Higher-Order Uncertainty

Here is an example of the standard decision-theory-illustrating table that we all know and love[1]:

	S1 (0.6)	S2 (0.2)	S3 (0.2)	Total EU
a1	10	2	4	7.2
a2	4	8	6	5.2
a3	2	6	50	12.4

I am fairly confident of the dictum that I relay to my undergraduate students on a termly basis:

> **MAXEU**: When faced with the standard decision-theory illustrating table, what I should do is *maximize expected utility* – choose the act with highest expected utility.

But when I say "fairly confident", I mean exactly that. I'm not *certain* of MAXEU. I don't give this proposition credence 1. That is partly because a good Bayesian such as myself reserves credence 1 for logical truths, and MAXEU is not a logical truth. But it is also because there are rival theories of how I should respond to the standard decision-theory-illustrating table, many of which have arguments in their favor that are sufficiently compelling for me to give them some credence. For example, I have some credence that I should maximize risk-weighted expected utility (Buchak 2014) rather than regular expected utility. But I recognize that the truth of this alternative theory is inconsistent with the truth of MAXEU. So I cannot coherently assign credence 1 to the latter.

I also have *some* credence – not a lot, but *some* – in a theory that we might call "Do What Is In Fact Best" (or DWIB):

> **DWIB**: When faced with the standard decision-theory illustrating table, I should do whatever *is in fact best*, regardless of my credences.

DWIB is the view that my credences are irrelevant to the question of what I should do. This is because I actually *am* in just one state of the world, not any of the others in which I have some credence. One act is actually best for me, and this is what I should do. On one possible version of DWIB, the best act is that which maximizes actual utility. On another possible version, my *utilities* are also irrelevant to the question of what I should do; this version of DWIB posits objective criteria of prudential value, which yield an ordering over outcomes in terms of their goodness for me in a manner that does not depend on my preferences over these outcomes (cf., for example, Hurka 1993; Sen 1992, 1993). This theory then claims that, when faced with the standard decision-theory-illustrating table, I should perform whichever act will yield the most objective prudential value for me in the actual state of the world, ignoring the table entirely.

---

[1] This table features random numbers for the sake of illustration.

DWIB might seem like an odd view of what to do when faced with the standard decision-theory-illustrating table, especially in the second of its two versions. After all, why would a theory of what to do when faced with the standard decision-theory-illustrating table recommend *ignoring the table entirely*? But this theory is the prudential analogue of a view about the irrelevance of *moral* uncertainty that has recently received some sophisticated defenses (see especially Weatherson 2014, Harman 2015). I think the arguments for this view about moral uncertainty are sufficiently compelling for me to give the view *some* credence. And I think that the analogy is close enough to also lend some support to DWIB.

This puts me in an interesting predicament. I'm uncertain about what the world is like, and also uncertain about what I should do when I'm uncertain about what the world is like. But the truth or falsity of MAXEU is itself a state of the world; it is a way the world is *normatively*. So my state of uncertainty about what I should do when I'm uncertain about what the world is like is itself a state of uncertainty about what the world is like. It can be represented in a higher-order decision-table, like so:

	MAXEU (0.85)	Risk-weighted MAXEU (0.1)	DWIB (0.05)	Total
a1	7.2	5.44	?	?
a2	5.2	4.4	?	?
a3	12.4	5.28	?	?

Three questions now arise. First: how do I fill out the "DWIB" column? Second: how do I then fill out the "Total" column? Third: once I have finished drawing my table, what should I *do*?

This paper explores a range of possible answers to these questions, identifies problems with each set of answers, and thereby raises a challenge to MAXEU. The challenge is to formulate a version of MAXEU that intelligibly applies to higher-order uncertainty about whether this theory is itself true, such that it (a) generates clear verdicts as to what I should do (b) in a way that is sensitive to my credences. I first show that simple MAXEU is unable to generate clear verdicts for agents – like me – who have any credence in DWIB and who are also non-normatively uncertain, as these agents cannot fill out the DWIB column of their higher-order decision table, and it is impossible to calculate expected utility if one of the columns in one's table has no values in it. I then consider a more complex version of MAXEU that places upper and lower bounds on each act's possible degree of higher-order expected utility. I show that this version thereby generates clear verdicts in *some* of the cases in which simple MAXEU fails, but not *all* such cases. Lastly, I consider a yet-more-complex version of MAXEU that repartitions logical space so as to represent my normative and non-normative uncertainty together in a single decision-table. I show that this view generates clear verdicts in *all* cases, but that it does so in such a way that the amount of credence I invest in DWIB as opposed to MAXEU ceases to ever make any difference to what I should do. This view therefore secures clear verdicts at the expense of sensitivity to my credences. I conclude with skepticism as to whether any possible higher-order version of MAXEU can secure both of these two features – the two features that provide key motivations for the theory.

WORD COUNT: 999 words.

REFERENCES

Buchak, Lara (2014). "Risk and Tradeoffs". *Erkenntnis* 79: 1091-1117.

Harman, Elizabeth (2015). "The Irrelevance of Moral Uncertainty". *Oxford Studies in Metaethics* 10: 53-79.

Hurka, Thomas (1993). *Perfectionism.* Oxford: Clarendon Press.

Sen, Amartya (1992). *Inequality Re-examined.* Oxford: Clarendon Press.

Sen, Amartya (1993). "Capability and Well-being," in Nussbaum and Sen (eds.), *The Quality of Life,* Oxford: Clarendon Press, pp. 30–53.

Weatherson, Brian (2014). "Running Risks Morally". *Philosophical Studies* 167(1): 141-163.