4 ALLAIS'S PROBLEM AND THE INDEPENDENCE AXIOM

Katie Steele

4.1 Introducing the concept of "risk"

This chapter serves as an opening to the second part of this thesis. My general preoccupation in these chapters (4–6) is whether the subjective expected utility (SEU) theory treatment of "risk" is the only rational approach. In other words, is there scope for an agent to evaluate acts according to a different calculus, or does rationality require that one should always seek to maximise expected utility? This basic question has lead to the development of a number of alternative decision models. (There have been even more doubts raised about the descriptive adequacy of SEU theory, but recall that I am concerned only with the normative case here.) Importantly, the various alternative decision models involve different sorts of adjustments to SEU theory, depending on how the concept of "risk" is interpreted, and how it is thought to affect choice. I cannot here give a comprehensive survey of the ways that "risk" has been defined, and all the associated modifications of SEU theory. I focus, rather, on a particular class of risk measures—those that pertain to an individual act and its distribution of outcomes (or, more precisely, its distribution of outcome utilities). This class of risk measures might include the spread of outcome utilities for an act, or the utility of the worst-case outcome. My interest is whether it is rational to take account of this kind of risk in decision-making.

A simple example might help to make the above comments about "risk" clearer. Consider the decision problem in Figure 4-1 below.

Figure 4-1

	Coin lands heads	Coin lands tails
Bet 1	Gain $50	Gain $50
Bet 2	Gain nothing	Gain $100

If we assume that the utility-money relationship is linear (even though this assumption generally doesn't hold), the two bets in Figure 4-1 have the same expected utility. In such cases, SEU theory demands that a rational agent be indifferent between them. But the two bets involve varying degrees of risk, that is, if we define "risk" in terms of the distribution of outcomes for an act. In the first case, the agent is assured of $50 no matter how the world turns out. In the second case the agent will win $100 if the coin lands tails, but the coin may land heads, in which case the agent will come away with nothing. We might say that the second bet is more "risky", either because its worst-case outcome has less utility than that of the first bet's worst-case outcome, or because its outcomes have a greater spread in utility (or greater variance). In any case, the general class of risk measures I am interested in are those that concern a single act and its distribution of outcome utilities. By way of contrast, it is worth noting that the "Regret Theories" of Bell (1982) and Loomes and Sugden (1982) define "risk" in terms of the relationship between outcomes of different available acts that fall under the same state. According to this alternative class of risk measures, the two bets in Figure 4-1 are similarly risky, because when measured against each other, they each involve the same gains and losses, given the two ways that the coin might land. (If the coin lands heads, then Bet 1 is up $50 compared to Bet 2, but if the coin lands tails, Bet 2 is up $50 compared to Bet 1.)

We can associate the idea of "regret" with the concept of "risk". Indeed, the two compliment each other because both have to do with uncertain outcomes, and hinge on the relationship between a realised outcome, and the alternative outcomes one might otherwise have feared or hoped for. Above, I contrasted the class of risk measures that will be my interest in these next few chapters with an alternative concept of risk that involves comparisons between acts. We might make the same distinction using the language of regret. In the contrast case, the focus is on how the agent rates the act that they actually chose against the other acts that they might have chosen. For example, if the coin in the Figure 4-1 decision problem ends up landing heads, there is regret associated with the choice of Bet 2, because if the agent had chosen Bet 1, they would have been $50 better off. As stated, the sort of regret attitude that I am concerned with depends rather on the relationship between the outcomes of an individual act. The agent may regret receiving a particular outcome, not because they might have performed a different act, but because they are considering how the world might otherwise have turned out (in which case the same act would have led to a different outcome). There may be yet other ways to conceive of "risk/regret", but I have surely described two of the major classes of risk/regret measures. From now on, however, whenever I refer to "risk" or "regret", I am talking about a measure of the distribution of outcomes for a single act.

4.2 Allais's problem and the independence axiom

There are two well-known problems in the decision theory literature that press the issue as to whether it is rational for risk/regret considerations to affect an agent's preferences/choices. I am referring to the Allais and Ellsberg problems. In this chapter I will focus on Allais's (1953) problem, and in particular, the relationship between "risk-sensitivity" and the "independence" axiom of SEU theory. The Ellsberg problem raises some further questions about the relative precision of our beliefs; I will take this up in Chapter 5.

I turn then to the details of Allais's problem, and, in particular, to Savage's presentation of the decision scenario (Resnik, 1987, p. 105): we have two choice situations, *A* and *B*, and in each choice scenario, there is 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 monetary amounts, as per Figure 4-2 below. The punter is asked to nominate their preferred lottery in each of the two choice situations *A* and *B*.

Figure 4-2

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

A number of empirical studies have shown that many people make inconsistent choices, by the lights of SEU theory, when faced with Allais's problem—they opt for the sure 1 million in problem *A*, but then choose the 5 million gamble in problem *B*. (I will call this combination of choices the "Allais-choices", and the people who select these options the "Allais-choosers".)[1] The Allais-choices are inconsistent according to SEU theory because the choice of the sure 1 million in problem *A* indicates that the agent holds the following to be true:

[1] Allais's initial empirical finding was that the majority make the "Allais-choices" (MacCrimmon and Larsson 1979).

$$U(1\ \text{million}) > 0.10 \times U(5\ \text{million}) + 0.89 \times U(1\ \text{million})$$

$$0.11 \times U(1\ \text{million}) > 0.10 \times U(5\ \text{million})$$

The choice of option *c* over option *d* in problem *B*, on the other hand, indicates that the agent holds this expression to be true:

$$0.10 \times U(5\ \text{million}) > 0.11 \times U(1\ \text{million})$$

Clearly these statements are contradictory.[2] But we might wonder why the Allais-choices seem so reasonable to many. Some argue that the Allais-choosers are sensitive to the varying riskiness of the options in problems *A* and *B*, and that SEU theory does not adequately accommodate this kind of risk-sensitivity.

Before continuing, I want to point out that the above analysis highlights a particularly valuable aspect of Allais's decision problem—the problem isolates, as far as possible, the influence that risk/regret considerations might have on choice. In fact, both Allais's and Ellsberg's respective problems are formulated in such a way as to rule out the relevance of the diminishing marginal utility of goods/money phenomenon. We can see from the expressions above that the Allais-choices are inconsistent, whatever utility the agent attributes to the stated outcomes. The same cannot be said of the simple example that I gave in Figure 4-1 above. In that case, we could try to argue that SEU theory is inadequate, and make some appeal to the varying riskiness of the bets in order to defend a preference for Bet 1 over Bet 2. But a SEU theory-defender could simply reply that there is a much more obvious explanation for the preference for Bet 1—the agent's utility for $100 might be less than double their utility for $50, in which case the two bets do not have equivalent expected utility after all. While a concave money-utility curve like this is sometimes referred to as a "risk-averse" function, I think it captures an attitude towards the good in question (in this case money) rather than an attitude towards risk. In any case, this is not the type of risk-sensitivity that I am interested in here; my question is whether it is legitimate for an agent's choices to be sensitive to the spread of outcome utilities associated with an

[2] I am assuming that the agent's subjective beliefs correspond to the (arguably) objective probabilities for each of the three outcomes. But this assumption is not necessary. The agent's choices can be shown to be inconsistent regardless of what probabilities they assign to the three states.

act. Allais's problem, in particular, is designed to press this very question.

Savage presents Allais's problem in the manner depicted in Figure 4-2 so as to highlight the relevance of the independence axiom. This axiom is the backbone of SEU theory's approach towards risk. But I will say a bit more about the connection between the handling of risk and the independence axiom shortly. 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 (acts) 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 $\mathbf{A}$ 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, independence holds that "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)[3]

In exposing agents' risk-sensitivity, Allais's decision problem effectively challenges the independence axiom of SEU theory. If we model the problem as per Figure 4-2 above, 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. (Or else you can be indifferent between both

[3] I gave this same presentation of the independence axiom in Chapter 1.

lotteries in each choice scenario.)

The reason Allais's problem is considered paradoxical is that while many people think it is reasonable for choice to be affected by the spread of outcome utilities, as well as the expected utility, of an act, many also regard independence as a compelling constraint on rational choice.[4] Now there is clearly much intuitive appeal to the independence axiom when it is stated in general terms as the rule that "choice between acts should not depend on circumstances in which the acts yield identical outcomes". But it must be noted that it is precisely this axiom that precludes any kind of risk measure from playing a formal role in the decision calculus. The risk measures that I have been referring to rest on "global" properties of acts. If such a measure was incorporated in the decision calculus then acts could not be evaluated on a state-by-state, or an outcome-by-outcome, basis. Individual outcome utilities would be important, but so too would be the relationship between outcomes. This entails a violation of independence because it allows preferences between acts to shift, depending on how the outcomes that are common to the acts in question affect their respective spreads of outcome utilities.

4.3 The significance of empirical results about choice behaviour

It is important to be clear about just what we can learn from experiments like those involving Allais's decision problem. Since Allais first proposed his famous problem, there have been various empirical tests to ascertain what factors affect the choices people make in Allais-type decision situations. MacCrimmon and Larsson (1979, pp. 350–59) reference the historical results of Allais and Hagen, in addition to outlining some experiments of their own. As might be expected, the results are sensitive to the particular group of people who are sampled, how the problem is presented to them, and the parameter values that are used. Both Allais and Hagen found that the majority

[4] MacCrimmon and Larsson (1979) cite empirical findings that support this combination of sentiments.

of their subjects made choices that seem to violate the independence axiom (and in particular, the majority of choices showed an "aversion towards risk"). But MacCrimmon and Larsson later found that this result is highly sensitive to the probabilities and utilities involved. When monetary amounts are extreme (in the order of millions of dollars) and the probabilities are very disparate (such that the agent must compare a 100% chance with a 5% chance), a number of tests confirm that there is a tendency to make choices that appear to violate independence. But when the parameters have values that the average person is better able to comprehend, such that they can make more meaningful comparisons between acts, we do not see the same trend.

These empirical results are of course very valuable, and are particularly pertinent to the descriptive study of choice. But it is important to question what we should ultimately take away from findings about how people actually choose, when considering the normative question of how people should choose. Kahneman and Tversky (e.g. 1982), for instance, have drawn attention to a number of mistakes that the majority of people tend to make when reasoning with probabilities. The lesson here is that we should be very cautious about basing the properties of a normative decision model on the choice behaviour of the majority. But this is not to say that trends in actual choice behaviour cannot offer any insights into the normative study of choice. For one thing, convincing empirical results serve to focus attention on aspects of the normative decision model that have implications we might want to challenge. This is precisely what has happened in the case of Allais's problem and the independence axiom of SEU theory. Secondly, while broad statistical results about choice behaviour in Allais's betting scenario are ultimately neither here nor there when it comes to the normative decision model, principles of rational choice must at least appeal to reasonable-seeming people, upon reflection. And a number of reasonable people (in particular, some decision theorists) have declared that they would choose the combination of bets in Allais's decision scenario that appears to violate independence, even after reflecting on this fact. In my opinion this means that we cannot take the logical necessity of the independence axiom for granted.

4.4 Two readings of Allais's "paradox"

Even if we restrict our attention to the reasonable and reflective Allais-choosers, it is hard to know what conclusions to draw from the apparent disparity between their choice behaviour and what the independence axiom prescribes. Indeed, this is a highly contested issue, and it is the focus of this chapter from here on in. We could categorise the different positions in a number of ways, but for present purposes I am interested in whether those who opt for the sure 1 million in problem *A*, but who would take the 5 million gamble in problem *B*, do indeed violate the independence axiom. Accordingly, the first response is simply to acknowledge that the Allais-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 problem in turn.

As briefly stated above, our first reading is to affirm that there really is a "paradox" here. This is to accept that a significant number of reasonable people (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 maintains 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* (as per Figure 4-2), or else they would be indifferent between the two lotteries in each situation. The paradox is dissolved in this way: the seemingly reasonable Allais-choosers are not so reasonable after all. But as stated above, I think this response is a bit quick. In my opinion, Allais's problem demands a more substantial defence of the independence axiom.

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 problem is a genuine challenge to SEU theory. According to this line of thought it is then a question of finding a suitable replacement theory that involves some relaxation of the independence axiom. Machina (1989, p. 1631) lists a number of potential alternative decision models, each with its own risk/regret rationale (that will allow for an explanation of the Allais-choices). While I will not pursue the details of these models, in the next Section I do discuss how there are more and less defensible ways to violate independence. For now, all we need to know is that there are a number of alternatives to SEU theory that involve some relaxation of independence.

Then there is the second reading of Allais's problem. A number of decision theorists (e.g. Jeffrey 1982, Weirich 1986 and Broome 1991) seek to reconcile the paradox with SEU theory.[5] Broome (1991, p. 107) claims, not unlike Savage, that violation of the independence axiom is just outright irrational. But unlike Savage, and in common with those just referred to, Broome seeks ways of redescribing Allais's problem (specifically the act outcomes) so that the common response to the problem turns out to be consistent with SEU 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

[5] I note that there is a third response to Allais's paradox that I am not emphasizing 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 SEU theory (arguably an axiom more dispensable than independence). For instance, Levi (1997) argues that his decision theory, which relaxes ordering (and thus allows for incommensurability of acts), can explain both the Allais and Ellsberg paradoxes. I will give the details of Levi's theory in Chapter 5, and will apply it Ellsberg's problem. When it comes to Allais's problem, in brief, Levi thinks it is plausible that the agent's utility function is indeterminate in 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*. (I do not agree with Levi's treatment of secondary security considerations, as will become clear in Chapter 5.) 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 "Allais-choices". (I don't see how relaxing ordering alone will give the risk-sensitive choices, but again, this general issue will be taken up in Chapter 5.)

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. For instance, we might factor into the relevant outcomes the typical agent's comparative happiness when it comes to outcomes that are more secure or closer to a sure deal. Accordingly, the outcomes for act *a* (refer to Figure 4-2) might be specified as 1 million + ∂, rather than 1 million apiece, where 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. The rest of this chapter is mainly concerned with whether this move—including risk/regret sentiments in outcomes—is legitimate, or whether it is at odds with SEU theory. In the latter case, if we persist in claiming that risk/regret considerations legitimately affect choice, then we are effectively demanding a relaxation of the independence axiom of SEU theory.

4.5 Why not relax independence?

Before I discuss the latter "redescribing outcomes" response to Allais's paradox, let me elaborate on the motivations for this line of response. Broome and others in this camp think that the "Allais-choices" 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. As mentioned, I do not think this is a good enough defence of the axiom. Given that Allais's problem and the broader phenomenon of risk-sensitivity presents a challenge to independence, we might say that it is begging the question to respond with a restatement of the intuitive reasonableness of the axiom. In a moment, I will consider possible avenues for a more substantial defence of independence. But it is worth bearing in mind throughout this discussion that even if there is no decisive argument one way or the other, independence remains a compelling constraint on rational choice. So if Broome and co. can accommodate risk/regret attitudes within the SEU model, then surely this is the path of least resistance for anyone who wants to take seriously the Allais-choices, and risk/regret sensitivity in general.

A more substantial defence of independence will require some further appeal to the consequences of violating the axiom. It is generally agreed that not all violations of independence are on the same footing, and some violations lead to obviously bad outcomes. Consider the original "Prospect Theory" proposed by Kahneman and Tversky (1979).[6] This decision calculus has the form:

$$\mathrm{U}(X) = \boxed{\qquad\qquad}$$

where U is the agent's utility function, and option X has possible outcomes $O_1 \ldots O_m$ corresponding to states $S_1 \ldots S_m$, and

$\pi(.)$ is the agent's subjective risk function, subject to $\pi(0) = 0$ and $\pi(1) = 1$.

Machina (1989, p. 1634) notes that, if the function $\pi(.)$ is not linear (because in that case the agent's preferences would just conform to the SEU axioms), then there will exist probabilities over outcomes $\Pr(S_1) \ldots \Pr(S_m)$ summing to unity, such that

$$\boxed{\qquad\qquad} \neq \pi(1)$$

Take the case where

$$\boxed{\qquad\qquad} > \pi(1).$$

(The reverse case yields a similar result.) Then there will exist outcomes

$$O_1 < O_2 < \ldots < O_{\mathrm{m}} < O_{\#} \qquad \text{such that}$$

$$\boxed{\qquad\qquad} > \mathrm{U}(O_{\#}) \times \pi(1)$$

The agent is thus directed to choose the left-hand option, which means that they will miss out on a sure amount d, where $d > \mathrm{U}(O_{\#}) - \mathrm{U}(O_m)$! Another way of putting the

[6] The comments in this paragraph closely follow Machina's (1989, p. 1634) discussion of the original Prospect Theory.

above point is that Prospect Theory, as outlined above, violates independence in such a way that it also violates "first-order stochastic dominance", or what is simply referred to as "dominance"—there will be cases where, no matter how the world turns out, one act yields a preferred outcome, and yet the agent opts for the alternative act. We could say that the agent's decision calculus robs them of a sure amount, and this can only be considered a bad thing.[7]

But independence is a stronger constraint on choice than simple dominance, and so it is possible to relax independence while still respecting dominance. The major difference between the two is that dominance deals in simple or "basic" outcomes, while independence deals in "compound" outcomes. (The difference between the two is that compound outcomes are described in terms of a probability distribution over basic outcomes. Basic outcomes cannot be described in terms of a probability distribution over more basic outcomes.) Dominance stipulates that if you prefer the basic outcome O_x to another basic outcome O_y, then when you are faced with two acts that are exactly the same, except that the first yields outcome O_x in one particular state (or with some probability p), while the other yields outcome O_y for the same state, then you should prefer the first act. The independence axiom looks very similar to this, except that it applies to compound outcomes as well as basic outcomes. As per Joyce's formulation of the axiom given above, independence states that, for any two acts A and A^*, if you prefer A to A^*, (note that these are acts rather than basic outcomes), then you should prefer the compound act that for some partition of the state space yields A, to an otherwise identical compound act that yields A^* for the given set of states.

Whatever we might think about the rationality of risk/regret considerations affecting choice, it is surely unwise to accept a decision calculus that violates simple dominance, as the above analysis of a simplistic version of Prospect Theory shows. But more sophisticated decision theories have been developed that explicitly

[7] Kahneman and Tversky (1979) were not unaware of this problem with Prospect Theory. They thus recommended a two-stage decision process, where the first stage would involve removing first-order stochastically dominated options from the option set.

incorporate risk/regret considerations, including "Cumulative Prospect Theory".[8] These theories relax independence while retaining dominance, and thus are not so easy to dismiss. It might be thought that we could mount a Dutch Book argument against these more sophisticated independence-violating theories. But this will not be a fruitful avenue for criticism, because 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, and for similar reasons. In Chapter 2, I discussed this relationship between independence and the "value additivity" assumption of the Dutch Book argument, i.e. the assumption that the sum of fair bets is itself a fair bet. (My discussion draws on the work of Armendt (1993) and Schick (1986).) In short, there is no obvious problem with relaxing the independence axiom while retaining dominance.

It could be said, then, that the case for independence being an inviolable constraint on choice needs some further support. There are indeed some arguments to this effect that make reference to the sequential-choice framework. An analysis of these arguments will require some considerable space, and indeed, this will be my concern in Chapter 6—I consider what, if anything, the sequential-choice framework can reveal about the qualities of the SEU axioms of preference, in particular the independence and ordering axioms. But even if the arguments in support of independence are not entirely conclusive, I want to emphasise that if all other things are equal, it is surely better to retain the axiom. After all, there is a great deal of intuitive appeal to the independence axiom. And the position of Broome and Weirich is precisely that all other things are equal—given that SEU theory can accommodate risk-sensitivity, why move to a theory that relaxes the independence axiom? Recall that this is precisely the issue I want to pursue here: can SEU theory be shown to accommodate the kind of risk/regret attitudes that seem to motivate the search for a

[8] Machina (1989, p. 1631) notes that most of the independence-violating decision theories he lists (the latter five, anyhow) respect simple dominance, provided their component functions respect some reasonable monotonicity conditions. Cumulative Prospect Theory is just one example (and in fact Machina refers to this model as "Anticipated Utility"). I will not detail the model here, but note that it involves the transformation of cumulative rather than individual probabilities. A version of the theory is outlined by Tversky and Kahneman (1992), who credit the earlier work of Quiggin (1982), Schmeidler (1989), Yaari (1987) and Weymark (1981).

less stringent normative theory of choice? If so, there is no need to pursue a relaxation of the independence axiom. If, on the other hand, it is shown that SEU theory cannot accommodate the full spectrum of risk-sensitivity, then we are faced with a choice—we can rule that some kinds of risk-sensitivity are simply irrational, or we can pursue a relaxation of independence, despite whatever arguments there may be against this move. So having described what is at stake, let me return to the question of whether SEU theory can indeed account for risk-sensitivity.

4.6 Savage's theory and the content of outcomes

We can continue to use Allais's problem as the focus for investigations of risk-sensitivity. As mentioned, 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. Interestingly, Broome and Weirich both argue that Savage's theory, to its detriment, doesn't allow such a move. I want to consider the general issue here: Do specific SEU theories (taken in all their detail) constrain the content of outcomes, and in particular, do they prohibit risk/regret properties in outcomes? In such case, we would have a very clear-cut answer to the question of whether SEU theory (or at least the versions of SEU theory in question) can accommodate the Allais-responses. Savage's (1954) theory is a good place to start on this issue, seeing as Broome and Weirich have already argued that it does not permit act outcomes to involve sentiments towards risk.

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".[9] The assumption is not an intuitive requirement of

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

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 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.[10]

Broome (1991, p. 115) states what he thinks are the consequences of this assumption. He considers a lottery, and how we might describe the outcome of losing. If the agent is sensitive to risk or regret, it might seem best to describe the outcome as "receive no money and feel disappointment at not winning". Broome (1991, p. 115) goes on to say why such an outcome is at odds with Savage's SEU theory:

> 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 (as per the "rectangular field assumption") because such risk properties intimately depend on the precise

---

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

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 preclude risk sentiments, but I do not think 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 a continuous utility function (unique up to positive linear transformation) and a corresponding unique probabilistic belief function for an agent. We cannot assume that an agent with such extraordinary discerning powers actually exists. It is impossible for us ordinary mortals to entertain a complete preference ordering over the infinitely rich option space that Savage’s theory requires. 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. To top it off, many state-outcome combinations will not merely be impossible for an agent to achieve, but will be 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”. Surely if the weather is fine at my location I cannot be walking in the rain. 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 do not think Savage’s theory, in particular, rules out risk/regret sentiments from featuring in the description of act outcomes. Moreover, there is even less reason to think that Jeffrey's theory constrains outcomes in this way. There is no "rectangular field assumption" in Jeffrey's theory, and it is made explicit that outcomes include properties of the act—for Jeffrey, an outcome is the conjunction of an act and a state. In fact, Broome and Weirich agree with me on this point; they think it is a virtue of Jeffrey's theory that it can incorporate risk/regret sentiments in outcomes. But not only do I think that Jeffrey's theory shouldn't be singled out in this respect, in my opinion, we should be very cautious about how expansive (or "comprehensive") we want to make the description of outcomes.

4.7 A vacuous decision theory?

I think there are significant reasons to be concerned about an SEU model that allows risk/regret sentiments in outcomes, even if what is at stake is an independence-compatible explanation of the Allais-choices. The move threatens to trivialise SEU 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 SEU model seems ad hoc or overly permissive. If we allow such moves then it is questionable whether we can produce counter-examples against SEU 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.[11] The transitivity axiom would be vacuous if every time it seemed as if an agent were violating it, e.g. their preferences over prospects *A*, *B* and *C* appeared to have the following cyclical form—

---

[11] 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.

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

(recall that "$\succ$" denotes the strict preference relation)

—the preferences could simply be redescribed (in the following way for instance) such that adherence to transitivity is really an assumption rather than a result to be tested:

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

(where *B1* and *B2* are two other basic outcomes)

(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.

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. An agent's preferences, then, are simply assumed to be rational, and the task is to determine the agent's probability (credence) and utility (value/desire) functions such that he/she can be represented as an expected utility maximiser. I do not agree with this account, however, because I think SEU theory is intended to provide a "thicker" notion of rationality than this. In other words, I think SEU theory is intended to provide us with some substantial guidance when it comes to our preferences amongst acts/outcomes. This means that it should be possible for an agent to *fail to comply* with the theory. And there should be a way to distinguish rational from irrational preferences, at least in principal. The debate about risk-sensitivity indicates that we can only make such distinctions if there are restrictions on what can count as an outcome.

Of course, we do not want to constrain the content of outcomes in a way that privileges particular theories of value. SEU 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 SEU theory with whatever ethical constraints seem most reasonable (Colyvan *et al.* to appear). When determining whether an agent is merely rational or coherent, we do not want to take controversial stands on what aspects of the world should be valued, and in what way. We are looking for more minimalist constraints on the content of outcomes.

My particular interest here, of course, is 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. Perhaps the differences between some outcomes are 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 here we are 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. This is the approach that I favour. 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.[12] 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, meaning that it will be possible to determine instances when an agent violates rational principles such as transitivity. 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 there are any kinds of risk/regret properties that can be associated with individual outcomes, or whether any such properties will be "global" properties of an act. If we are not talking about properties of individual outcomes, then it doesn't look like any kind of "property-consistency" criterion can be sensibly applied, and there is no hope for reconciling risk-sensitive choice behaviour with SEU theory. There is reason to think that risk/regret attitudes can be properties of individual outcomes, however. An obvious move is to claim that risk/regret attitudes are simply a certain type of emotional response, either anxiety or excitement about the perceived riskiness of an act, or regret (whether positive or negative) that is experienced when some outcome rather than another actually eventuates. It is likely that expected emotional responses consistently matter to an agent. Most plausibly, the agent will always value excitement and elation favourably, and anxiety and regret unfavourably. In such case, it is quite legitimate for the presence of such emotions to be cause for distinguishing between what would otherwise be identical outcomes.

Jeffrey (1982) discusses an example problem where an apparent violation of

[12] 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.

independence can be explained away by taking into account obvious emotional responses that distinguish between otherwise identical outcomes.[13] 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 (act *A*) or have some champagne and cheese (act *B*). What this example is supposed to illustrate is that it would be very reasonable for Mr Smith to prefer act *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 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 having failed 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 have here then a way to explain the Allais-choices. 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 problem that is commonly put forward.) It is 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. Moreover, even if the parameters in Allais's problem are slightly modified, so that there is no qualitative certain/uncertain distinction, an

---

[13] Jeffrey in fact 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.

appeal to emotional response can still be compelling. It is not inconceivable that an agent may be highly emotionally sensitive to the varying amount of riskiness associated with Allais-type options, whatever probabilities and monetary amounts are at play. In general, SEU theory can accommodate Allais-type choices to the extent that there are tangible risk-related emotions involved, and such a move does not make the theory vacuous.

4.8 Can we take risk/regret sensitivity further?

An explanation of the Allais-choices is all well and good, but Broome and Weirich seek a broader and more robust account of how choice may be affected by risk/regret considerations. They both hold that risk/regret sensitivity need not be linked to a tangible emotional reaction. And incidentally, while I have indicated that it is by no means impossible, 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, particularly when the original probabilities are adjusted in such a way that there is not the qualitative certainty/uncertainty distinction. To the extent that Weirich and Broome try to marry more sophisticated risk/regret sensitivity with the SEU 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 would thus be running the risk of having a vacuous decision theory. Any set of preferences could be defended as rational.

Weirich's (1986) account of how SEU theory can handle risk/regret could also be interpreted as resting on counterfactual properties of outcomes.[14] Weirich allows an outcome to be sensitive to the precise distribution of other possible outcomes associated with the act. Again this seems 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 SEU 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.[15] 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

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

[15] Explicit modifications of SEU theory that involve the subtraction of a variance term (from the expected utility of the act) do not respect first-order stochastic dominance (Machina 1989, p. 1631).

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 SEU theory is maintained. And again, I do not think that specific distributions of outcomes are the right sort of things to count as properties of individual outcomes.

4.9 Risk/regret conclusions

So is redescribing outcomes to include global-property-inspired risk/regret sentiments a legitimate way to accommodate these sentiments within SEU theory, and given our problem of interest, a legitimate way to explain the Allais-choices? I have argued that a possible danger of this move is that it threatens to make SEU theory vacuous. Those who think that SEU theory merely elucidates inevitably 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 SEU theory is to have any normative content. But I don't think we can find these constraints amongst the assumptions of Savage's, or any other, 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. In the Allais case, for instance, an extra property of relief or confidence might distinguish the outcomes associated with the certain prize in problem *A*. In this way, we will be able to explain a lot of apparent violations of independence. All that is required is a plausible story about the risk-related emotional state of the agent, and how such a state consistently affects the agent's evaluations of outcomes.

But some, including Broome and Weirich, seek a more sturdy account of risk/regret sensitivity, one that does not appeal to emotions but rather to more objective properties of outcomes. Both Broome and Weirich can be interpreted as appealing to counterfactual properties of outcomes, where the counterfactuals refer to what alternative outcomes might otherwise have eventuated from the act in question. Counterfactuals offer a way to depict what appear to be "global properties" of an act as properties of individual outcomes. I find this account problematic, however, because such counterfactual propositions do not lend themselves to natural property groupings. Therefore, the move effectively makes SEU theory vacuous.

The upshot of all this, I think, is that if we want to accommodate any kind of systematic risk/regret sensitivity, then we should face up to the fact that we are talking about a violation of independence. This might be a rather predictable conclusion, given my starting position that we should be able to distinguish between SEU theory and alternative accounts of choice under risk/uncertainty. One could say that it is precisely because I do not think that SEU theory should be vacuous that I come to the conclusion that not all kinds of risk/regret sensitivity can be accommodated by the theory. For now, I will leave open the obvious next question as to whether a relaxation of independence should be considered rationally permissible. At stake is the normative status of theories like cumulative prospect theory. Such theories are much more flexible than SEU theory when it comes to determining whether an agent has rational preferences. There is essentially an extra degree of freedom with respect to representing the agent's ordinal preferences—as well as the utility and credence functions, we have a subjective risk factor to play with. Of course, in such case, the agent's ordinal preferences may be consistent with more than one (coherent) credence and value function combination, which certainly complicates the usual representation theorem story. The Dutch book justification of probabilism (which I analysed in Chapter 2) would also be compromised by a relaxation of the independence axiom (because the move would undermine the similar value-additivity-of-bets principle that the DBA depends on).

In brief, relaxing the independence constraint on rational choice has some significant consequences. In Chapter 6 I address this core question; I consider whether the sequential-choice context provides some stronger arguments for retaining the independence axiom. In the meantime, however, I will turn to another example decision problem that is illustrative of a general challenge to SEU theory—Ellsberg's

(1961) problem. There is much in common between the Allais and Ellsberg problems, but the latter raises an additional question—should our representation of belief make a distinction between known risk and uncertain risk, or, more accurately, between "sharp" and "vague" degrees of belief? This is a further dimension to the risk/regret challenge to SEU theory.