's theory, where the objects of desire are interpreted as acts and formally modelled as functions from the state space to the set of consequences. Second, Jeffrey's theory depends on much weaker assumptions than Savage's, which enables us to evaluate proposals for constraints on rational attitudes to risks and chances without taking too much for granted.

Before formally introducing Jeffrey's theory, and our extension of it, let us try to explain informally how our framework solves the problems we have been discussing. First, by extending the set on which Jeffrey's desirability function is defined to propositions describing chance distributions, we make room for the possibility that rational agents can take conative attitudes to chances that differ from the way in which orthodox EU theory assumes that such agents evaluate chances. In particular, we make room for the possibility that people can like or dislike the chances of obtaining a concrete good, relatively independently of how they evaluate quantities of these goods. And we

model different risk attitudes by different forms of the desirability function over such chance propositions, thereby formalizing our view that risk attitudes are a special type of desire.

Second, we show that by extending the desirability function to chance propositions, we can account for both ambiguity and risk attitudes in terms of the form of this function. In particular, we show that the same desirability function over chance propositions can account for both ambiguity aversion and the aforementioned four-fold pattern of risk attitudes, which means that ours is the first model that can simultaneously make sense of the two types of preference patterns that have, historically, created the biggest challenges for orthodox EU theory.

4.1 Linearity, chance neutrality, and risk aversion

The aim of this section is to formally introduce our framework, and explain how it differs from orthodox EU theory. In Jeffrey's theory, which forms the basis of ours, the degrees of belief of a rational agent are measured by a subjective probability function, P , on a Boolean algebra of propositions, Ω . Her degrees of desire are measured by a corresponding desirability function, V , defined on the same algebra but with the logically contradictory proposition \perp removed and satisfying for all $A, B \in \Omega - \{\perp\}$:

Desirability: If $A \wedge B = \perp$, and $P(A \vee B) \neq 0$, then:

$$V(A \vee B) = \frac{V(A)P(A) + V(B)P(B)}{P(A \vee B)}.$$

Necessary and sufficient conditions for preferences to be represented by such a pair of functions, P and V , were established by Bolker ([1966]). None of these conditions have anything special to say about preferences for risky prospects. Indeed there are no lotteries in the basic Jeffrey–Bolker framework, so it is not possible to model risk preferences within it (let alone cardinalize utility on the basis of them). This is a limitation that we now need to address. We do so by explicitly introducing propositions about chances and then identifying lotteries with conjunctions of such propositions.

Let Z be a Boolean subalgebra of the background Boolean algebra Ω . Intuitively, Z contains those propositions to which it is meaningful to ascribe chances.¹² Let $\Pi = \{ch\}$ be the set of all probability functions on Z and let $\Delta = \wp(\Pi)$ be the set of all subsets of Π . The elements of Δ serve here as what we will call chance propositions. In particular, for any $X \in Z$, and $x \in [0, 1]$, let $Ch(X) = x$ denote the chance proposition defined by $\{ch \in \Pi : ch(X) = x\}$.¹³

¹² Canonically we take the base propositions to be sets of possible worlds, but nothing hangs on this particular treatment of them.

Intuitively, $Ch(X) = x$ is the proposition that the chance of X is x (and the chance of $\neg X$ is $1 - x$). To construct propositions corresponding to the lotteries that are the basic objects of choice in the orthodox (vN–M) theory of decision-making under risk, let $\mathcal{X} = \{X_1, \dots, X_n\}$ be an n -fold partition of Z , with the $X_i \in \mathcal{X}$ being the prospects that constitute the various possible ‘prizes’ of a lottery or, more generally, outcomes of some stochastic process. Let $\bigcap_{i=1}^n (Ch(X_i) = x_i)$ denote the conjunction of the corresponding n propositions $Ch(X_1) = x_1, Ch(X_2) = x_2, [\dots]$, and $Ch(X_n) = x_n$, where the x_i are such that $\sum_{i=1}^n x_i = 1$. A proposition $\bigcap_{i=1}^n (Ch(X_i) = x_i)$ expresses the chances of realizing each of the X_i , thereby serving as the propositional equivalent, in this framework, of a lottery over the X_i .

The focus of our interest is the product set $\Gamma = Z \times \Delta$ whose elements are combinations of factual and chance propositions. Since Γ forms a Boolean algebra we can simply apply Bolker’s theorem to establish, for preferences over the propositions in $\Gamma - \{\perp\}$ satisfying the Bolker axioms, the existence of probability function P and desirability function V , respectively, on Γ and $\Gamma - \{\perp\}$, measuring the agent’s degrees of belief in, and desire for, the propositions contained in these sets; including, of course, propositions concerning chances. We thus have a framework in which it is meaningful to ask what the relationship is between agents’ attitudes to concrete goods and their attitudes to chances of such goods, including lotteries over them. And in particular whether or not agent’s preferences for lotteries must generally satisfy the requirements of the vN–M theory or not.

We can give an immediate answer to the latter question. As we have seen, the vN–M theory postulates that agents’ utilities for lotteries are linear in the chances. This is captured in our framework by the following condition on the desirability of chance propositions:

Linearity: For any n -fold partition \mathcal{X} of Γ and set $\{x_i\}$ such that $x_i \in [0, 1]$ and $\sum_{i=1}^n x_i = 1$:

$$V\left(\bigcap_{i=1}^n (Ch(X_i) = x_i)\right) = \sum_i x_i \cdot V(X_i).$$

Linearity says that the desirability of any lottery is a sum of the desirabilities of the lottery’s prizes weighted by the chances accorded to them by the lottery.

¹³ Strictly speaking, each chance function, and corresponding chance proposition, should be time-indexed (as discussed by Stefánsson and Bradley [2015], pp. 613–14). This is important to keep in mind when interpreting these propositions, since it is often the case that the desirability of a particular chance distribution for some outcome differs depending on when the distribution holds. For instance, the desirability of Donald Trump having, say, a 30% chance of winning the presidential election presumably depends on how close to the election it is true that he has that particular chance. However, since we will not, in this article, discuss examples where the chance of some outcome changes, we can safely ignore the time index for now, to simplify our notation. (We thank a referee for *BJPS* for reminding us of the need to address this issue.)

Informally, we can think of it as capturing the idea that chances do not matter intrinsically to the agent; they matter only instrumentally, as means to the attainment of the prizes that they are chances of. More exactly, as shown in Stefánsson and Bradley ([2015], Theorem 1), linearity encodes the neutrality of chances, the idea that one should not care about the chance of a (maximally specific) outcome once one knows whether or not the outcome has been realized. Formally (and adopting the convention that the status quo has desirability zero):

$$\text{Chance Neutrality: } V(\text{Ch}(X) = x|X) = 0.$$

So the question of whether or not an agent's risk preferences within the extended framework must conform with the vN–M theory boils down to that of whether it is rationally permissible or not to attach any (dis)value to chances over and above the extent to which they make various goods of which they are the chances of more or less likely. We will not discuss this normative issue in any detail, since it has been argued at length elsewhere (see Stefánsson and Bradley [2015]) that linearity and chance neutrality are not requirements of rationality, on the standard decision-theoretic conception of rationality as consistency. But to explain our view briefly, we think that chance neutrality makes a substantial value claim and is not a mere consistency condition on desire. In particular, the claim that one cannot rationally care about the chance of an outcome once it obtains, is a claim about what one can value, rather than a claim about what relationship must hold between one's values. Moreover, the demands imposed by linearity are too stringent, in our view, to be considered general requirements of rationality. In addition to condemning preferences like Allais's and Ellsberg's, the principle entails that relatively modest risk aversion when it comes to small stakes is only consistent with what seems to be absurd levels of risk aversion for larger stakes. For instance, linearity means that a person who turns down, when her total wealth is less than \$300,000, a 50:50 gamble that results in her either losing \$100 or winning \$125, must, when her wealth is \$290,000, turn down a 50:50 gamble that results in her either losing \$600 or winning \$36 billion (Rabin [2000]).

Let us now return to the opening observation that dislike of risk *per se*, rational or otherwise, is psychologically very different from the decreasing marginal desirability of quantities of concrete goods, even though the two phenomena may give rise to the same choice behaviour. Can we make sense of this observation within our framework? To do so we must show how we can distinguish the two kinds of attitudes and then show how these attitudes relate to risk-averse behaviour, that is, a preference for a lottery over any mean-preserving spread of it. We will consider each task in turn.

4.2 Distinguishing risk attitudes

Our central thesis is that an agent's like or dislike of risk involving a good is captured by the properties of her desirability function for chances of this good. Consider a lottery that pays \$100 with probability one-half and nothing otherwise and an agent whose desirabilities for modest amounts of money are linear in quantities of it. If this agent attaches a desirability to the half-chance of winning the \$100 equal to the desirability she attaches to winning \$50, then we can say that she is neutral with regard to the risk of (not) winning the \$100. But if the desirability of the half-chance of winning the \$100 is less than that of \$50, then she must value the half chance of \$100 at less than half the value of the \$100. So in this case the agent will display behavioural risk aversion (she will prefer the \$50 to the lottery), not because of her attitude to quantities of money (which were assumed to be linear), but because of her attitude to the risk *per se* of (not) winning the \$100. So unlike the vN–M theory, our framework can distinguish between these two quite different motivations for the same choice behaviour.

In general, in our framework, an agent's preferences amongst lotteries will depend not just on her desirability function for quantities of the good at stake, but also on her desirability function for chances of these goods (indeed on the relationship between the two). To substantiate this claim let us restrict attention to a class of simple lotteries identified by propositions of the form 'The chance of \$100 is x ' where, of course, $x \in [0, 1]$. In Figure 1, we plot some example desirability functions for these propositions against values for x (the chances). For convenience, we set the desirability of a zero chance of \$100 to zero and the desirability of \$100 for sure to 1. If the vN–M theory is correct, then the relationship between chances of \$100 and their desirability is linear and so the graph will be a straight line. On the other hand, if the chances have diminishing marginal desirability then the graph will be concave. But there are other possibilities: chances could have increasing marginal desirability, or a combination of increasing and decreasing marginal desirability at different chances, as illustrated by the snake-shaped curve plotted in Figure 1. Evidently, it is an empirical matter as to what attitudes to chances agents actually display and there is no *a priori* reason why the graph should have one shape rather than another (and no shape is imposed by the adoption of our framework).

To see how these curves capture the agent's attitudes to risk, let us define a risk function R on chances for an agent from her degrees of desire by setting $R(x)$ equal to $V(\$y)$ where $\$y$ is the amount of money such that the agent is indifferent between getting it with certainty and getting \$100 with chance x . So:

$$V(x, \$100) = V(\$y) = R(x).$$

Now the function R will behave much like the risk functions deployed in cumulative prospect theory and risk-weighted EU theory since, given our choice

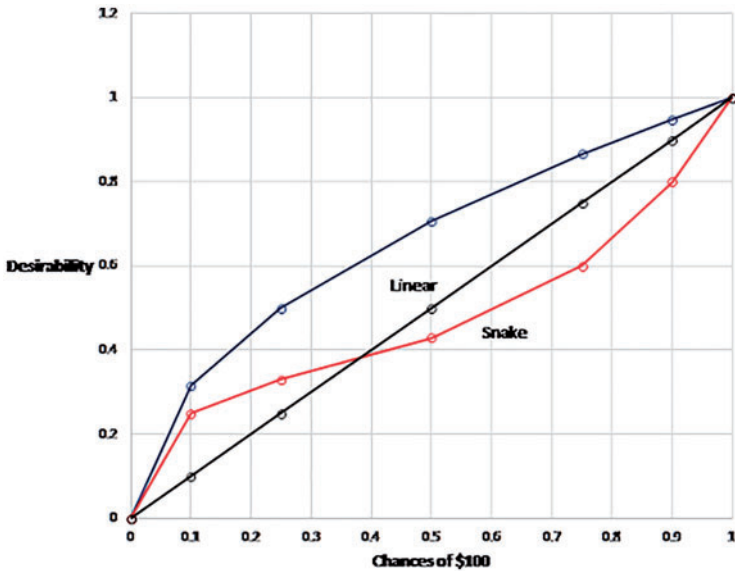


Figure 1. Candidate desirability functions for chances of \$100.

of scaling of the desirability function, $V(x, \$100) = R(x) \cdot V(\$100)$. That is, the desirability of some chance of \$100 will equal the risk-weighted desirability of \$100 for certain. (Nothing depends on this choice of scaling, but without it a somewhat more complicated definition of R would be required.) So, we can interpret the graphs in Figure 1 as candidate risk curves, representing the agent's attitude to risk, with the linear one being the vN–M risk curve and the snake-shaped one being the curve postulated by cumulative prospect theory. Crucially, however, the risk curves, so defined, are features of the agent's desires and not some distinct attitude. More precisely, these risk curves are determined by the relationship between the agent's desires for concrete goods and her desires for the chances of these goods. Hence, there is no reason to expect that properties of an agent's attitudes to the chances of one good, say money, will be the same as her attitudes to the chances of another, say health. A professional gambler who bets only to maximize expected monetary value when in the casino may be extremely averse to taking risks with his health.

It is the way in which the attitudes that agents take to goods combine with the attitudes they take to chances of goods that gives rise to behavioural risk aversion as initially defined, namely as a preference for lotteries over mean preserving spreads of them. For instance, diminishing marginal desirability of quantities of a good combined with constant marginal desirability of chances of the good—the only attitude to chances that the orthodox theory allows—gives rise to such risk-averse behaviour. But so does constant marginal desirability of quantities of the good combined with increasing marginal

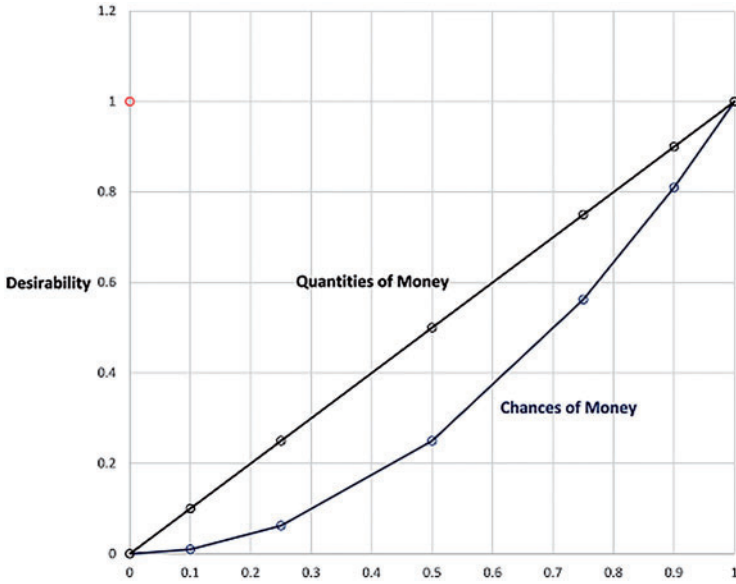


Figure 2. Linear desirability for quantities of money and increasing marginal desirability for chances of \$100.

desirability of the chances. This case is illustrated in Figure 2.¹⁴ Similarly, an agent may display risk neutrality in their choices (that is, indifference between a lottery and any mean-preserving spread of it) because they assign increasing marginal desirability to both quantities of the good and chances of it, or because they assign constant marginal desirability to both, or because they assign decreasing marginal desirability to both!

4.3 Ambiguity and the four-fold pattern

Let us now turn to the explanation of the two empirical phenomena that standard EU theory has such difficulty accommodating. The explanation we offer of an agent’s ambiguity attitudes towards some good is very straightforward; they are simply the reflection of the shape of her desirability function for the chances of the good. In particular, ambiguity aversion with respect to actions with consequences that are chances of monetary prizes reflects the diminishing marginal desirability of chances of money. So on our account,

¹⁴ In this figure and the next, we assume for convenience that $V(1, \$100) = V(\$100)$ and $V(0, \$100) = V(\$0)$. But these identities are not required by our theory.

ambiguity attitudes to goods are simply attitudes to the chances of these goods.

Consider again the example of bets on the two different coins that we gave earlier. A bet that the coin that is known to be fair will land heads has a 'sure' consequence of a chance of one-half of winning \$10. Hence, the desirability of this bet is equal to the desirability of a half chance of winning either \$10 or nothing. But the corresponding bet on the other coin has different consequences in different states of the world: if the coin is two-tailed then the bet has no chance of delivering the \$10, but if it is two-headed then it is certain to result in a win of \$10. Its desirability, assuming that the agent assigns equal probability to both possibilities (as revealed in an indifference between betting on heads and betting on tails), is the average of the desirability of no chance of winning \$10 and the desirability of winning \$10 for sure. Now if these chances have diminishing marginal desirabilities, then, by definition, the difference between a half chance and a zero chance of winning is greater than that between certainty of winning and a half chance of doing so. So then rationality requires that the agent prefer the bet on the fair coin, that is, that she displays ambiguity aversion.

The same explanation can be offered of the pattern of preferences exhibited in the Ellsberg paradox. Far from being inconsistent with Bayesian rationality, these preferences are, as Bradley ([2016]) shows, required of Bayesian agents who attach diminishing marginal desirability to chances of monetary outcomes. More generally, an agent with diminishing marginal desirabilities for the chances of some good will exhibit the kind of preference for hedging chances that is characteristic of ambiguity aversion.

The explanation of the four-fold pattern of risk preferences (in particular, the Allais preference) reported by Kahneman and Tversky ([1979]) is equally straight-forward. These results should be read as reporting the exchange rates between quantities of money and chances of winning a fixed amount of money induced by the agents' degrees of desire for these two types of good. So we can infer from them what sorts of relationships must hold between these desires. For instance, they report, for lotteries yielding \$100 with different chances, the following median cash equivalents: \$14 for chance 0.05, \$25 for chance 0.25, \$36 for chance 0.5, \$52 for chance 0.75, and \$78 for chance 0.95. These exchange ratios reflect the relationship between the agents' attitudes to monetary amounts in the range \$0 to \$100 and their attitudes to different chances of \$100. What cannot be determined from such data is the shape of the desirability function over each. We can however conclude that the subjects value gains in chances of \$100 more highly than the corresponding expected gains in monetary amounts when both the absolute chances and monetary amounts are small, but the other way around when both are large. So, agents will pay

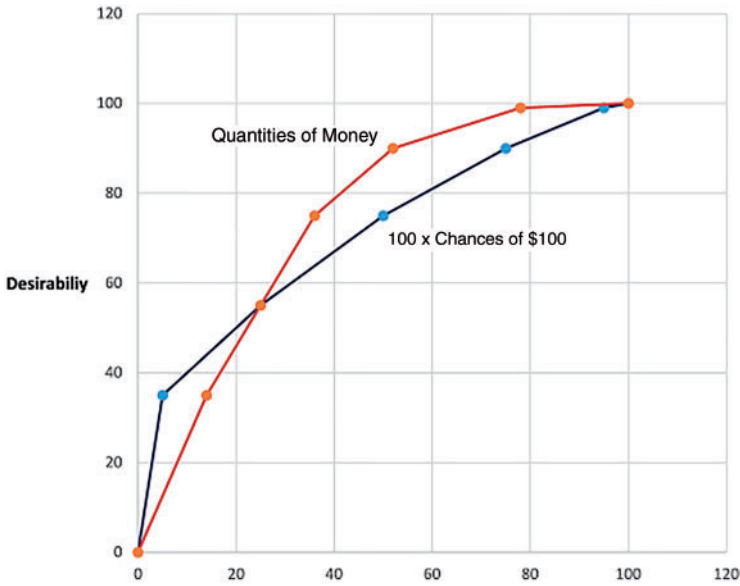


Figure 3. Desirability function for quantities of money and chances of \$100.

much larger sums of money for gains in chances when absolute chances are small than when they are large.

One way in which these constraints could be satisfied is if agents have desirability functions for money that are roughly linear in quantities and snake-shaped in chances, concave for low probabilities and convex for the very high ones. This is what is predicted by cumulative prospect theory. Such a postulate implies however that a Bayesian agent will exhibit ambiguity seeking preferences in situations involving lotteries with high chances of winning \$100.¹⁵ The empirical evidence offers little support for this implication; indeed, it is not consistent with the ambiguity averse patterns of preference frequently observed in the Ellsberg paradox.

In contrast, the four-fold pattern of risk preferences is perfectly consistent with ambiguity aversion in our framework. An agent will display both the four-fold pattern and ambiguity aversion when her desirability function is concave over both quantities of money and chances, but relatively less so over low chances of some monetary amount than small percentages of the

¹⁵ In cumulative prospect theory the chances of money have steeply increasing marginal utility close to certainty (the ‘certainty effect’). This implies that a Bayesian agent who is indifferent between say a 98% chance of \$100 if E and the certainty of \$100 if $\neg E$ and a 98% chance of \$100 if $\neg E$ and the certainty of \$100 if E (so subjectively regards E and $\neg E$ as equiprobable), will prefer either of these lotteries to a 99% chance of \$100 (whether E or not). But this is just what it is to be ambiguity seeking at those chances.

amount and relatively more so over high chances of the amount than over large percentages of the amount. To illustrate, suppose that agents have the desirability functions over monetary gains and chances of \$100 depicted in Figure 3. Note that although they have diminishing marginal desirabilities for both, the shapes of the function over the two are different, with the desirabilities of chances initially rising more rapidly than those of the monetary amounts, but less rapidly later on. Such agents would display precisely the risk-seeking preferences at low chances and risk aversion at high chances that Tversky and Kahneman report. For instance, they would be willing to pay more than \$1 to achieve a 0.01 chance at a \$100 prize when their chances of the prize are 0, and would also be willing to pay more than \$1 to avoid a 0.01 drop in the chance of \$100 when their chances are 1 (that is, when they have already secured the prize). They would also of course exhibit ambiguity aversion because of the concave shape of the desirability function for chances.

5 Conclusion

The orthodox treatment of risk attitudes in decision theory seems both conceptually and empirically inadequate. Conceptually because it fails to distinguish attitudes to concrete goods from attitudes to risks regarding these goods; empirically because it neither offers a satisfactory explanation of the four-fold pattern of risk behaviour observed in choice experiments, most famously in the Allais paradox, nor of ambiguity attitudes observed in setups such as the Ellsberg paradox. We have offered a framework in which it is both possible and natural to distinguish attitudes to concrete goods from risk attitudes, and show how these two types of attitudes can combine to determine an agent's choices in a way that is consistent with both the four-fold pattern of risk preference and with ambiguity aversion.

Acknowledgments

We would like to thank two referees for *BJPS*, the Practical Philosophy group in Uppsala, and participants in the 'Reasons and Mental States in Decision Theory' workshop at the LSE, for comments and suggestions that helped us improve this article. Stefánsson's work on this article was supported by the AXA Research Fund, with a grant for The (Dis-)Value of Risks and Chances (14-AXA-PDOC-222), for which he is grateful. Bradley gratefully

acknowledges support from the Arts and Humanities Research Council grant Managing Severe Uncertainty (AH/J006033/1).

H. Orri Stefánsson
Institute for Futures Studies
Stockholm, Sweden
orri.stefansson@iffs.se

Richard Bradley
London School of Economics and Political Science
London, UK
r.bradley@lse.ac.uk

References

- Allais, M. [1953]: ‘Le Comportement de l’Homme Rationnel Devant le Risque: Critique des Postulats et Axiomes de l’Ecole Americaine’, *Econometrica*, **21**, pp. 503–46.
- Bermúdez, J. L. [2009]: *Decision Theory and Rationality*, Oxford: Oxford University Press.
- Bolker, E. D. [1966]: ‘Functions Resembling Quotients of Measures’, *Transactions of the American Mathematical Society*, **124**, pp. 292–312.
- Bradley, R. [2016]: ‘Ellsberg’s Paradox and the Value of Chances’, *Economics and Philosophy*, **32**, pp. 231–48.
- Bradley, R. [2017]: *Decision Theory with a Human Face*, Oxford: Oxford University Press.
- Broome, J. [1991]: ‘Utility’, *Economics and Philosophy*, **7**, pp. 1–12.
- Buchak, L. [2013]: *Risk and Rationality*, Oxford: Oxford University Press.
- Diamond, P. [1967]: ‘Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparison of Utility: Comment’, *Journal of Political Economy*, **75**, pp. 765–66.
- Dreier, J. [1996]: ‘Rational Preference: Decision Theory as a Theory of Practical Rationality’, *Theory and Decision*, **40**, pp. 249–76.
- Ellsberg, D. [1961]: ‘Risk, Ambiguity, and the Savage Axioms’, *Quarterly Journal of Economics*, **75**, pp. 643–69.
- Hansson, B. [1988]: ‘Risk Aversion as a Problem of Conjoint Measurement’, in P. Gärdenfors and N.-E. Sahlin (eds), *Decision, Probability, and Utility*, Cambridge: Cambridge University Press, pp. 136–58.
- Harsanyi, J. C. [1977]: ‘On the Rationale of the Bayesian Approach: Comments on Professor Watkinss Paper’, in R. Butts and J. Hintikka (eds), *Foundational Problems in the Special Sciences*, Dordrecht: D. Reidel, pp. 381–92.
- Jeffrey, R. [1965]: *The Logic of Decision*, Chicago, IL: The University of Chicago Press.
- Kahneman, D. and Tversky, A. [1979]: ‘Prospect Theory: An Analysis of Decision under Risk’, *Econometrica*, **47**, pp. 263–92.
- Lewis, D. [1980]: ‘A Subjectivists Guide to Objective Chance’, in R. C. Jeffrey (ed.), *Studies in Inductive Logic and Probability*, Berkeley, CA: University of California Press, pp. 263–93.

- List, C. and Dietrich, F. [2016]: 'Mentalism versus Behaviourism in Economics: A Philosophy-of-Science Perspective', *Economics and Philosophy*, **32**, pp. 249–81.
- Okasha, S. [2016]: 'On the Interpretation of Decision Theory', *Economics and Philosophy*, **32**, pp. 409–33.
- Quiggin, J. [1982]: 'A Theory of Anticipated Utility', *Journal of Economic Behavior and Organization*, **3**, pp. 323–43.
- Rabin, M. [2000]: 'Risk Aversion and Expected Utility Theory: A Calibration Theorem', *Econometrica*, **68**, pp. 1281–92.
- Savage, L. [1954]: *The Foundations of Statistics*, New York: John Wiley and Sons.
- Stefánsson, H. O. [2014]: 'Risk and Rationality by Lara Buchak', *Economics and Philosophy*, **30**, pp. 252–60.
- Stefánsson, H. O. [2015]: 'Fair Chance and Modal Consequentialism', *Economics and Philosophy*, **31**, pp. 371–95.
- Stefánsson, H. O. and Bradley, R. [2015]: 'How Valuable Are Chances?', *Philosophy of Science*, **82**, pp. 602–25.
- Tversky, A. and Wakker, P. [1995]: 'Risk Attitudes and Decision Weights', *Econometrica*, **63**, pp. 1255–80.
- von Neumann, J. and Morgenstern, O. [2004]: *Games and Economic Behavior*, Princeton, NJ: Princeton University Press.
- Voorhoeve, A., Binmore, K. and Stewart, L. [2012]: 'How Much Ambiguity Aversion?', *Journal of Risk and Uncertainty*, **45**, pp. 215–38.
- Wakker, P. [2010]: *Prospect Theory: For Risk and Uncertainty*, Cambridge: Cambridge University Press.
- Watkins, J. [1977]: 'Towards a Unified Decision Theory: A Non-Bayesian Approach', in R. Butts and J. Hintikka (eds), *Foundational Problems in the Special Sciences*, Dordrecht: D. Reidel, pp. 345–79.