What can we rationally value?

Katie Steele

There are many responses to Allais's (1953) Paradox, some reconciling it with expected utility (EU) theory, others holding that the two cannot be reconciled (for better or worse for the theory). I here focus on a particular aspect of the debate—the extent to which EU axioms like Independence constrain the content of (not just the relationship between) our preferences. Many accept that an agent's feelings towards risk/regret can legitimately be considered part of an act's outcomes. This is nonetheless contentious, because risk and regret do not apparently reside in any particular outcome, but rather arise from the way in which potential act outcomes are related to one another, i.e. they are "global properties" of the outcome space or decision problem. I argue that in the normative context, even if "global properties" can be located in individual outcomes, the appeal to risk/regret attitudes to distinguish outcomes must answer to some kind of consistency considerations. Otherwise there is the danger of EU theory being irrefutable or lacking content. Furthermore, more permissive ways of including risk/regret in outcomes have implications for sequential decision scenarios that are analogous to the kind associated with violating Independence. So if avoiding these kinds of consequences is the main motivation for not violating Independence, then allowing the content of outcomes to be such that the same problems arise even if Independence is respected is not going to be any better a response to Allais's Paradox.

1 Introduction

In this paper I discuss a particular type of response to Allais's paradox. The paradox is formulated so as to isolate and thus highlight a specific tendency in some people's choice behaviour—the tendency to avoid particular kinds of risk or possible regret. Some think the risk/regret attitudes that Allais exposes concern the qualitative difference between certain lotteries and uncertain lotteries. This is a possible interpretation of the paradox. In my opinion, however, Allais's paradox demonstrates a broader kind of risk/regret attitude that does not necessarily hinge on any special properties of lotteries with certain prizes. It is a kind of risk/regret attitude that is relevant to any given choice scenario. (Allais's paradox is just nicely formulated so as

to make very clear some people's intuitions in a particular case as well as the consequences of these intuitions.) Granted that Allais's paradox has wide relevance, any response to the paradox should be carefully scrutinised because it will have far-reaching import. Indeed, the response I will focus on in this paper asks us to consider what expected utility (EU) theory has to say about the sort of things that can influence our preferences, or in other words, what we can rationally value.[1]

Before I go on to outline the paradox itself, let me say a little about the kind of risk/regret attitude that it exposes. The problem is formulated in such a way as to rule out the relevance of the diminishing marginal utility of goods/money phenomenon. In any case, while concave utility curves are sometimes referred to as risk-averse functions, I think they capture attitudes towards goods (usually money) rather than attitudes towards risk.[2] Theorists who have objectified what might properly count as risk/regret in order to explicitly incorporate such phenomena in decision models have defined it in different ways.[3] I am not here interested in evaluating these different conceptions of risk/regret; I draw attention only to the sort of risk/regret attitudes that are pertinent to Allais's paradox. Such attitudes hinge on the relationship between the different outcomes of an act. We could say that an act is less risky than another with the same expected utility if its possible outcomes are less varied in utility. I talk about risk and regret interchangeably here because, in the sense that I am using the terms, both types of sentiment have to do with uncertain outcomes, and hinge on the relationship between a realised outcome and the alternative outcomes one might

[1] Tversky makes this claim about how Allais's paradox raises the issue of what we can rationally value. In his own words, "The key issue…is not the adequacy of the axioms, but rather the appropriateness of the interpretation of the consequences" (1975, p. 171). In this paper I focus on Weirich's (1986) and Broome's (1991) accounts of what can legitimately affect our preferences or evaluation of consequences. Jeffrey (1982) and Sen (1985), in particular, also address this topic.

[2] Tversky (1975, p. 164) makes this point: "…it is somewhat misleading to refer to the measurement of the utility for money as 'the measurement of attitudes towards risk'".

[3] Most of these decision models that explicitly incorporate risk/regret are inconsistent with EU theory in that one or more of the axioms of EU theory are violated. See, for instance, Machina (1982 and 1991), Bell (1982) and Loomes and Sugden (1982). (Machina's models could be said to define risk/regret in terms of the relationship between outcomes in different states of the same act. The "Regret Theories" of Bell and Loomes and Sugden define risk/regret in terms of the relationship between outcomes of different acts that fall under the same state.)

otherwise have feared or hoped for. In other words, I am talking about attitudes that are not associated with any particular outcome of an act but that depend, rather, on the "global properties" of the act, i.e. the relationship between the outcomes of different states.

2 Two readings of Allais's paradox

Consider Savage's presentation of Allais's paradox (Resnik, 1987, p. 105): we have two choice situations, A and B, and in each choice scenario, we have a choice between two lotteries (*a* or *b* in situation A, and *c* or *d* in situation B). For each lottery there are 100 tickets; one ticket is chosen at random and different ticket numbers yield different money amounts, as per the table below. The punter is asked to choose their more preferred lottery in each of the two choice situations A and B.

Table 1

	Ticket Number		
	1	2-11	12-100
A:			
a	1 million	1 million	1 million
b	0	5 million	1 million
B:			
c	0	5 million	0
d	1 million	1 million	0

Savage presents the decision problem in this way so as to highlight the relevance of the Independence Axiom. Let me first give a formal expression of the Independence Axiom. Joyce (1999, p. 85) presents Savage's version of it as follows:

Suppose that A and A* produce the same outcomes in the event that E is false, so that $A_{\sim E} = A^*_{\sim E}$. Then, for any act $B \in \mathbf{A}$ (where **A** is the set of all acts, including constant acts), one must have

$A > A^*$ if and only if $A_E \,\&\, B_{\sim E} > A^*_E \,\&\, B_{\sim E}$

$A \geq A^*$ if and only if $A_E \,\&\, B_{\sim E} \geq A^*_E \,\&\, B_{\sim E}$

"In other words, a rational agent's preference between A and A* should not depend on what happens in circumstances where the two yield identical outcomes." (Joyce, 1999, p. 86)

Independence is relevant to Allais's decision problem because the lottery pairs in situations A and B are identical in the last column (*a* and *b* both have winnings of 1 million for tickets 12–100, and *c* and *d* both have winnings of 0 for tickets 12–100). So the choice between *a* and *b* in situation A should depend solely on whether the agent prefers 1 million for all 11 tickets, or whether they want to take a gamble on one ticket yielding nothing and the other 10 yielding 5 million. But this is exactly what the choice in situation B depends on. So if you choose *a* in situation A then for consistency you should choose *d* in situation B. And if you choose *b* in A then you should choose *c* in B.

The reason Allais's problem is considered paradoxical is that many reasonable people, when presented with the two choice situations, tend to do precisely what the Independence Axiom prohibits (and yet Independence seems a compelling constraint on rational choice). But it is hard to know what conclusions to draw from this disparity between what the Independence Axiom prescribes (in terms of the decision model laid out above) and what a number of reasonable-seeming people tend to do. Indeed, this is a highly contested issue. We could categorise the different positions in a number of ways, but for present purposes I am interested in whether the *a* and *c* choosers do indeed violate Independence. Accordingly, the first response is simply to acknowledge that *a* and *c* choosers violate Independence. (Then there are divisions in this camp as to what such a violation means.) The second response is that the decision problem must be incorrectly specified, because any well-specified problem would not

make a violation of Independence seem attractive to reasonable agents. I will discuss these two main responses to Allais's paradox in turn.

As briefly stated above, our first reading is to affirm that there really is a "paradox" here of sorts. This is to accept that a significant number of people (even after reflection) indeed violate the Independence Axiom when responding to Allais's problem. Of course, even if we accept the violation of Independence, there are different things that can be said about this. Savage (1954, p. 101 ff.) for instance thinks that Allais's problem simply exposes a common flaw in people's reasoning. The fact that many people stick to their faulty choices after reflection only goes to show how seductive the *inconsistent* choices are in this particular kind of scenario. Unerringly rational decision-makers, Savage would maintain, do not violate independence in this way; they would choose *a* and *d* or *b* and *c*, or else they would be indifferent between the two lotteries in each situation.

There are others who agree with Savage that many people violate Independence in the Allais scenario, but who take the moral of the story to be quite the opposite from what Savage takes it to be. They claim that Allais shows a fault in the Independence Axiom rather than a fault in people's reasoning. In other words, Allais's paradox is a genuine challenge to Savage's EU theory. According to this line of thought it is then a question of finding a suitable replacement theory that does not feature the Independence Axiom. Machina (1991, p. 51) lists a number of potential alternative decision models, each with its own risk/regret rationale. I won't pursue the details of these models; all we need to know is that they violate Independence.

Then there is the second reading of Allais's problem. A number of decision theorists (e.g. Jeffrey 1982, Weirich 1986, Broome 1991) seek to reconcile the paradox with EU theory.[4] Broome (1991, p. 107) claims, not unlike Savage, that violation of the

[4] I note that there is a third response to Allais's paradox that I am not emphasising here because I am interested in the tension between the other two responses. The third path is to explain the common responses to Allais's problem by relaxing some other axiom of EU theory (arguably an axiom more dispensible than Independence). Levi (1997) argues that his decision theory that relaxes ordering (and thus allows incommensurability of acts) can explain both the Allais and Ellsberg paradoxes. For Allais's paradox, Levi thinks it is plausible that the utility function is imprecise in

Independence Axiom is just outright irrational. But unlike Savage, and in common with those referred to above, Broome seeks ways of redescribing Allais's problem (specifically the potential outcomes) so that the common response to the problem turns out to be consistent with EU theory after all. Broome's conclusion is much the same as others in this camp—as far as the typical agent is concerned, the proper way to describe the stakes involved in Allais's problem is to include in the description/evaluation of act outcomes the agent's attitudes towards risk/regret. In this case, we might factor into the relevant outcomes the typical agent's comparative happiness or preference for outcomes that are more secure or closer to a sure deal. According to this account, we might specify the outcomes in option *a* as (1 million + ∂) rather than (1 million) apiece. The term ∂ represents the extra satisfaction that the agent experiences given that they know the 1 million monetary gain is a certainty. Once this move is made, the symmetry of the choice situations A and B is broken, and we no longer have a case where the Independence Axiom applies. This paper is mainly concerned with whether this move—including risk/regret sentiments in outcomes—is legitimate, or whether it is somehow at odds with EU theory.

3 Why not violate Independence?

Before I go on to discuss the latter "redescribing outcomes" response to Allais's paradox, I will give some motivation for this line of response. (I will be referring back to these motivations later.) Broome and others in this camp think that the common choices to Allais's problem are reasonable, and yet want to explain them without sacrificing Independence. So why is it so important to retain the Independence Axiom? Many claim that the reasonableness of Independence is self-evident. Indeed the axiom is intended to be an intuitive requirement of rationality. But even highly

---

such a way that in both choice situations the two acts are incommensurable with respect to expected utility comparisons. Where there is no preference between options, Levi claims that we can appeal to secondary security considerations—a person who shares his intuition to maximise the worst-case scenario will choose options *a* and *c*. Others also claim that Allais's paradox can be explained by relaxing ordering. Schervish *et al.* (1990) make the claim in passing. Bell (1982) and Loomes and Sugden (1982) claim that their respective versions of "regret theory" (which also relax ordering) can account for the choice of *a* and *c* in Allais's problem.

intuitive logical laws such as Excluded Middle have not been immune from challenge![5] And given that Allais's paradox is specifically formulated to challenge Independence, we might say that it is begging the question to respond to the paradox with a restatement of the intuitive reasonableness of Independence.

A more substantial defence of Independence will require some further appeal to the consequences of violating the axiom. Note that some argue that violation of Independence makes an agent vulnerable to a Dutch Book being made against them. The problem with this line of defence is that the Dutch Book argument assumes something akin to Independence, and so it is likely that anyone who challenges Independence is going to challenge the Dutch Book argument as well.[6] Armendt (1993) discusses this connection between Independence and the "value additivity" assumption of the Dutch Book argument (i.e. the sum of two fair bets is itself a fair bet). He cites Schick's (1986) argument to the same effect. The upshot of these claims is that those who want to relax Independence are not going to be impressed by Dutch Book criticisms of their position because Dutch Book arguments assume "value additivity" of bets and this is precisely what such critics challenge.

It could be said that the case for Independence being a requirement of rationality needs some further support. There are indeed some arguments to this effect that make reference to sequential choice situations or decisions represented in extensive form. Raiffa (1968, pp. 82–5) for instance represents the two choice situations in Allais's paradox as sequential decisions, and argues that Independence-violators are irrational in the sense that they will deviate from their initially chosen strategy, even if their preferences over outcomes are unchanging/stable. Machina (1991, p. 57) illustrates Allais's problem in sequential or extensive form with two branching tree diagrams, illustrated in Figure 1.

[FIGURE 1. See last page]

[5] Consider many-valued logics.
[6] Machina (1991) concedes that violating "stochastic dominance" (a criticism of standard prospect theory) makes an agent vulnerable to a Dutch Book. But we can focus on decision rules that violate Independence while satisfying "stochastic dominance".

If the agent chooses *a* in situation A, then it means that they prefer to choose down at the first choice node, once they get to this position. But this is the same choice node as occurs in situation B, so even if the agent *planned* to violate Independence (relative to their choice in situation A) and thus to choose up at the first choice node, *in the case that this node is reached* they would presumably change their mind and choose down instead. The presumption that the agent will change their mind is based on the claim that they choose at intermediate nodes as if confronting a fresh decision over the remaining portion of the tree. So the agent's choice at the intermediate node should match their choice in the case where the portion of the tree in question stands alone as a decision problem. The fact that the agent will deviate from his/her original plan in this way is considered a form of dynamic inconsistency, and reason not to plan to violate Independence in the first place.

Machina (1991) responds to this argument at length. As per the supposed Dutch Book consequences of violating Independence, Machina claims that the presumption that rational agents in sequential decision situations should make choices at intermediate decision nodes as if these were new and isolated decisions is precisely what is at issue in debates about Independence. I think Machina is right to claim that Independence violators by definition will not switch strategies (if their preferences remain stable) because what is important to them is the whole sequential decision problem, including the parts of the sequential tree that have already been "closed off" by circumstance. The fact that decisions at intermediate choice nodes must be made in light of the risks *already born* is part and parcel of violating Independence.

But this extra burden on the Independence-violator to maintain proper perspective on series of bets or sequential decisions seems to be precisely what is questionable about this account. Will it always be obvious to the agent whether they are supposed to be choosing from scratch or whether they should rather remain faithful to a strategy previously decided upon? Should all of our current choices depend on risks already born sometime in the distant past? There is something fishy about requiring of an agent that they continually bear in mind every outcome or choice situation that they might have faced, had the world turned out differently.

I note that Seidenfeld (1988) advances an even stronger argument for the dynamic incoherence of Independence-violators. Seidenfeld finds that an agent with stable preferences who obeys ordering (and stochastic dominance) but violates Independence will in some cases appraise extensive decision trees differently *at the outset* when outcomes at terminal nodes are exchanged for indifferent outcomes. This is a stronger argument because it is an inconsistency that cannot be avoided (as opposed to inconsistency due to short-sightedness of the agent). However, I think the previous argument about the onerousness of having to keep track of earlier branches of an extensive decision that are no longer possible is very plausible motivation for retaining Independence as a constraint on rational choice. It is this weaker motivation for not violating Independence that I will return to later when it comes to evaluating the "redescribing outcomes" response to Allais's paradox.

4 Savage's theory and the content of outcomes

As discussed, both Broome (1991) and Weirich (1986) seek to explain Allais's problem while not giving up the Independence Axiom. Their common strategy is to redescribe the relevant outcomes in Allais's problem. Moreover, both argue that Savage's theory, to its detriment, doesn't allow such a move. I will investigate the claim that Savage's (1954) theory does not permit act outcomes to involve sentiments towards risk. I think it is important to settle this issue, but note that the question remains as to whether it is in fact productive or desirable to incorporate risk/regret sentiments in act outcomes.

The argument is not that the Independence Axiom itself rules out risk sentiments being incorporated in outcomes or prizes. That would be begging the question. It is another assumption in Savage's theory that has been brought under scrutiny; both Broome (1991, pp. 115–117) and Weirich (1986, p. 424) point the finger at what Broome calls the "rectangular field assumption".[7] The assumption is not an intuitive

[7] Broome uses the term "rectangular field assumption" because the assumption in question concerns a product set, and "a produce set occupies a rectangle or a series of rectangles in vector space" (1991, p. 80).

requirement of rationality. It is what Joyce refers to as a "structure" axiom in Savage's theory, or what Hájek (2006 manuscript) calls "an *idealisation* of our *theory of rationality*" as opposed to "an *ideal* of rationality itself." It is described as follows:

> …there is a set of possible states of the world and a set of possible consequences, and that any function from possible states to possible consequences is an option. It follows that a possible consequence can be produced by a variety of options—by options that yield the consequence in every state and by options that yield it only in a single state.[8]

Broome (1991, p. 115) here states what he thinks are the consequences of this assumption:

> Take, for instance, the outcome shown in Table 15…as '0 and disappointment'. [Broome is referring to one of the outcomes of a gamble, the disappointment referring to the fact that the punter missed out on the better prize.] The rectangular field assumption says your preference ordering includes all arbitrary prospects. Amongst them is the prospect that leads to this particular outcome for sure. This prospect determines, whatever lottery ticket you draw, that you get no money and also feel disappointment. But this feeling of disappointment is supposed to be one you get as a result of bad luck in the draw. It is hard to see how you could feel it if every ticket in the lottery would lead to the same boring result. So this prospect seems causally impossible, and that may make if doubtful that it will have a place in your preferences.

To summarise, the claim that Broome makes here (and that Weirich also effectively makes in his further discussion of the issue) is that we cannot just randomly attribute outcomes involving risk sentiments to states, because such outcomes intimately depend on the precise combination of states/outcomes involved in an option.

But I think Weirich and Broome over-interpret the "rectangular field assumption", and single out Savage's theory unnecessarily, at least with respect to risk/regret. I agree that the assumption strongly suggests that the description of outcomes should

[8] This is a direct quote from Weirich (1986, p. 424). Weirich attributes the first sentence to Savage (1954, end papers and p. 14 f).

preclude risk sentiments, but it is not obvious that we should interpret it so literally. As discussed, the "rectangular field assumption" is an *idealisation*; it describes the preference space for an ideal agent in such a way that Savage's (1954) representation theorem gives us the nice uniqueness result that we are after. We cannot assume that an agent with such extraordinary cognitive powers actually exists. In fact, it is impossible for us ordinary mortals to have a complete preference ordering over the infinitely rich option space that Savage asks us to entertain. Further, many of the options in the option space described by the "rectangular field condition" will not be physically possible. And not just due to any deficiencies us non-ideal agents might have, but because the actual world constrains the set of actions that any agent, ideal or otherwise, is able to carry out. Just because we can conceive of an abstract map from states to some combination of outcomes doesn't mean that the act in question is, will be, or ever was, a viable possibility in the actual world. Moreover, many state-outcome combinations will not merely be impossible for an agent to achieve, but outright contradictory. Schervish *et al.* (1990, p. 842) ask how it is possible, for instance, for the outcome "walk to work in the rain" to occur in a state such as "bright sunny morning". So there is more than one way in which the acts in Savage's assumed outcome space are fictitious, whether or not we want to further introduce risk sentiments. Thus it is not clear why Broome, in his statement above, invokes causal impossibility as an unassailable obstacle for incorporating risk sentiments, *in particular*, into outcomes.

Given the ideal nature of Savage's "rectangular field condition", I think any attempt to draw from it concrete conclusions about the contents of act outcomes is questionable. I don't think Savage's theory in particular rules out risk/regret sentiments from featuring in the description of act outcomes. But again in contrast to Broome and Weirich, this is not to say that I think the incorporation of risk/regret sentiments in an EU decision model is necessarily a good thing!

5 A vacuous decision theory?

I think Savage's "rectangular field condition" does not legislate one way or the other with respect to risk/regret in outcomes. (In fact given that this assumption seems to disregard causal possibility in more ways than one, I think it is hard to draw from it any conclusions about the content of outcomes, such as whether outcomes can incorporate descriptions of the act that brought them about.) In any case, I think there are significant reasons to be concerned about an EU model that allows risk/regret sentiments in outcomes, even if what is at stake is an Independence-compatible explanation of Allais's paradox. The move threatens to trivialise EU theory (whichever representation theorem is assumed as its justification). Simply altering the representation of preferences or redescribing outcomes (such that winning $1 million in the comfort of it being a certainty and winning $1 million in a gamble are different outcomes) to force compliance with the EU model seems ad hoc or overly permissive. If we allow such moves then it is questionable whether we can produce counter-examples against EU theory, which would make it essentially vacuous.

Consider Broome's analogy between what it means for an agent to comply with the Independence axiom and what it means for the agent to comply with the (perhaps less controversial) axiom of transitivity.[9] The transitivity axiom would be vacuous if every time it seemed as if an agent were violating it, i.e. their preferences appeared to have the following cyclical form—

1. $A \succ B$ and $B \succ C$ but also $C \succ A$

—the preferences could simply be redescribed in the following way, for instance, such that adherence to transitivity acts as an assumption rather than as a result to be tested:

2. $A \succ B1$ and $B2 \succ C$ and $C \succ A$

(Nothing is said in the second set of preferences above about the relationship between B2 and A, so there is no violation of transitivity.) The same kind of argument applies to the Independence axiom, which is what is at stake in Allais's problem, and in discussions about risk sensitivity in general. The axiom is essentially vacuous if it acts as an assumption in our model of an agent's preferences/choice behaviour rather than as a result to be tested.

[9] Broome (1991) eventually concludes that risk/regret sentiments *are* a legitimate aspect of outcomes, but he discusses at length the problem of being too permissive with respect to defining act outcomes.

Let me note that some might be quite comfortable with decision theory being trivial in the sense alluded to above. Axioms such as transitivity and Independence might be considered logical relations that an agent cannot but satisfy. Preferences are assumed then to be rational, and the task is to determine the agent's probability (credence) and utility (value) functions such that he/she can be represented as an expected utility maximiser. I don't agree with this account, however, because I think EU theory is intended to provide a thicker notion of rationality than this. As a normative theory, it should be possible for an agent to fail it. We should be able to cast judgment on whether a particular set of preferences is rational or irrational. And such discriminatory judgments are possible only if there are some kinds of restrictions on what can count as an outcome.

Of course, we don't want to constrain the content of outcomes in a way that privileges particular theories of value. EU theory is intended to be value-neutral. It is not supposed to be the final word on practical choice. The theory is silent on substantive matters of value—what it demands is just that one's preferences are consistent, and this may well be satisfied by an agent who prefers genocide to a walk in the hills. But we could always tighten up our model of choice by supplementing EU theory with whatever ethical constraints seem most reasonable (Colyvan *et al.* to appear). When determining whether an agent is merely *rational*, however, we don't want to take controversial stands on what aspects of the world should be valued, and in what way. We are looking for suitable minimal constraints on the content of outcomes.

Here I am particularly interested in whether risk/regret attitudes can legitimately feature in outcomes. But it is useful to first consider general proposals for determining how outcomes should be described, and whether such proposals shed light on the risk/regret issue. Broome (1991), for instance, claims that we seek a principle for determining whether it is reasonable to have a preference between two outcomes, given that they differ in a specified way. This is to say that the differences between some outcomes might be judged so inconsequential that it is irrational to be anything but indifferent between them. For example, it would seem irrational for me to prefer the state of affairs where I am reading the paper at position A to the state of affairs where I am reading the paper at position B, where B is one millimetre to the left of A,

and all other things are equal. But even these kinds of judgments involve taking a stand on values; in the case described, why is such a small difference in spatial location unimportant?

Instead of trying to make objective claims about what properties discriminate one outcome from another, some have suggested that what matters, rather, is whether an agent's preferences are in some sense consistent. Pettit (1991) provides a good way of thinking about this issue. He argues that we owe an explanation for why one outcome (or prospect) is preferred to another, and this explanation hinges on what properties of the outcomes in question affect our evaluations. If a particular kind of property matters to an agent in one situation, then discrimination between outcomes on this basis is legitimate to the extent that the property *always* affects the agent's evaluations. So an agent must demonstrate consistency with respect to how he/she distinguishes and evaluates outcomes in terms of the properties that the outcomes exhibit. Of course, the story about properties will be complicated, as we will still need to make judgments about what can be suitably called a "property"—it must be a sufficiently salient or natural grouping.[10] There will be a further problem associated with interactions between properties and how such interactions influence evaluations. I will assume that such problems are surmountable, and thus take "property-consistency" as outlined to be the right kind of criterion for determining the proper content of outcomes for a particular agent.

The above criterion facilitates a thicker sense of rationality, such that it is possible to claim that in some instances an agent violates transitivity, for instance. So what does the "property-consistency" constraint say about risk/regret attitudes, and determining what counts as a violation of Independence? For starters, it is still unclear as to whether risk/regret properties can be associated with individual outcomes, or whether they are in fact "global" properties of an act. If we are not talking about properties of individual outcomes, then it is unclear whether any kind of "property-consistency" criterion can be sensibly applied. Many do think that risk/regret attitudes are associated with properties of individual outcomes, however. A common move is to

[10] Some kind of "naturalness" condition is necessary, because otherwise properties could be any kind of gerrymandered grouping and would not serve to constrain the content of outcomes at all.

claim risk/regret attitudes are simply a certain type of emotional response, either anxiety or excitement about risk that is generated in the act of choice, or regret or elation experienced when some outcome (rather than another) actually eventuates. If the agent consistently values this sort of excitement and elation favourably, and anxiety and regret unfavourably, then it is perfectly legitimate for the presence of such emotions to be cause for distinguishing between what would otherwise be identical outcomes.

The above account might be thought somewhat unsatisfactory if EU theory is taken to provide an even stronger account of rational preference— rather than accommodating our existing gut feelings about risk and regret, the theory could be understood as telling us *how* to feel about acts with uncertain consequences. In other words, it seems strange for the model to make concessions for psychological attitudes we might have towards risk/regret *in spite of* the recommendations of EU theory itself. But it could well be argued that emotional responses are among the facts pertaining to a given state of affairs and should thus be treated similarly to other kinds of candidate properties of an outcome. Moreover, some acts of choice will be accompanied by clear and predictable emotional reactions, and in these cases it seems unreasonable *not* to include such sentiments in the description of outcomes.

Jeffrey (1982) discusses an example problem where an apparent violation of Independence can be explained away by taking into account obvious emotional responses that distinguish between otherwise identical outcomes.[11] Sen (1985) gives a similar example that goes roughly like this: Mr Smith may receive a letter, and its contents will be one of two things, either notification of a large cash win in case 1, or else a court summons for some semi-serious traffic law infringement in case 2. For both cases, if he does not receive the letter, then Mr Smith will either clean up the house (activity A) or have some champagne and cheese (activity B). What this example is supposed to illustrate is that it would be very reasonable for Mr Smith to prefer activity A over B (cleaning the house over drinking champagne) in the event that he fails to receive a letter that might have contained a large cash prize, and at the

[11] Jeffrey takes his example from Machina (1982), who does not incorporate risk/regret sentiments in outcomes and formulates the example to illustrate the deficiencies of the Independence Axiom.

same time reasonable for Mr Smith to prefer B over A in the event that he does not receive a letter that might have contained a court summons. On the face of it this seems like a violation of Independence (because the alternative outcome of receiving a cash win/fine is identical for each option in both cases 1 and 2), but most of us think that the choice situations are not symmetrical because drinking champagne having avoided a court summons is quite a different outcome from drinking champagne after failing to win a prize. We can well imagine in this situation that there will be strong emotions at play—disappointment at not having won the prize versus relief at having avoided the fine. To ignore such obvious emotions in the decision model just seems wrong.

We might have here then a way to explain Allais's paradox. If an agent has some amount of anxiety when faced with uncertain options as compared to certain options, and anxiety is a property that consistently affects the agent's evaluation of an outcome, then further distinctions can be made amongst the outcomes in Allais's problem such that there is no violation of Independence. (Indeed this is the kind of "redescribing outcomes" explanation of Allais's paradox that is commonly put forward.) It seems a reasonable account because it is very plausible that the qualitative difference between certain and uncertain gambles will be linked to a noticeably different emotional reaction or sentiment. We might challenge even this claim, however. I think the case for detectable emotional responses is not so strong in the Allais case as it is for the Mr Smith-letter case. More importantly, it is questionable whether appeal to emotions is going to allow any kind of sophisticated account of risk/regret sensitivity. The move might explain Allais's problem, but it is doubtful that it can explain a subtler version of the "common consequence effect" (that doesn't involve a certainty/uncertainty distinction).[12] In any case, Weirich (1986) and Broome (1991) think Allais's problem raises issues about risk/regret sensitivity that go deeper than a certainty/uncertainty distinction. Moreover, they are both unsatisfied with loose appeals to emotional reactions. Weirich claims that choosing *a* and *c* in Allais's problem is rational even if the agent has completely

[12] Machina (1991, p. 49) notes that the Allais Paradox "is not an isolated example, but rather a member of a whole class of violations of replacement separability known as the 'common consequence effect'…namely a swing in preference from more risky to less risky sublotteries in one branch of a compound lottery as the sublottery in the other branch improves in the sense of first-order stochastic dominance."

neutral emotions all round. I agree that Allais's paradox is best interpreted as a far-reaching challenge to EU theory—it uncovers more widespread risk/regret sensitivity than just coarse distinctions between certainty and uncertainty. But to the extent that Weirich and Broome try to marry more sophisticated risk/regret sensitivity with the EU framework, I think their accounts run into problems.

Broome appeals to counterfactual properties of outcomes as a way to accommodate sophisticated risk/regret attitudes. In this way, risk/regret attitudes are effectively linked to "global properties" of an option, but importantly, these "global properties" are present in individual outcomes in the form of counterfactuals. In the Mr Smith case, for instance, there will be two different outcomes "drink champagne given that a money prize would otherwise have been received" and "drink champagne given that a fine would otherwise have been received". We can distinguish these two outcomes without recourse to emotional response. (Appeal to emotions gives us the outcomes "drink champagne feeling disappointed" and "drink champagne feeling elated".)

My concern about the counterfactual tactic amounts to the old problem that it is too permissive with respect to the way outcomes are described. Recall the suggestion that outcomes may only be distinguished in terms of properties that consistently affect the agent's evaluations. The question is whether there can be any salient groupings of counterfactual statements that could serve as this kind of property? What group or property should "drink champagne given that a fine would otherwise have been received" belong to? We might be able to make some coarse-grained distinctions—for instance, we could group outcomes that involve the counterfactual statement that something much better or much worse would otherwise have eventuated. But if we were to treat outcomes differently depending on the precise combination of other possible outcomes associated with the same act, then we are talking about a large number of very specific outcome properties, and are thus running the risk of having a vacuous decision theory. Any set of preferences could be defended as rational.

Weirich's (1986) account of how EU theory can handle risk/regret could also be interpreted as resting on counterfactual properties of outcomes.[13] Weirich allows an outcome to be sensitive to the precise distribution of other possible outcomes associated with the act. This might seem overly permissive, but Weirich proposes that the contribution of "global properties" to the value of outcomes be constrained by a specific algorithm or decision rule. He experiments with a rule whereby the value of an outcome is modified by a function of the variance of the act's overall outcome distribution. (This is supposedly consistent with EU theory because once outcomes have been modified to include a variance factor, the best act is the one that maximises expected utility.) I think there are reasons to think that a variance factor is not a good choice of rule for modifying outcomes.[14] But whichever rule is decided upon, I think it is suspect to categorise this response to risk/regret sensitivity as a mere "redescription of outcomes". Effectively an alternative decision rule is being used to decide upon the ordering of acts, but the details of the rule are hidden away in the evaluation of individual outcomes, so that on face value, consistency with EU theory is maintained. Again it is not clear whether specific distributions of outcomes associated with an act should be treated as properties of each singular outcome.

7 Revisiting the motivations for respecting Independence

In determining when to allow risk/regret sentiments to be part of an outcome's description, it is not just a vacuous EU theory that must be guarded against. The other important consideration is that there is no point turning to a "redescribing outcomes" response to Allais's paradox if this strategy falls prey to the same kinds of problems associated with violating Independence that motivated the move to this alternative strategy in the first place. Of course some may maintain that Independence is intuitively a rational requirement despite Allais's challenge, or that it is important for

[13] Weirich himself does not appeal to counterfactual properties. But I think his account would be strengthened if he did rationalise it in this way.

[14] If a variance factor is not incorporated in individual outcomes but is rather treated as a modification of EU theory, then it is possible for a dominated act to be selected if the dominating act has sufficiently large variance (where less variance in outcomes is better).

a decision rule to be justified by a representation theorem. But I think there are more substantive arguments for retaining the axiom. In particular, I earlier outlined concerns about the implications of violating Independence in sequential decision scenarios. My worry with the "redescribing outcomes" approach is that an analogous kind of problem arises in the context of sequential decisions, and so in this sense the approach fares no better than outright violation of Independence.

If the description of outcomes includes counterfactual propositions about what other act outcomes might have eventuated had the world turned out differently, then essentially outcomes include both intrinsic and relational properties, and the relational properties may change throughout a sequential decision. In this way the agent might perceive terminal outcomes differently at different points in the sequential decision. At intermediate nodes, only a subset of the original set of outcomes will still be possible, and so it would be natural for the agent to perceive "global properties" (counterfactual properties) with respect to this subset of outcomes—the ones that are still possible. But this would amount to recognising different terminal outcomes at different choice nodes in the sequential decision. And changes in terminal outcomes will in some instances lead to the agent acting against their initially chosen strategy, despite having stable preferences. Of course, we can stipulate that the agent must be mindful of continuing to make choices with respect to the original descriptions of outcomes (which reference the full range of possible outcomes). In other words, the agent must at no time forget about branches of the decision tree that have already closed. But, if the need for an agent to maintain perspective on the whole sequential decision problem (in order to stick to a strategy) is considered too burdensome and thus reason not to violate Independence, then so too should it be considered reason not to incorporate "global properties" in outcomes.

So is redescribing outcomes to include global-property-inspired risk/regret sentiments a legitimate way of explaining Allais's paradox? I have argued that the first danger of this move is that it makes EU theory vacuous. Those who think that EU theory merely elucidates logical relations among preferences may be unconcerned about this. I am not of this opinion, however. I think outcomes must be constrained in some way if EU theory is to have any normative content. But I don't think we can find these constraints amongst the assumptions of Savage's representation theorem. The idea that outcomes may be distinguished only with respect to properties that *consistently*

affect the agent's evaluations of states of affairs is I think the right way to limit what counts as a new or different outcome. It is very plausible that an agent would consistently regard anxious sentiments as having disutility, and feelings of relief as having positive utility. So it is possible to explain Allais's paradox if we tell a story about the emotional state of the agent—an extra property of relief or confidence distinguishes the outcomes associated with the certain prize in problem A. But both Broome and Weirich seek a more sturdy explanation of the Allais-type choices, one that does not appeal to emotions but rather to more objective properties of outcomes. Both can be interpreted as appealing to counterfactual properties of outcomes, counterfactuals regarding what alternative outcomes might have otherwise eventuated from the act in question. I find this account problematic because it is not clear that such counterfactual propositions can be grouped into "properties". Furthermore, if outcomes are partially defined in terms of the "global properties" of an act, it is not so straightforward for an agent to follow an initially chosen strategy in a sequential decision scenario, even if their preferences remain stable. And this is precisely what some consider problematic about decision models that violate Independence! So if avoiding complications with sequential decisions is taken to be the main motivation for not violating Independence, a sophisticated "redescribing outcomes" approach is no better an alternative.

References

Allais, Maurice. 1953. Fondements d'une théorie positive des choix comportant un risque et critique des postulats et axioms de l'école americaine. Econométrie 40:257–332.

Armendt, B. 1993. Dutch Books, Additivity and Utility Theory. Philosophical Topics 21:1–20.

Bell, David E. 1982. Regret in Decision Making Under Uncertainty. Operations Research 30 (5):961–981.

Broome, John. 1991. Weighing Goods: Equality, Uncertainty and Time. Oxford; Cambridge, MA: Basil Blackwell Ltd.

Colyvan, Mark, Damian Cox and Katie Steele. *To appear.* Modelling the Moral Dimension of Decisions.

Hájek, Alan. *To appear*. Problems for Probabilism.

Jeffrey, Richard. 1982. The Sure Thing Principle. PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1982 (Volume Two: Symposia and Invited Papers):719–730.

Joyce, James M. 1999. The foundations of causal decision theory. Cambridge; New York: Cambridge University Press.

———. *Forthcoming*. Decision Theory. In The Philosophy of Science: An Encyclopedia, edited by J. Pfeifer and S. Sarkar. Routledge.

Loomes, Graham, and Robert Sugden. 1982. Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty. The Economic Journal 92 (368):805–824.

Machina, Mark J. 1982. "Expected Utility" Analysis without the Independence Axiom. Econometrica 50.

———. 1991. Dynamic Consistency and Non-expected Utility. In Foundations of Decision Theory, edited by M. Bacharach and S. Hurley. Oxford and Cambridge, MA: Basil Blackwell.

Pettit, Philip. 1991. Decision Theory and Folk Psychology. In Foundations of Decision Theory: Issues and Advances, edited by M. Bacharach and S. Hurley. Oxford: Blackwell.

Resnik, Michael D. 1987. Choices: an introduction to decision theory. Minneapolis:

University of Minnesota Press.

Savage, Leonard J. 1954. The Foundations of Statistics. New York: Wiley.

Schervish, Mark J., Teddy Seidenfeld, and Joseph B. Kadane. 1990. State-Dependent Utilities. Journal of the American Statistical Association 85 (411):840–847.

Schick, Frederic. 1986. Dutch Bookies and Money Pumps. Journal of Philosophy 83:112–119.

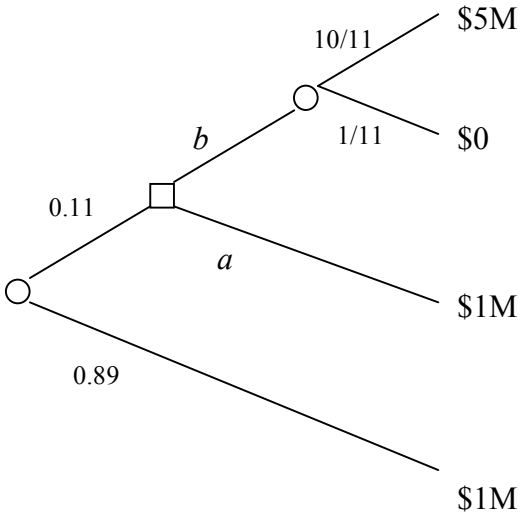
Seidenfeld, Teddy. 1988. Decision Theory Without "Independence" or Without "Ordering". Economics and Philosophy 4:267–290.

Sen, Amartya. 1985. Rationality and Uncertainty. Theory and Decision 19:109–28.

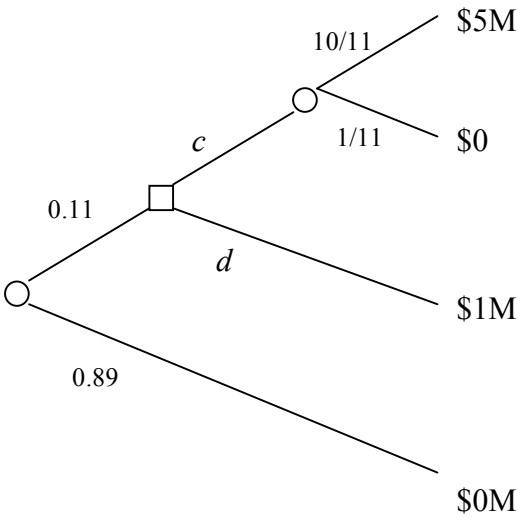
Tversky, Amos. 1975. A Critique of Expected Utility Theory: Descriptive and Normative Considerations. Erkenntnis 9:163–173.

Weirich, Paul. 1986. Expected Utility and Risk. British Journal of Philosophy of Science 37 (4):419–42.

FIGURE 1



Problem A



Problem B