On Risk and Rationality

Brad Armendt
School of Life Sciences
Arizona State University
Box 874501
Tempe, AZ 85287-4501
armendt@asu.edu

Abstract

It is widely held that the influence of risk on rational decisions is not entirely explained by the shape of an agent's utility curve. Buchak (Erkenntnis 2013; Risk and rationality, Oxford University Press, Oxford, in press) presents an axiomatic decision theory, *risk-weighted expected utility theory (REU),* in which decision weights are the agent's subjective probabilities modified by his risk-function *r*. *REU* is briefly described, and the global applicability of *r* is discussed. Rabin's (2000) calibration theorem strongly suggests that plausible levels of risk aversion cannot be fully explained by concave utility functions; this provides motivation for *REU* and other theories. But applied to the synchronic preferences of an individual agent, Rabin's result is not as problematic as it may first appear. Theories that treat outcomes as gains and losses (e.g. prospect theory and cumulative prospect theory) account for risk sensitivity in a way not available to *REU*. Reference points that mark the difference between gains and losses are subject to framing, many instances of which cannot be regarded as rational. Yet rational decision theory may recognize the difference between gains and losses, without endorsing all ways of fixing the point of reference. In any event, *REU* is a very interesting theory.

1. Risk-weighted expected utility.

Sensitivity to risk is a pervasive influence on our decision-making. Sometimes we are averse to risk, sometimes we seek it. Is that rational? How should a decision theory represent those influences, and which influences does a rational decision theory condone? Lara Buchak addresses these questions in a formal theory, *risk-weighted expected utility* (*REU*), and the answers she develops are well worth considering. A concise description of central parts of the account is given in (Buchak, 2013); a richer, more detailed exposition will appear in a forthcoming book, (Buchak, in press). Here I will discuss the theory, and along the way make a few more general observations. I think Buchak's work on this topic is excellent—she has developed a preference-theoretic foundation for a general account of risk-sensitivity that encompasses a wide range of possible risk-sensitive attitudes. It is a sophisticated theory, accompanied by a rich and enlightening discussion of related philosophical issues and literature.[1]

[1] Especially as it appears in the book-length treatment. Buchak presented, and I commented on, a much earlier stage of her work at the FEW 2005 in Austin TX. Her theory has advanced far since then, and many of the questions I raised then have been well answered. In particular, the formal foundation of theory has now established the possibility of a globally consistent treatment of risk-functions along the lines she presents, and the theory deals with both risk-aversion and risk-seeking.

I express doubts about one central feature of Buchak's theory of the influence of risk on choice, but I do not argue for any specific rival account.

A motivation for *REU* is expressed in this passage from Buchak (2013, Sect. 2):[2]

> "An agent might be faced with a choice between one action that guarantees that he will get something he desires somewhat and another action that might lead to something he strongly desires, but which is by no means guaranteed to do so. Knowing how much he values the various ends involved is not enough to determine what the agent should do in these cases: the agent must make a judgment not only about how much he cares about *particular* ends, and how likely his actions are to realize *each* of these ends, but about what *strategy* to take towards realizing his ends *as a whole*. The agent must determine how to structure the potential realization of the various aims he has. This involves deciding whether to prioritize definitely ending up with something of some value or instead to prioritize potentially ending up with something of very high value, and by how much: specifically, he must decide the extent to which he is generally willing to accept a risk of something worse in exchange for a chance of something better. This judgment corresponds to considering global or structural properties of gambles."[3]

REU is offered as the normative theory of how to do this. As Buchak makes very clear, *REU* shares features of other formal theories, and in a way synthesizes two other formal results. It is innovative in providing that synthesis, and in providing, as other theories do not, the formal account of risk-sensitivity for a subjective utility theory, where the probability measure p is to be consistently interpreted as subjective degree of belief. For a given agent at a given time, a risk function $r(p)$ characterizes the way that risk influences the value of available actions; since $r(p)$ is a function of the rational agent's subjective probabilities p, it characterizes the agent's attitude toward *perceived* risk. The intended interpretation is that in deliberation, the weights given to outcomes are the rational agent's subjective beliefs, adjusted by the risk function r.

REU treats the agent's risk-sensitivity as uniform over risky choices, in the sense that choices with equal risk profiles (equal p values on their outcomes) are discounted at the same rate; for example, every gamble of the form $\{A, p; B, 1\text{-}p\}$ is discounted by the same $r(p)$, $r(1\text{-}p)$. So the function r, together with the combination rule *REU*, describes how the influences of beliefs (probabilities) are modified in light of their relevance to the set of strategies the decision-maker considers, and among which he chooses. Together with p and U, r provides a fuller characterization of an agent's rational efforts to make the best choice.

The combination rule that yields the *REU* of an act is a cumulative utility theory, which provides a worst-case-up recipe for calculating the effects of r: The utility of an act is the utility of its worst outcome plus the sum of the risk-weighted increments of utility attributable to the possibility of getting the next-worse outcome, and the next-worst after that, and so on. The function r is designed to be such that, in the context of this combination rule, it captures the

---

[2] Unless otherwise specified, references to Buchak's work are references to this paper.

[3] A point that I will raise later: notice the use in the last sentence of 'global.' It contrasts nicely with 'particular', but so might an adjective like 'regional.'

decision-maker's risk-sensitivity to the possible results of the strategy of performing the act. Or, we should say, Buchak's theory establishes the existence of a unique function *r* that, in the context of the combination rule, does that.

Having mentioned the combination rule, let us note that, as Buchak points out, the rule corresponds to a natural way of representing an agent who focuses first on the minimum guarantee: Start with the worst case, the guaranteed minimum outcome, and upgrade the act's utility from there for each successively better possible outcome, with each upgrade consisting of the increment in utility its outcome would provide over the next-worst one, weighted by the risk-adjusted probability of getting that increment. An alternative best-case-down construction, could be given by a different theory (and has been, elsewhere), and it could be seen as a natural way to capture the thinking of agents who focus first on the best possible outcome, and decrement from there. Given that risk functions *r* can be concave or convex, either way can capture risk aversion or risk seeking. And one way can be transformed into the other by modifying the risk function. Buchak makes all of that clear. To this I would just add that after the theory both a) chooses a rule that specifies one way or the other way of doing this, and b) provides a unique global *r* for a given rational agent, it is committed to being able to use one of these ways, not both, to informally capture the agent's thinking throughout the preference system.

2. Rabin's result.

If risk-sensitivity could be adequately represented by the shape of the agent's utility curve in a standard expected utility theory, then there would be no need for a revision of expected utility theory like *REU*. Focus on risk aversion for the moment. One striking argument against the idea that concave utility functions provide adequate representations of risk averse preferences has been developed by Matthew Rabin (2000), and Buchak uses Rabin's result in her discussion of the need for an alternative such as *REU*. Though I do not wish to defend the claim that risk aversion can be represented by concave utility functions alone, I do have a comment on Rabin's calibration theorem.

The basic problem that Rabin points to is this: if agents display a small degree of risk aversion at modest stakes, whatever their wealth level, and if their utility curve represents that risk aversion with appropriate local concavity, it turns out that the curve's larger-scale concavity depicts an extreme and implausible degree of risk aversion for higher-stakes decisions. To use examples from Rabin that Buchak presents: If the rational agent would turn down a fair coin toss yielding either a gain of $110 or a loss of $100, regardless of his initial wealth level, then that agent will turn down a fair coin toss that yields either a loss of $1,000, or a gain of *any amount of money*. An agent whose utility curve represents a similar pattern of turning down tosses that yield either a gain of $1,050 or a loss of $1,000 will turn down a coin toss that yields a loss of $20,000 or a gain of $*x*, *no matter how large x is*. If risk aversion is to be represented by a concave utility function, modest-stakes choices that exhibit some minimal degree of risk aversion at various wealth levels imply implausibly extreme degrees of risk aversion in decisions where stakes are higher.

This is significant for treatments of risk in economic models. Suppose we are interested in a population of agents in a market, interacting over some period of time, who may have a variety of present or lifetime wealth levels, and we have a utility curve whose shape is thought to capture risk-aversion for all of them, or for some typical member, during that time. If the curve is such that for *whatever the wealth level*, risk aversion for moderate stakes generates some threshold level of local concavity, it will turn out to have large-scale concavity that predicts implausibly extreme risk aversion at high stakes.

But if instead we are interested in the risk-aversion of a single decision-maker at a given time, we should focus on his preference system and utility function at that time; we can expect that at later times his preference system will have evolved. Applied to a single utility curve representing my preferences at a given time, the condition corresponding to the requirement that moderate-stake risk-aversion exceeds some threshold at all wealth levels would be something like this: moderate-stake risk-aversion is a feature of my current preferences about choices I would make, whether at my actual wealth level, or at any number of counterfactual wealth levels, different from my actual wealth.

So perhaps I would turn down a fair coin toss yielding either +$110 or -$100; I judge that the risk of losing outweighs the prospect of gain. Do I also judge that if I had $100,000 more than I do, I should turn down the coin toss? That is far from obvious. Personally, I would take that bet.[4] And if I would, if there is no minimum threshold of risk-aversion that I display concerning all such counterfactual situations, then the condition needed to guarantee that Rabin's result applies to my preferences is not satisfied.

Rabin's result also implies that implausible results arise when the threshold level of concavity only holds over a significant range of wealth levels, not all the way up. For example (Rabin's again), turning down a fair gamble on -$100 or +$105 at all wealth levels up to $350,000 leads, at the $340,000 wealth level, to turning down a fair gamble on -$4,000 or +$635,670. But a range of counterfactual situations large enough to generate the implausible results likely includes situations in which, according to my present preferences, the threshold is not met.

As I indicated, I do not make this point in order to claim that concave and convex utility functions are all we need to represent risk aversion and risk seeking. But I suggest that, for subjective *EU* in rational choice theory for individuals, the situation is not as dire as Rabin's result may at first make it seem.

3. On the global character of risk-sensitivity in *REU* theory.

REU is offered as an improved theory of rational choice, compared with *EU* theory. Buchak argues that risk-sensitive preferences may be incompatible with principles of standard *EU* theory, but nevertheless be rational. So compared with standard *EU* theory, *REU* theory relocates the boundary between rational preferences and choices and irrational ones. The main issue I will raise here concerns the global character of risk functions $r(p)$ in the *REU* theory.

[4] So I now judge. Notice, though, that I can judge that, conditional on my being a lot richer, I would take the gamble, and at the same time think that if I were lot richer, my preferences may shift in a way that would lead me to then reject the gamble.

To a given set of rational preferences that satisfy its axioms, *REU* attributes *U*, *p*, and a unique risk function *r*. *r* is a function of *p*, one that applies to values of *p* in whatever context they appear—for example, whether the stakes are high or low. In other words, agents whose preferences satisfy *REU* display a characteristic sensitivity to risk, common to their high-stakes and low-stakes choices, and to choices very favorable (in the sense that the worst outcome is very good) or less favorable (in the sense that it is not). A question that naturally arises, I think, is whether there are good normative reasons in favor of the unique, global *r*.

Consider a fanciful analogy with norms of motoring. In Ohio, let us say, the legal requirement is that you must obey the posted speed limit. But imagine that in California, the law allows many ways of deviating from the posted limit, as long as you do it in the same way on all occasions. (Same way = same function of the posted limit; many possible functions.) The California requirement is more lenient; it allows a wider variety of driving behavior than does the Ohio requirement. But it surely elicits the question, why the same way on every occasion? What is the normative justification for that? Why not Nevada's requirement, which requires, let us imagine, that your driving be appropriately inspired by the posted limit, as fits the occasion?

EU is what we get from *REU* when *r* is the identity function $r(p) = p$. So *REU* is a genuine generalization, subsuming *EU* as a special case. Its requirements count a wider range of preferences as rational, it allows more latitude in how a rational agent's preferences may be arranged. So it may seem inappropriate to ask, why does *REU* impose a norm of global uniformity on risk-sensitivity? Inappropriate because what it revises, *EU*, does the same thing, even more stringently (allowing only one trivial *r*). Still, once we attend to the requirement it is natural to ask, why is it a norm of rational preference that an agent's risk-sensitivity is globally fixed in this way?

The question I just asked is about the normative *result* that the formal theory provides. A great virtue of providing the formal theory is that we can look for an answer in the *axioms*. So let me ask a corresponding question, intended not as a challenge, but a request for help: What is the axiomatic source of the unique, global *r* function? Buchak provides a good discussion of what *makes room* for the *r* function in the theory: the restriction of tradeoff consistency to sets of pairwise comonotonic acts (comoncones), rather than requiring tradeoff consistency over all acts.[5] But the axiomatic source of the uniqueness of the global *r* is more obscure (to me, at least), and more guidance would help us judge the normative plausibility of that axiomatic source, as opposed to possible alternatives.[6]

---

[5] Buchak (2013, Sect. 5).

[6] Buchak (in press) suggests that her Axiom 1, the comparability axiom, would be in question, since if gambles in different domains subject to different regional *r* functions (as I describe it) were comparable, what would be the explanation of different action-guiding norms in different domains? But a) comparability does not seem to be the entire source of the unique global *r*, though weakening comparability may well undermine it, and b) it is not obvious that a norm permitting different risk sensitivity values in different regions of preference amounts to having multiple norms.

Consider an illustration, simple in one sense, in that each act has only two outcomes. It is less simple than it might be, in that both positive and negative outcomes are in the picture. There are two gambles, G_1 and G_2. I describe their payoffs in terms of utility, though I also offer cues concerning what those utilities might measure.

> G_1 offers me a chance at +10 utiles with probability p, and a chance at -10 utiles with probability 1-p. Think of +10 utiles as something like the value of receiving an unexpected \$100. If I am risk sensitive, my ranking of G_1 will be influenced by a risk function r.
>
> G_2 offers me a chance at plus one billion utiles with probability p, and a chance at minus one billion utiles with probability 1-p. One billion utiles is really good, and minus one billion utiles is really bad. Think of minus one billion utiles as something like slow death by torture, or brutal incarceration for life. If I am risk sensitive, my ranking of G_2 will be influenced by an r function. *REU* requires that it be the same function that captures my risk-sensitivity about G_1.

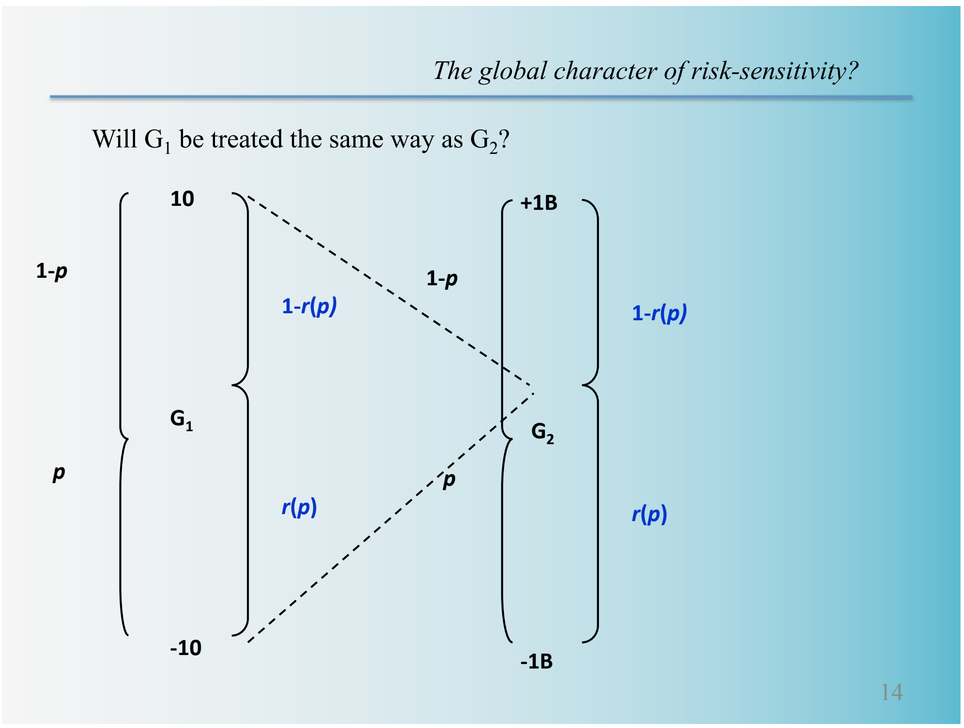

Figure 1 Are G_1 and G_2 treated in the same way?

Since the payoffs are given in utility, if there is a difference in my risk sensitivity, it seems that we cannot explain it by the shape of my utility curve over \$ or torture. If I satisfy *REU*, r is fixed, and my risk sensitivity will influence the utility of each gamble in equal proportions; it does not matter how large the interval between the good and bad outcomes are, or where they are located in my utility ranking. (Note that the figure is intended to display the equal weightings of p in the two gambles; not indifference between the gambles.)

Is that correct? Would it be irrational of me to be more cautious, more risk-sensitive concerning G_2 than G_1?

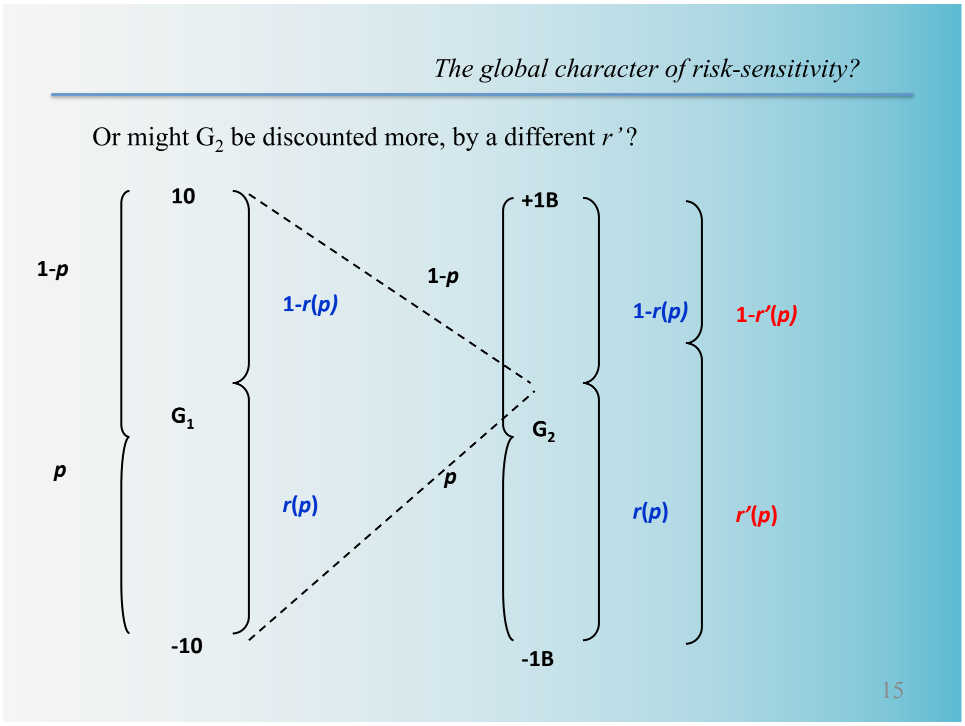

Figure 2 Or is G_2 discounted more, by a larger $r'(p)$ on the unfavorable outcome?

The question for *REU* is whether it is a mistake for risk-sensitivity to differ in decisions like these. Or is risk-sensitivity a more regional matter than *REU* allows?

4. Gains and losses.

There is strong empirical evidence that gains and losses do not weigh equally in our deliberations. This is a descriptive point. Decision theories that treat outcomes as prospects, as deviations from the status quo, are an alternative to theories that treat outcomes as total wealth levels, or complete worlds. By far the best known of such theories are Kahneman and Tversky's (1979) prospect theory, and its successor designed to cover many-outcome choices, cumulative prospect theory (Tversky and Kahneman 1992, Wakker and Tversky 1993). Prospect theory is first of all intended to be a good descriptive theory, and it is designed to represent many features of our decision-making, including some that would be very hard to accept as rational. One of its key features is the asymmetry between the prospect of gain and the prospect of loss.

Gains and losses matter to risk-sensitivity; we are risk-seeking with regard to small probabilities of gain (we buy lottery tickets from profit-seeking governments), risk-averse with regard to small probabilities of loss (we buy insurance from profitable companies), risk-averse with regard to larger probabilities of gain (we take the sure \$490 over a 50:50 chance of \$1000), and risk-seeking with regard to larger probabilities of loss (we take the 50:50 chance of losing

$1000 over the sure loss of $490). These effects cannot all be simultaneously captured by a given risk-function *r*, used with a consistent build-from-the-worst-case-up combination rule. Buchak does not seek to subsume them into *REU*.

Now what counts as a gain or a loss depends on a point of reference, and another feature of deliberation that Kahneman and Tversky famously explored are the effects of *framing*. Framing of outcomes involves characterizing them in comparison to a point of reference, and it is well established that as a descriptive matter, we are very flexible, and influenced by many things, in doing this. Different descriptions of problems can lead us to choose points of reference that differ, to treat hypothetical points of reference as the status quo, and to reverse our choices when arguably everything about the problem is the same except for the way in which it is described. I do not suggest that it is rational to do that. But the idea that a reference point can matter to us, and to our sensitivity to risk, need not commit us to the view that any way we choose it is rational. The irrationality of some combinations of evaluations to which we are led by framing effects need not imply the irrationality of treating risks of gains and losses differently.

The influences of a point of reference on risk-sensitivity are pervasive. They also seem relevant to the recommendation Buchak made in the passage I quoted earlier, that a rational agent can and should legitimately care about the structure of his attainable outcomes, about the strategy he chooses, and not just about the values of the outcomes considered individually. After all, a strategy has a starting point. It is reasonable to think that the point from which a strategy is executed is a significant part of it, and that the starting point is relevant to the assessment of the strategy's value within the decision structure at hand. So we might think that the general motivation Buchak offers for rational risk-sensitivity arising from probability-weights applies equally to risk-sensitivity tied to a point of reference.

If this is correct, then *REU*'s unique global risk function cannot capture some risk-sensitivities of an agent who exhibits all of the commonly observed patterns mentioned above. Unless we rule out some of those patterns as irrational, *REU* thereby falls short of the goal of providing a full account of the influence of perceived risk in rational decision-making. I will not attempt to judge here the normative credentials of the patterns of risk-sensitivity in question, but it is an issue that an advocate of *REU* should address.

5. Conclusion.

REU and the theorem that backs it up demonstrate the tenability of accounting for at least part of the influence of risk on rational deliberation with a nonlinear weighting of probabilities within subjective *EU* theory. *REU* is an important contribution to the pursuit of a normative subjective decision theory for more-realistically-modeled agents. Perhaps we suspect, as I have suggested, that there are further sources of rational risk-sensitivity. I do not want my remarks to be seen as saying, "Well this theory is cool, but why isn't there even more?" Instead, I would point out, *REU* theory moves the boundaries of what is counted as normatively acceptable when it comes to preferring and deciding. The main issue I raise is whether there are good *normative* reasons for setting the boundaries exactly where it does.

Buchak's work on this topic is excellent and enlightening. Philosophers interested in rational choice theory, and particularly in the influence of risk on rational decision-making, will find it well worth reading.

Acknowledgements

Thanks to Lara Buchak for the very interesting work I discuss here, for providing a draft of her book (Buchak, in press), and for helpful conversations about *REU*. Thanks to Jim Joyce, Klaus Nehring, Wlodek Rabinowicz, and other participants in the Formal Epistemology Workshop 2012 hosted by the Center for Mathematical Philosophy at the LMU for discussions related to this paper. Special thanks to Wes Anderson for very helpful conversations and comments on the paper.

References

Buchak. Lara (2013). Risks and tradeoffs. Erkenntnis, this issue.
Buchak, Lara. Risk and Rationality. (Oxford: Oxford University Press), (in press).
Kahneman, Daniel & Tversky, Amos (1979). Prospect theory: an analysis of decision under risk. Econometrica, 47, 263-291.
Rabin, Matthew (2000). Risk aversion and expected-utility theory: a calibration theorem. Econometrica, 68, 1281-1292.
Tversky, Amos & Kahneman, Daniel (1992). Advances in prospect theory: cumulative representation of uncertainty. Journal of Risk and Uncertainty, 5, 297-323.
Wakker, Peter & Tversky, Amos (1993). An axiomatization of cumulative prospect theory. Journal of Risk and Uncertainty, 7, 147-176.