Weak rationality and imprecise choice

Seamus Bradley

seamusbradley.net

seamus@seamusbradley.net

May 29, 2013

There is a growing body of literature that suggests people are dissatisfied with the standard Bayesian insistence on precise values for degree of belief for every possible proposition. Vague or imprecise belief is a popular way to sidestep this problem. Despite its popularity, imprecise credence has yet to be provided with a fully satisfying decision theory. I explore decision rules for imprecise believers. I argue that we cannot hold imprecise decision to the same strong standard of rationality as better informed decisions can be held to.

1. Imprecise belief

We want to explore what rationality requires of your choices. What choices seem good or bad depends on your beliefs and your values. As a first approximation we could say that your degrees of belief are represented by numbers that capture the strength of belief. The stronger you believe a proposition, the higher the number we give it. But consider the following delightfully odd example from Adam Elga.

> A stranger approaches you on the street and starts pulling out objects from a bag. The first three objects he pulls out are a regular-sized tube of toothpaste, a live jellyfish, and a travel-sized tube of toothpaste. To what degree should you believe that the next object he pulls out will be another tube of toothpaste? Elga (2010, p. 1)

It seems strange to require that you have some specific sharp degree of belief in such a proposition. This motivates the idea that we should allow degrees of belief to be imprecise, unsharp. I present a formal version of this position in the next section.

Such imprecise probabilities are becoming increasingly popular,[1] but they have also been criticised on a number of grounds. One important problem with imprecise probabilities at the moment is that they haven't really been given a satisfying decision theory. That is, how imprecise probabilities inform choice hasn't been fully worked out. I survey some of the literature in this area, and attempt to provide a framework for assessing further work.

My conclusions are modest. I argue that there are certain rational constraints on imprecise choice, but these constraints don't always give us the same kind of determinate advice that the choice rule for precise credences does. I argue that this is not a flaw in imprecise credence: it's merely a reflection of the difference in epistemic position between imprecise and precise credence. I argue that rationality requires *no more* than certain constraints on choice: rationality needn't determine choice.

Some criticisms of imprecise choice rely on distinctively *sequential* or *diachronic* choices (Elga 2010). I won't have the space to go into this issue here.[2] The choices I consider are choices that take place at a time. As we will see, there is plenty to say even about this simple case.

2. Preliminaries

This section collects the formal machinery I will need to discuss imprecise choice in the next sections. We are interested in choice, and so, we must have a formal model of choice. A *choice situation* or *decision problem* consists of three elements. There are the objects of value, the objects of belief and the objects of choice. This is "object" in a grammatical sense: these are objects in that they are the object of a verb. In the current context, it is the decision maker who is the subject of those verbs. For example, the objects of value are the things the subject values. I don't mean to suggest that all objects of value are *material objects.*

2.1. Value, Belief, Choice

We are operating here with a fairly standard decision theoretic framework. The aim is to modify one particular element – the credence function – and see how that modification changes things.

We have some objects of value: a set of outcomes – $\mathcal{O}$ – which I am going to assume are real numbers. These represent monetary values. I am going to sweep a great deal of debate under the carpet here by assuming constant marginal utility for money, and risk neutrality. Debates about risk aversion or diminishing marginal utility of money

[1] Advocates of imprecise credences include R. Bradley (2009); Joyce (2005, 2011); Kaplan (1996, 2010); Kyburg (1983); Kyburg and Pittarelli (1992); Levi (1974, 1985, 1986); Sturgeon (2008). For more formal treatments of the position see Cozman (n.d.); Halpern (2003, 2006); Walley (1991, 2000).

[2] But see S. Bradley and Steele (ms.); Steele (2010).

are besides the point of the current debate, and arguably orthogonal to it. Thus, making these (fairly standard) assumptions doesn't bias the following discussion in any way. The objects of belief are built up from a set of mutually eclusive and exhaustive states that form the set Ω. We take sets of states to be *events*. Thus, you have beliefs about elements of 2^{Ω} The objects of choice are acts. Elements of $\mathcal{A}$ – the set of acts – are functions from 2^{Ω} to $\mathcal{O}$.

Let's say we have acts φ and ψ. Say we have some kind of random device that outputs a 1 with probability p and a 0 otherwise. $p\varphi + (1-p)\psi$ is the act "get whatever φ gets you with probability p, get whatever ψ gets you otherwise".[3] Anscombe and Aumann (1963), and von Neumann and Morgenstern ([1944] 2004) allow themselves this kind of mixed act. There might be some subtlety attached to how mixed acts are generated, but I don't think this will impact on my current project (Schwarz ms.). If A is a set of acts, $pA + (1-p)\psi$ is the set of acts of the form $p\varphi + (1-p)\psi$ for $\varphi \in A$.

We define a *probability function* over 2^{Ω} as a function $\mathbf{pr} : 2^{\Omega} \to \mathbb{R}$ with the following properties:

- $\mathbf{pr}(\emptyset) = 0$ and $\mathbf{pr}(\Omega) = 1$
- $\mathbf{pr}(\emptyset) \leq \mathbf{pr}(X) \leq \mathbf{pr}(\Omega)$ for all X
- $\mathbf{pr}(X \vee Y) + \mathbf{pr}(XY) = \mathbf{pr}(X) + \mathbf{pr}(Y)$ for all X, Y

In the standard precise setting, we would use a probability function to represent your belief. So you would have some function **pr** that represents your belief and assigns numbers to the events. Bigger numbers mean stronger belief in that event. Consider the simple case where we have beliefs about a fair die. The states Ω would be the possible outcomes – $\{1,2,3,4,5,6\}$ – and the events would be sets of such outcomes. For example the event "The die lands even" is the set $\{2.4.6\}$.

We are interested in choice with imprecise probabilities – what I call *imprecise choice* – so we represent your degree of belief by a set of probability functions. Let's imagine that instead of our well behaved die, we are considering bets on a coin of unknown bias.[4] No single probability function represents your state of uncertainty with respect to this coin. There is, however, a *set* of probability functions that represents your belief. Call this set $\mathcal{P}$. With a little abuse of notation, we can define a function $\mathcal{P}(H)$ which maps H to the set of values that the probability functions in $\mathcal{P}$ give to H. So $\mathcal{P}(H) = \{\mathbf{pr}(H) : \mathbf{pr} \in \mathcal{P}\}$. We can then define $\overline{\mathcal{P}}(H)$ and $\underline{\mathcal{P}}(H)$ as the min-

[3]If the outcomes of φ and ψ are monetary amounts, one might instead interpret the mixed act as the act that wins you $p\varphi(X) + (1-p)\psi(X)$ when X obtains. This corresponds to betting £0.50 on each of two horses. But this option isn't available if the goods that the acts win aren't "divisible" in the way money is. Nothing in what follows hangs on this difference anyway.

[4]The phrase "coin of unknown bias" is a little awkward, since there's no sense in which what is unknown is a *bias*. Bias with respect to what? The coin could only said to be biased if we had a pre-existing idea that the coin *should* land tails as often as heads. We should perhaps speak of a "coin with unknown chance of heads". Or if chance-talk is not allowed, "coin with disposition to generate unknown statistics". Hykel Hosni was helpful in pressing me on this point.

imal and maximal values that the probabilities in $\mathcal{P}$ assign to H.[5] These "summary statistics" give us objects that somehow represent the belief and are easier to handle than the full set of probability functions. One should always be aware, however, that it is $\mathcal{P}$ – the set – not $\mathcal{P}(\cdot)$ – the function – that represents the belief. Following van Fraassen (1990), we call this your *representor*. It is sometimes helpful to interpret $\mathcal{P}$ as a *credal committee*.[6] Each $\mathbf{pr}$ in $\mathcal{P}$ is a member of a committee in charge of your actions. Collectively, the committee must decide how to act. I should note here that, despite being the somewhat standard terminology, "imprecise probability" doesn't really do justice to the phenomenon of interest. As Steele (2007) argues, it would be better to speak about *indeterminate probability*. Despite this, I shall stick with the standard terminology.

We define your *expectation* $\mathrm{E}_{\mathbf{pr}}(\varphi) = \sum_{w \in \Omega} \mathbf{pr}(w)\varphi(w)$. That is, your expectation – or expected value – for an act is a weighted sum of what the act gets you in each state, weighted by how likely you consider that state.

We can define an imprecise expectation by taking the set of the expectations for each $\mathbf{pr} \in \mathcal{P}$. That is, $\mathcal{E}_{\mathcal{P}}(\varphi) = \{\mathrm{E}_{\mathbf{pr}}(\varphi), \mathbf{pr} \in \mathcal{P}\}$. We often drop the subscript and just talk about $\mathcal{E}$ when it is obvious what $\mathcal{P}$ is at issue. We can define $\underline{\mathcal{E}}(\varphi)$ and $\overline{\mathcal{E}}(\varphi)$ as the smallest and largest expectations assigned to φ by members of $\mathcal{P}$. How are we to choose with imprecise expectations? The first thing to note is that we can't simply "choose the biggest". The $\mathcal{E}$s for the various acts will typically be sets of numbers: there's no obvious sense in which one collection of numbers is *bigger* than the other. The sets can overlap. So we need to think a little more carefully about what imprecise choice involves.

In all of the examples we will consider in what follows, we will be concerned with acts that are effectively bets on events of various kinds of simple chance set ups. There will be two kinds of chance set ups: those where you know the probability. These will be bets on a fair die. We assume that your credence in the die landing on an even number $\mathcal{P}(E) = \{0.5\}$. There will also be bets where you don't know the probability: these will be the cases where you must represent your uncertainty by a set of probabilities. These will be bets on a coin of unknown bias. We assume throughout that your credence that the coin lands heads is $\mathcal{P}(H) = [0, 1]$.

2.2. Choice functions

The main object of study in this paper will be various forms of *choice function*. A choice function will take a set of available acts and output a subset of choiceworthy acts. A choice function is a function $\mathcal{C}\colon 2^{\mathcal{A}} \to 2^{\mathcal{A}}$ such that for all $A \subset \mathcal{A}$ we have $\mathcal{C}(A) \subseteq A$ and $\mathcal{C}(\mathcal{C}(A)) = \mathcal{C}(A)$. That is, the function outputs a subset of the acts available: it would be unhelpful if the choice function gave you the advice to perform some act that wasn't available to you. We also require that the choice function is sta-

[5]More properly, these should be the greatest lower bound and the least upper bound, since we aren't sure that the extrema are attained. Nothing hangs on this.

[6]This is how Joyce (2011) talks, though he attributes the term to Adam Elga.

ble in a certain sense. That is, applying the function a second time has no effect. Call the set that the choice function outputs – $\mathcal{C}(A)$ – the *choice set*. The majority of this paper will be about what properties we can impose on choice functions, and which of those properties it is reasonable to demand in the imprecise case. We will explore some well-known imprecise choice functions and discover which properties they do or do not satisfy.

$\mathcal{C}(A)$ is meant to represent or encode what it is that rationality requires of you when you must make a choice among the members of A. There are many ways of interpreting $\mathcal{C}(A)$.[7] A "Strong" interpretation would say that acts in $\mathcal{C}(A)$ are all equally the best act: there is nothing to choose between the acts in $\mathcal{C}(A)$ and you should be equally happy to take any of them. $\varphi \in \mathcal{C}(A)$ is here considered an endorsement of act φ. A weaker interpretation might be to say that all the acts in $\mathcal{C}(A)$ are better[8] than the acts not in $\mathcal{C}(A)$. This interpretation does not preclude there being strict preference between the acts in $\mathcal{C}(A)$. $\varphi \in \mathcal{C}(A)$ isn't now such a strong endorsement of φ; but $\psi \notin \mathcal{C}(A)$ is still considered a real flaw in ψ. An even weaker interpretation would be just to say that $\mathcal{C}(A)$ includes the best act, but to make no further assumptions about how acts in $\mathcal{C}(A)$ compare to those outside $\mathcal{C}(A)$ or among themselves.

In short, we can think of standards of rationality as giving sufficient conditions for being acceptable, or we can think of the standards of rationality as giving necessary conditions for being acceptable. The former accords with the positive understanding of rationality: endorsing elements in $\mathcal{C}(A)$. The latter accords with the negative understanding of rationality: those elements outside $\mathcal{C}(A)$ are advised against.

I think that at least some of the scepticism about particular imprecise choice functions can be defused by being clear about what kind of interpretation we give the function. In precise Bayesianism, we can always take the strong interpretation of the choice function, but this interpretation might not always be applicable in the imprecise case.

Obviously, a weaker kind of choice function is less useful, but it may be that that is all we can achieve in this context. Rationality can provide necessary conditions for an act to be choiceworthy – or rather, sufficient conditions for an act to be unchoiceworthy – but it is harder to see why rationality should always provide necessary and sufficient conditions for choiceworthiness. One criticism that has been levelled at the imprecise probabilities approach is that imprecise choice functions will fail to always determine an act that was optimal: they might fail to give you any advice (Williamson 2010, p. 68–72). To criticise imprecise choice for this flaw is to hold the theory to too high a standard. In cases of severe uncertainty, such determinate answers *shouldn't* always be available. Consider Elga's toothpaste/jellyfish example: what bets should you take on the next item's being a tube of toothpaste? I would be suspicious of a choice rule that purported to give a fully determinate and perfectly precise answer to

[7] A talk at the LSE Choice Group seminar by Conal Duddy on "shortlisting" inspired much of this discussion of interpretations of choice sets.

[8] Note that such "betterness" needn't determine an order on the acts. Consider the case where φ is better than ψ just in case that φ doesn't have some obvious flaw that ψ does. A choice rule that returned the set of acts without this flaw would be an example of this weaker sort of choice rule.

that question. I will argue that no choice rule that admits of a strong interpretation satisfies all of the properties we want a choice rule to satisfy. Thus, I argue, rationality only requires so much: rationality needn't determine choice.

If you knew whether or not X, then you could pick the act that would do best. That is, by the lights of your epistemic state, you could recognise which act will be best. However, if you only had a (precise) probability for X, you couldn't do that. You can, however, pick the act that will do best "on average". On average, by the lights of your credence, that is. The move from full belief to (precise) partial belief requires a shift in attitude to decision. You can't demand that you pick the act that will get you the best outcome any more; you can only demand that you pick the act with the best *on average* outcome. That is, you need to move from maximising actual outcome to maximising *expected* outcome. The move from precise to imprecise probability requires a similar shift in focus: you need to be more modest in your ambitions with respect to choice. You can no longer tell which act will do best on average, you can still tell that some acts won't do best on average. You cannot demand that you pick the best on average outcome any more. But choice is still constrained: you can rule out some acts as flawed. Criticising imprecise belief for this weakness of its decision theory is like criticising precise partial belief for failing to pick out the best act every time. There is no way to give a reasonable *Strong* choice rule for imprecise decision. Rationality only requires that you satisfy certain constraints; rationality does not require that your choice is always determined.

Perhaps it's best to keep two projects separate. First we want to know what rationality demands of imprecise decision; second we want to know how we should act in cases of imprecision. The answer to the former question might not fully determine an answer to the latter; in fact I will argue that it doesn't. In the precise credence case, it might do however. But in the cases where it doesn't, we still need an answer to the second question. We will have to accept that the answer might not be fully rational. That's not to say that it will be *irrational* – violating rationality – but just that it will be *arational*: without rationality. Rationality can only get us so far. I focus here on the first project. That is, I explore how imprecise choice ought to be rationally constrained; we should try to get comfortable with the fact that those constraints might fail to determine choice.

2.3. How to constrain imprecise choice

What does a reasonable imprecise choice rule look like? There are many places in the literature where enterprises like this this have been developed. There are a great many ways we could approach the question of how best to settle on an imprecise decision rule. I survey some ways here.

One major source of inspiration is the literature on decision under ignorance. This will suggest some kinds of choice rule, as well as some of the conditions for reasonable imprecise choice. The most important sources here will be Milnor (1951) and Luce and Raiffa (1989).

Another important inspiration will be work on social choice theory: how to aggregate many individuals' preferences into a single group choice. If we think of each probability in your representor as a member of a credal committee that has to vote on what you should do, then the parallel between imprecise decision and social choice becomes clear. Here I will draw on Arrow's theorem (Gaertner 2009) and the work of Amartya Sen (Sen 1970, 1977).

There are two ways one might frame the discussion: in terms of an ordering over the acts (Arrow, Milnor), or in terms of a choice rule (Luce and Raiffa, Sen). I will talk in terms of choice rules, but we will see that relations will also play an important role.

There are several ways we could describe conditions on the choice function. One is just to put conditions on the functional form of the choice function. That is, we could impose intuitive conditions on the function with respect to how it interacts with unions and intersections of sets of acts. For example, is it reasonable to demand that if $B \subseteq A$ then $\mathcal{C}(B) \subseteq \mathcal{C}(A)$? It turns out the answer is no.[9] But that's an example of the kind of condition we might think to impose on a choice function.

There is another way we might want to impose constraints on reasonable choice functions. This is by restricting various kinds of relation associated with the choice function. Various properties of relations that it will be useful to appeal to are described in Appendix A.

For this, we need some definitions. A choice function $\mathcal{C}$ *pairwise satisfies* a relation $\succeq$ when, for all $\varphi, \psi \in \mathcal{A}$:

- If $\varphi \succeq \psi$ then $\varphi \in \mathcal{C}(\{\varphi, \psi\})$
- If $\varphi \succ \psi$ then $\{\varphi\} = \mathcal{C}(\{\varphi, \psi\})$

If $\succeq$ is understood as preference relation then pairwise satisfying a relation means never picking a dispreferred option in pairwise choices. A choice function $\mathcal{C}$ *satisfies* a relation $\succeq$ when, for all $\varphi, \psi \in A \subseteq \mathcal{A}$:

- If $\varphi \succ \psi$ then $\psi \notin \mathcal{C}(A)$
- If $\varphi \sim \psi$ then $\varphi \in \mathcal{C}(A) \Leftrightarrow \psi \in \mathcal{C}(A)$

Satisfying a relation can be understood as never picking a dispreferred option in *any* choice. We could then constrain reasonable choice by demanding that the choice function (pairwise) satisfies some particular relation defined on the acts. If $\mathcal{C}(A)$ is nonempty for all nonempty A[10] and satisfies $\succeq$ then it pairwise satisfies it, but the converse need not be true. This is proved in Appendix B (Theorem 1).

A relation can also determine a kind of choice function. The *maximal set* for a relation $\succeq$ is $\mathcal{M}_{\succeq}$:

$$\mathcal{M}_{\succeq}(A) = \{\varphi \in A : \neg\exists \psi \in A, \psi \succ \varphi\}$$

[9] Consider a case where $\varphi \in A$ but $\varphi \notin B$ and φ is better than all acts in B. Then it is unreasonable to have any elements of B (thus of $\mathcal{C}(B)$) in $\mathcal{C}(A)$.

[10] We will call this property Decisive later.

Interpreting the "$\succeq$" as a relation of preference, this $\mathcal{M}_{\succeq}$ is the set of acts that aren't strictly dispreferred to anything else in the set. Here are some facts about $\mathcal{M}_{\succeq}$.

(i) $\mathcal{M}_{\succeq}$ is a choice function

(ii) $\mathcal{M}_{\succeq}$ pairwise satisfies $\succeq$

(iii) If $\succeq$ is acyclic on A where A is finite then $\mathcal{M}_{\succeq}(A)$ is non-empty

(iv) If $\succeq$ is transitive, then $\mathcal{M}_{\succeq}$ satisfies $\succeq$.

These are proved in the appendix (Theorem 2).

Call a choice rule $\mathcal{C}$ *more discriminating* than $\mathcal{C}'$ when $\mathcal{C}(A) \subseteq \mathcal{C}'(A)$ for all A. $\mathcal{M}_{\succeq}$ is the least discriminating choice function that satisfies $\succeq$. That is, if $\mathcal{C}$ satisfies $\succeq$ then $\mathcal{C}(A) \subseteq \mathcal{M}_{\succeq}(A)$ for all A. This is also proved in the appendix (Theorem 3).

Sometimes we will talk about the relation generated by a function F, $\succeq_F$. We understand this to be the relation such that $\varphi \succeq_F \psi$ iff $F(\varphi) \geq F(\psi)$. For instance, $\varphi \succeq_{\mathrm{E}_{\mathbf{pr}}} \psi$ iff $\mathrm{E}_{\mathbf{pr}}(\varphi) \geq \mathrm{E}_{\mathbf{pr}}(\psi)$. We can think of relations as pairs of elements of the domain of the relation, so it makes sense to talk about the intersection and union of relations, and of one relation being a subset of another.[11] I will help myself to such ways of talking in what follows.

When your credences are precise, your choice rule is $\mathcal{M}_{\succeq_{\mathrm{E}_{\mathbf{pr}}}}$. That is, you choose among the things that do best by the criterion of expected value.

What if, instead of talking about maximality, we talked about *optimality*? The *optimal set* for a relation $\succeq$ is: $\mathrm{Opt}_{\succeq}(A) = \{\varphi \in A : \forall \psi \in A, \varphi \succeq \psi\}$. What we will find is that optimality – which is stronger than maximality – is too strong a property. That is, $\mathrm{Opt}_{\succeq}$ is often empty. Consider the set $\{\varphi, \psi\}$ where no relation holds between the two options. For this set, there are no optimal acts – although both acts are maximal in the sense of $\mathcal{M}_{\succeq}$. If the relation is complete, reflexive and acyclic then $\mathrm{Opt}_{\succeq}$ is nonempty (Sen 1977, p. 55). When $\mathrm{Opt}_{\succeq}(A) \neq \emptyset$, and $\succeq$ is transitive then $\mathrm{Opt}_{\succeq}(A) = \mathcal{M}_{\succeq}(A)$ (Theorem 4). This means that talking about optimality is superfluous. Maximality is the more interesting concept in general. The two happen to coincide for complete, transitive relations but when we have incomplete relations, optimality can be empty while maximality won't be.

There is a third way to judge choice rules: we look at how they fare in certain decision problems. That is, we compare what the choice rule says about a decision problem to what our intuitions say about that problem. While it is often the case that our intuitions are not at all strong about what is the right choice in cases of imprecise decision, there are at least some clear cut cases of correct and incorrect imprecise choice: an imprecise decision rule should at least get these right.

In summary, we want to analyse what sort of choice rule makes sense for imprecise decision. We are going to proceed by imposing certain intuitive constraints on choice and showing that certain decision rules violate these principles. The principles will

[11] That is, define $R_{\succeq} \subseteq \mathcal{A} \times \mathcal{A}$ by: $(\varphi, \psi) \in R_{\succeq}$ iff $\varphi \succeq \psi$.

come in three flavours: restrictions on the functional form of $\mathcal{C}$, relations that $\mathcal{C}$ must satisfy and getting the right answers in inuitive problems.

One might think that given the material I'm taking inspiration from, I would be aiming at a representation theorem (Luce and Raiffa, Milnor) or an impossibility theorem (Arrow, Sen). I am doing neither. Both representation and impossibility theorems start by making some assumptions about what rational choice requires. I don't think the conditions I discuss below are enough to generate an impossibility, nor do I think they are sufficient for any interesting kind of representation.[12] My main focus is on what we can say about *rational constraints on choice*.

3. Non-negotiable properties of choice

Let's consider a first possible decision rule. What about $\mathcal{C}(A) = A$. This satisfies the definition of a choice function. Unfortunately, it is thoroughly unhelpful! This choice rule is maximally permissive, or minimally discriminating. In general, we would like our choice rules to be as discriminating as possible. This is something of a guiding principle for this paper: *ceteris paribus*, more discriminating choice rules are better.

Now let's look at some conditions on choice functions that I take to be absolute minimum requirements on rational choice.

3.1. Decisiveness

It seems obvious that if $\underline{\mathcal{E}}(\varphi) \geq \overline{\mathcal{E}}(\psi)$ for all ψ, then φ should be chosen. If φ *cannot do worse* than *the best any other act can do*, then φ is clearly the best act. Let's define *Interval Dominance* as the choice rule that captures this intuition. $\mathrm{ID}(A) = \{\varphi \in A : \forall \psi \in A \ \underline{\mathcal{E}}(\varphi) \geq \overline{\mathcal{E}}(\psi)\}$ This is the decision rule that Henry Kyburg suggests in Kyburg (1983). He calls it "Principle III".[13] Unfortunately, ID is often empty. That is, there is often no unambiguously best act in this sense: where some acts have overlapping expectations, neither interval dominates the other. So neither act is in the choice set.

This sort of failure to choose any choiceworthy acts seems a flaw in a choice rule. This suggests the first absolutely non-negotiable principle on choice functions.

Decisiveness: If $\mathcal{C}(A) = \emptyset$ then $A = \emptyset$.

We need our choice functions to actually help us make choices. If the choice function can just "give up" then it isn't really that useful. There may be cases where "giving up" is in fact the right thing to do; for example in cases where you have to choose among acts of the form "n days in heaven followed by eternity in hell" for all $n \in \mathbb{N}$. It seems no choice of n is permissible, since there will always be a much larger n.[14] Even if we want to grant this, the ID rule looks like it is definitely giving up too

[12] Although the extremely general theorems of Evren and Ok (2011) or Chu and Halpern (2004, 2008) might apply.

[13] In response to Teddy Seidenfeld's comments (pp. 259–61), Kyburg changes his mind (p. 271). We will discuss this in due course.

[14] Thanks to Jim Joyce for pointing out this example to me.

easily. So perhaps Decisiveness should only hold for some restricted class of decision problems: those that don't involve awkward acts like the "n days in heaven" example. For the remainder of this paper, I am talking only about that restricted class of acts. I think it contains all choice problems we ever actually face.

3.2. Interval dominance

Despite failing as a choice rule, we can use this ID idea to further restrict reasonable choice rules: when some act does interval dominate all others, then the dominating act should be in the choice set. Define the relation $\varphi \succ_{\text{ID}} \psi$ iff $\underline{\mathcal{E}}(\varphi) \geq \overline{\mathcal{E}}(\psi)$. Note this is defined directly as an irreflexive relation, since it doesn't lend itself to having a reflexive part. This gives us another non-negotiable condition.[15]

Interval Dominance: $\mathcal{C}$ satisfies $\succ_{\text{ID}}$

$\succeq_{\text{ID}}$ is transitive and thus acyclic, so $\mathcal{M}_{\succeq_{\text{ID}}}$ is decisive. Often $\succ_{\text{ID}}$ is empty, so this condition will put no restrictions on choice. However, when it is not empty, the restrictions it puts on choice are reasonable.

3.3. Non-domination

Consider the following example:

Example 1: There is a coin of unknown bias. You are offered the choice between these two bets:

- a: win £1 if the next toss lands heads
- b: win £1 if the next ten tosses all land heads

The first of these is clearly better. Act b has an obvious flaw: namely that the expected value for a is always at least as high. That is, all the credal committee members agree that a is at least as good as b, and most think that a is better. We'd like our choice rule to take into account this sort of unanimity of the credal committee. Call an act ψ (weakly) dominated by act φ if $\varphi \succeq_{\mathrm{E}_{\mathbf{pr}}} \psi$ for all $\mathbf{pr} \in \mathcal{P}$. Call an act strongly dominated if it is always strict. It seems reasonable that in this circumstance we should prefer φ to ψ and in fact that we can safely ignore ψ when making our final decision. Let's consider the relation of dominance, $\succeq_{\text{Dom}}$, as a relation that we want our choice rule to satisfy. Note that $\succeq_{\text{Dom}} = \bigcap_{\mathcal{P}} \succeq_{\mathrm{E}_{\mathbf{pr}}}$. That is, the relation of dominance is the intersection of all the relations of higher expectation. φ dominates ψ if and only if every relation of expectation (in your representor) ranks φ and least as high as ψ. This motivates another important desideratum for imprecise choice.

Non-domination: $\mathcal{C}$ satisfies $\succeq_{\text{Dom}}$

[15]Note that $\varphi \succ_{\text{ID}} \psi$ and $\psi \succ_{\text{ID}} \varphi$ implies φ and ψ have the same precise expectation. So the second condition of the definition of "satisfies" is still reasonable in this odd case.

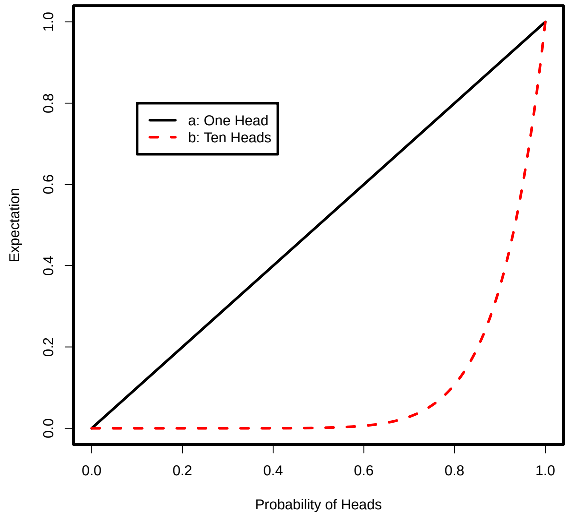

Figure 1: Graph of Example 1

Note that this is a stronger condition than Interval Dominance. That is, whenever φ interval dominates ψ, φ dominates ψ. Put another way, $\succeq_{\mathrm{Dom}} \supseteq \succeq_{ID}$. Another related condition is State-wise Dominance: $\varphi \succeq_{SWD} \psi$ when $\varphi(w) \geq \psi(w)$ for all $w \in \Omega$. Again, $\succeq_{\mathrm{Dom}} \supseteq \succeq_{SWD}$. No logical relationship holds between ID and SWD.

3.4. Precise limit

Another thing we want our decision rule to do is to reduce to the standard "maximise expectation" rule in the case of singleton representors.

Precise limit: When $\mathcal{P} = \{\mathbf{pr}\}$ then $\mathcal{C}(A) = \mathcal{M}_{\mathrm{E}_{\mathbf{pr}}}(A)$

This encodes the idea that we want imprecise choice to be an extension of, not a replacement for, standard decision theory.

4. Desirable properties of choice

No rule that violates the above properties should even be on the table as a possible choice rule. Now we turn to properties of choice rules that, while not absolute requirements, are still desirable properties of choice. These are all properties that restrict the functional form of the choice rule.

4.1. Contraction consistency

First, let's look at the intuition behind this property. Consider the following scenario. You go to a restaurant and see that the menu consists of Fish, Steak or Chicken. You

decide on Chicken. The waiter comes to take your order and tells you there is no more Fish. So you decide to have the Steak. This story seems a little odd. Why should the availability of an option you don't choose cause a switch in choice like the one exhibited in the move from Chicken to Steak? It seems like a reasonable choice rule should be somewhat consistent under various kinds of expansion or contraction of the option set.[16]

Contraction Consistency: $\mathcal{C}(A \cup B) \subseteq \mathcal{C}(A) \cup \mathcal{C}(B)$

This rule is more normally seen in one of these equivalent forms:

$$\text{If } \varphi \in \mathcal{C}(A), B \subseteq A, \varphi \in B \text{ then } \varphi \in \mathcal{C}(B) \tag{1}$$

$$\text{If } \varphi \notin \mathcal{C}(B), B \subseteq A, \varphi \in B \text{ then } \varphi \notin \mathcal{C}(A) \tag{2}$$

So in the preceeding story, $\mathcal{C}(S,C,F) = C$ but $\mathcal{C}(S,C) = S$. This violates the above property. This property is also known as Sen's alpha condition (Sen 1970, 1977). I am following Gaertner (2009) in calling it "contraction consistency", but it also somewhat restricts expansion of the option set. Luce and Raiffa have a version of (2) as their Axiom 7.

4.2. The Sure Thing Principle

The sure thing principle is often said to be violated by people's actual choices in games like those of the Ellsberg paradox (Camerer and Weber 1992; Ellsberg 1961). The property does, however, have some intuitive appeal, so it is worth discussing with a view to perhaps *limiting* if not eradicating violations of it. Luce and Raiffa's Axiom 8 corresponds to the "sure thing principle". As Luce and Raiffa put it:

> Consider a probability mixture of two decision problems with the same actions and states. If the payoffs in the second problem do not depend on the act chosen, then the optimal set in the mixed problem is the same as in the first problem. Luce and Raiffa (1989, p. 290)

Broome (1991) characterises STP in terms of what he calls *Separability*. The idea is that how good sub-parts of the decision problem are can be evaluated in isolation. It guarantees that the evaluation of the acts will, in a certain sense, be an aggregation of the evaluations of the sub-parts. In the standard case that aggregation is done by probability-weighted sums. We can cash out STP as:

Sure thing principle: $\mathcal{C}(pA + (1-p)\varphi) = p\mathcal{C}(A) + (1-p)\varphi$

[16] Imagine you are at a sushi restaurant. You prefer salmon over tuna, but find that salmon needs to be prepared by a good chef in order to be tasty. So you take the safe option and go for tuna. Now you notice that the restaurant offers fugu – pufferfish – which is poisonous unless prepared properly. You change your order to salmon. This seems like a failure of Contraction Consistency, but it isn't really. Learning that the restaurant offers fugu gives you evidence that they have a good chef. This changes *each* of the acts on offer, so the change isn't just a simple addition of a new act. Thus the axiom doesn't apply. Thanks to Nick Baigent for this example.

Perhaps the best way to understand STP is with an example.

> EXAMPLE 2: I am going to ask you to choose c or d. Then I'm going to roll a fair die and flip a coin of unknown bias. If the die lands even, you gain £6 if $\neg H$, nothing otherwise. If the die lands odd, c and d pay out as set out here:
>
> - c: Gain £10 if H, nothing otherwise
> - d: Gain £2 if H, £8 otherwise

The idea is that since what you choose – c or d – doesn't make a difference if the die lands even, then you should choose in order to get the better of the options when it matters (in the odd branch of the game).

Seidenfeld (1988) argues that violating STP leads to diachronic inconsistency. That is, if you use a choice rule that violates STP, you might end up making suboptimal sequences of decisions (see also Peterson 2009, pp. 176–9 and Steele 2010).

4.3. Union consistency

A third desirable property of choice is Union Consistency. Recall that Contraction Consistency put a sort of "upper bound" on $\mathcal{C}(A \cup B)$ by requiring that it be a subset of $\mathcal{C}(A) \cup \mathcal{C}(B)$. Union Consistency puts a *lower* bound on $\mathcal{C}(A \cup B)$.

> Union consistency: $\mathcal{C}(A) \cap \mathcal{C}(B) \subseteq \mathcal{C}(A \cup B)$

This is Sen's gamma condition. It is sometimes seen in this equivalent form:

$$\text{If } \varphi \in \mathcal{C}(A), \varphi \in \mathcal{C}(B) \text{ then } \varphi \in \mathcal{C}(A \cup B) \tag{3}$$

The motivation here is that if you would choose Steak out of Steak or Fish, and you'd choose Steak out of Steak or Chicken, then you should choose Steak when all three options are on the menu.

5. Optional properties of choice

We have now seen some non-negotiable properties of choice and some desirable properties of choice. For completeness, we shall also consider some properties of choice that are optional.

5.1. Continuity

Consider the following example.

> EXAMPLE 3: You are betting on a coin of unknown bias. It will be tossed twice. The following bets are available to you:

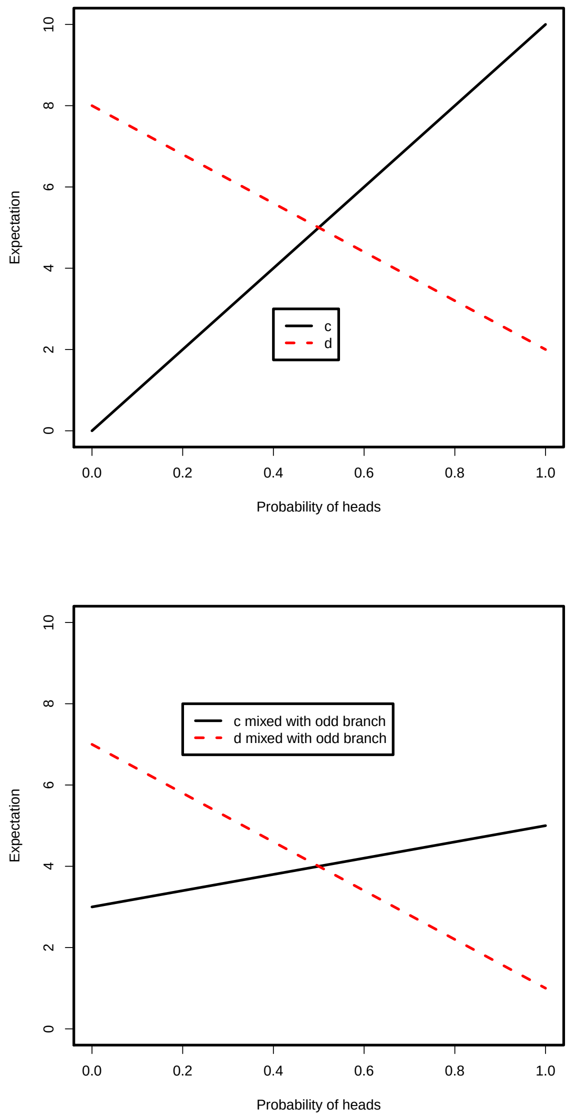

Figure 2: Graph of Example 2

- e: Win £0.50
- f: Win £$1+\varepsilon$ if the coin lands on different sides on the two tosses and win £ε otherwise

For very small values of ε, it seems like the sure £0.50 is the safer bet. The idea is that act f is "only ε away from being dominated" and so is not a good option for small enough ε. It seems like as ε gets smaller, at some point f should stop being choice-

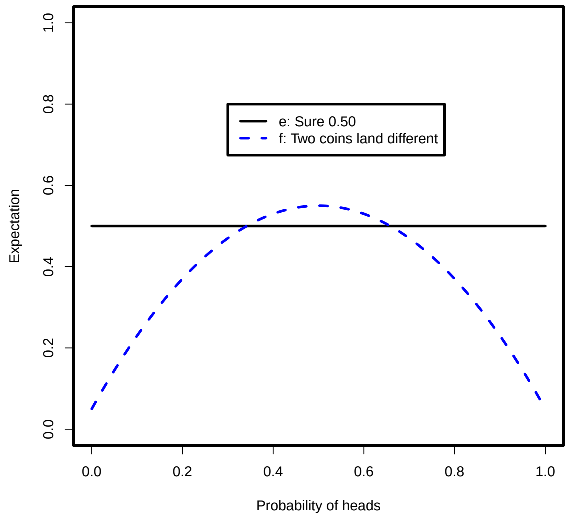

Figure 3: Graph of Example 3

worthy. This motivates a property termed *continuity*. Continuity is about sequences of decision problems. A sequence of decision problems is defined by a collection of sequences of outcomes. I borrow Ken Binmore's gloss on Milnor's "continuity" axiom here:

> Consider a sequence of decision problems with the same acts and states, in all of which $\varphi_j \succ \varphi_i$. If the sequence of matrices of outcomes converges, then its limiting value defines a new decision problem in which $\varphi_j \succeq \varphi_i$.
>
> Binmore (2008, p. 137, minor notational changes)

Continuity says that if we keep changing the payoffs slightly and the preference remains stable, then in the limit of our fiddling, the preference will still be stable. The change from "$\succ$" to "$\succeq$" is to allow this to be consistent with a case where act φ_j tends to φ_i. Or rather, each consequence of φ_j tends to the corresponding consequence of φ_i. In the limit, they are equal, so there cannot be strict preference between them.

In terms of optimal acts, if we keep changing the payoffs slightly and the set of optimal acts remains stable, then in the limit, the set of optimal acts is stable. The idea is that if you have some sequence of small bonuses ε_i that are added to some φ,

this determines a sequence of act sets A_i with $\varphi + \varepsilon_i \in A_i$ for each i. As the ε_i tend to the limit, ε, A_i tend to A. The stability of the choice function now becomes:

> Continuity: If $\varphi + \varepsilon_i \in \mathcal{C}(A_i)$ for all i, then $\varphi + \varepsilon \in \mathcal{C}(A)$

In the above example we appealed to the contrapositive of this property: if the "limit act" as ε tends to zero is not choiceworthy, then there should be some ε such that f is not choiceworthy.

5.2. All-or-nothing expansion consistency

Next let's consider a property whose violation I don't consider a flaw at all. Understanding why I think imprecise choice rules should be allowed to violate this property will point to an important difference between imprecise choice and precise choice. I shall call this property "all-or-nothing expansion consistency". It is called 7″ by Luce and Raiffa and "beta" by Sen. This says that if an old choiceworthy act is made non-choiceworthy by the addition of new acts, then *all* old choiceworthy acts are made non-choiceworthy.

> All-or-nothing: If $\varphi \in \mathcal{C}(A)$ but $\varphi \notin \mathcal{C}(A \cup B)$ then, for all $\psi \in \mathcal{C}(A)$, we have $\psi \notin \mathcal{C}(A \cup B)$

As Luce and Raiffa show, this "all-or-nothing" condition makes sense only when you are evaluating the acts on a single scale. Imagine you are picking your team for the school gymnastic competition. You need the best long jumper and the best high jumper in class A. So you have a choice function that picks out Laura as the best long jumper and Horatio as the best high jumper. That is, the choice function "best athletes" chooses these two students. Now imagine that we expand the set of available students to be the whole year group not just the class. As it happens, Herbert, in class B is a better high jumper than Horatio. But Laura is the best long jumper in the year. So the choice rule now picks out Laura and Herbert for the team. This rule violates the above all-or-nothing rule, because different members of the choice set are included based on different kinds of evaluations. The choice function for each sport will satisfy this condition, but the overall "best athletes" function does not. Such a choice function can't be given a strong interpretation. That is, each member of the choice set is better than all acts outside the choice set in some sense; but it is not the case that all members of the choice set are equally good. They are merely good in different ways. I claim that imprecise decision can be a little like this, and thus that all-or-nothing should not be required. It is a property that makes sense only for *strong* choice functions.

Sugden (1985) discusses a similar example where one race car is faster and another is more manoeuvrable: the first will win in a head to head race, but the second will win if there are other cars on the track. Thus the "race winning function", if you like, does not satisfy All-or-nothing. Single criterion choice (as characterised by All-or-nothing) and the strong interpretation of the choice set go hand in hand.

5.3. Convexity

I don't really have an intuition for the following property, but it appears in a number of places, so it is worth discussing briefly. This is the property of Convexity which relates how choiceworthiness interacts with mixed acts.

> Convexity: if $\varphi, \psi \in \mathcal{C}(A)$ then $p\varphi + (1-p)\psi \in \mathcal{C}(A)$ for all $p \in [0,1]$

This is Luce and Raiffa's Axiom 9. Milnor has a version of this that says that mixtures of acts are weakly preferred. That is, it says "if $a_1 \sim a_2$ then $pa_1 + (1-p)a_2 \succeq a_1, a_2$ for $p \in [0,1]$." Of course, this only makes sense if these "mixtures" are in A.

I find myself untroubled by violations of convexity. So why might we consider it a condition of rational decision? There is an idea that a mixture is at least as good as the worst part of the mixture, and at most as good as the best part. This is related to what R. Bradley (2007) calls the "Averaging Slogan".

> No prospect is better (or worse) than its best (worse) realisation in a set of mutually exclusive and exhaustive prospects. (p. 241)

In particular, if both parts of the mixture are good enough to make it into $\mathcal{C}(A)$, then so should the mixture be. But this seems to be trading on the same "single-criterion" intuition as the above All-or-nothing property. Say Horatio is a terrible long-jumper and Laura a terrible high-jumper. Someone who was a "mixture" between these people – call him/her Lauratio – would not be as good as either of them. If Lauratio is an average high-jumper and an average long-jumper, then she/he doesn't deserve a place on the athletics team. So All-or-nothing and Convexity both seem reasonable only for strong choice sets, or "single-criterion" choice.

6. Weak choice rules

Let's turn now to considering some specific choice functions. The aim is to assess these functions based on the properties we have outlined in the last three sections. We first consider some choice rules that are fairly "weak" in the sense that they are not very discriminating.

6.1. Non-domination

Let's look at restricting the space of reasonable acts. Instead of considering which acts are the "best" in some sense, I want to focus on which acts you can justifiably rule out. I want to look at the necessary conditions on an act's being good. This gives us a choice set that only admits of a weak interpretation. That is, all we can say of the acts in the choice set is that they are all better than those that have been excluded. But we cannot say that the acts remaining in the choice set are necessarily all good options (or comparable).

What about just taking $\mathcal{M}_{\geq_{\text{Dom}}}$ as our choice rule? That is, any acts that are not dominated are in the choice set. It is, perhaps, too permissive a rule. For example, it doesn't rule out act f in Example 3. And indeed, it doesn't satisfy the Continuity property. Also, this rule does not satisfy the All-or-nothing property. Here is an example of how $\mathcal{M}_{\geq_{\text{Dom}}}$ fails all-or-nothing expansion consistency.

Example 4: Consider the choice between g and h, and the choice between g, h and k.

- g: Gain £10 if H, nothing otherwise
- h: Gain nothing if H, £10 otherwise
- k: Gain £11 if H, £1 otherwise

k dominates g, so in the expanded decision problem, g is not choiceworthy. However, h is still undominated, so this violates all-or-nothing.

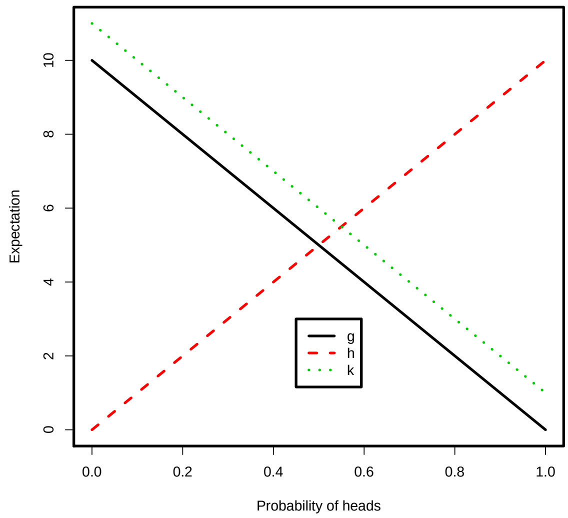

Figure 4: Graph of Example 4

A mixture of undominated acts can be dominated (see Table 1). Each of a_1 and a_2 are undominated, but the mixture is dominated by a_3. So convexity is not true for $\mathcal{M}_{\geq_{\text{Dom}}}$.

Would it be possible to construct a rule that never picks an act whose mixtures with other acts can be dominated? Would it be desirable? Forcing a choice rule to choose among $\mathcal{M}_{\geq_{\text{Dom}}}$ but also satisfy Convexity would make the choice rule violate Union Consistency: consider $\{a_1, a_2\}$ and $\{a_1, a_3\}$ (in each case we assume now that the act space is closed under mixtures). In each case, a_1 is undominated and mixtures of a_1 with other acts aren't dominated either. However, in the larger act space where a_1, a_2 and a_3 (plus their mixtures) are available, Convexity forces a_1 and a_2 to be

	s_1	s_2
a_1	2	−2
a_2	−2	2
a_3	1	1
$0.5a_1 + 0.5a_2 = a_4$	0	0

Table 1: A mixture of undominated acts can be dominated

unchoiceworthy since Table 1 shows that mixtures of them are dominated. So only a_3 is choiceworthy. This violates Union Consistency. So forcing a rule like $\mathcal{M}_{\succeq_{\text{Dom}}}$ to obey Convexity comes at a high price.

None of the above properties that $\mathcal{M}_{\succeq_{\text{Dom}}}$ violates are non-negotiable constraints on rational choice, so $\mathcal{M}_{\succeq_{\text{Dom}}}$ seems a good choice rule. It is perhaps too permissive, but it does at least serve as a constraint on rational imprecise choice. It satisfies all the desirable properties.

6.2. E-admissibility

The main problem with $\mathcal{M}_{\succeq_{\text{Dom}}}$ is that it isn't really discriminating enough. That is, the choice sets that that rule generates will often contain many acts. We would really like choice to be more constrained. Let's consider a more discriminating choice rule. Another restriction of the act set – "E-admissibility" – is due to Isaac Levi (Levi 1974, 1986). An act is E-admissible if there is some probability in your representor such that that act maximises expectation with respect to that probability function. In picture terms an act is E-admissible if it is "on top" for some probability value in the representor. E-admissible acts are the ones that some credal committee member thinks are best (by that member's standard of $\mathrm{E}_{\mathbf{pr}}$). Levi argues that you should only choose among E-admissible acts. A first attempt at cashing out this choice rule is:

$$L(A) = \bigcup_{\mathbf{pr}\in\mathcal{P}} \mathcal{M}_{\succeq_{\mathrm{E}_{\mathbf{pr}}}}(A) \tag{4}$$

This might be more perspicuously rephrased as:

$$L(A) = \{\varphi \in A : \exists\, \mathbf{pr} \in \mathcal{P}, \forall \psi \in A, \mathrm{E}_{\mathbf{pr}}(\varphi) \geq \mathrm{E}_{\mathbf{pr}}(\psi)\} \tag{5}$$

The intuition is that we ask each credal committee member to pick their favourite act(s): we then take the collection of each of these favourites.

As it stands, the definition of E-admissible isn't quite good enough. Recall Example 1 where we had the choice between a bet on heads and a bet on ten heads in a row. The latter maximises expectation for $\mathbf{pr}(H) = 0$ and $\mathbf{pr}(H) = 1$ so it is E-admissible. This act is, however, weakly dominated.[17] I expect that Levi would want to rule

[17] L never contains *strongly* dominated acts.

out choosing weakly dominated acts, even if they maximise for some **pr**. We could give Levi the benefit of the doubt and say that what he was really interested in was $\mathcal{M}_{\geq_{\mathrm{Dom}}} \circ L(A)$ where "$f \circ g$" is composition of functions. We shall call this $\mathcal{L}(A)$.

We know that $\mathcal{L}(A) \subseteq \mathcal{M}_{\geq_{\mathrm{Dom}}}(A)$. There are undominated acts that are not E-admissible. So we in fact know that $\mathcal{L}(A) \subsetneq \mathcal{M}_{\geq_{\mathrm{Dom}}}(A)$ for some A. So $\mathcal{L}$ is more discriminating that $\mathcal{M}_{\geq_{\mathrm{Dom}}}$.

One further motivation that might be offered for preferring $\mathcal{L}$ to $\mathcal{M}_{\geq_{\mathrm{Dom}}}$ is that undominated acts can be dominated by mixtures of other acts, but E-admissible acts can't be dominated by mixed acts. This seems like a strange argument. If the mixed acts are available acts, then the act dominated by mixtures won't be undominated. If the mixed acts aren't in the set of available acts, then it's irrelevant that there is something that dominates it.[18] That is, some "undominated" acts, φ can be dominated by mixtures $p\psi + (1-p)\rho$ where $\psi, \rho \in A$. If the mixture is in A, then φ isn't undominated and thus not choiceworthy. If the mixture is not in A (and thus not a feasible choice) then domination by the mixture is irrelevant. A second point against this is that not all undominated acts can be dominated in this way. That is, some undominated, but E-*in*admissible φ are such that no mixture of acts in A dominates φ.[19] In a moment, we will see such an undominated act that can't be dominated by mixtures: the act o in Example 6 is one such act.

Given that E-admissibility is more discriminating and given that non-domination is arguably too permissive (not discriminating enough), one might think that E-admissibility is obviously the better rule. But $\mathcal{L}$ seems to "get things wrong" in certain cases, where $\mathcal{M}_{\geq_{\mathrm{Dom}}}$ gets it right. That is, $\mathcal{L}$ rules out some intuitively reasonable acts as we shall see shortly. $\mathcal{L}$ doesn't help solve any of the problems with $\mathcal{M}_{\geq_{\mathrm{Dom}}}$.

$\mathcal{L}$ violates union consistency, as can be seen from considering Example 5.

> Example 5: You are betting on a coin of unknown bias. You can choose among these bets:
>
> - l: Gain £10 if H, lose 5 otherwise
> - m: Lose £5 if H, gain 10 otherwise
> - n: Gain 0 whatever happens

$\mathcal{L}(\{l, n\}) = \{l, n\}$ and $\mathcal{L}(\{m, n\}) = \{m, n\}$, but $\mathcal{L}(\{l, m, n\}) = \{l, m\}$. That is, n is choiceworthy in both pairwise choices, but if all three options are offered together, then n is ruled out.

As I mentioned earlier, Kyburg (1983) suggests his Principle III, which is essentially what we called Interval Dominance. In the commentary on Kyburg's paper, Teddy Seidenfeld suggested two improvements on Principle III which are effectively

[18] Every act φ can be dominated by another act φ' that gets you what φ got, plus £10. These acts aren't typically available to you, so are irrelevant to choice.

[19] More carefully, fix a set of acts A. Only some undominated acts in A are dominated by mixtures of acts in A. We could still have that the act is undominated in some larger set of acts A' such that it is dominated by mixtures of acts in A'.

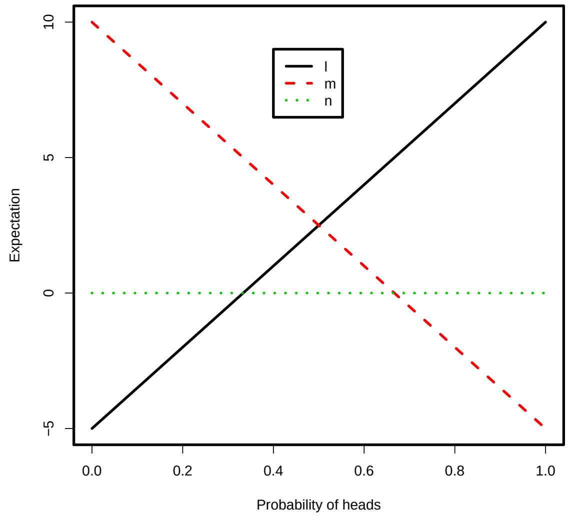

Figure 5: Graph of Example 5

$\mathcal{M}_{\geq_{\text{Dom}}}$ and $\mathcal{L}$. Seidenfeld notes that $\mathcal{L}$ violates Union consistency.[20] Ultimately Seidenfeld suggests that despite this failure, $\mathcal{L}$ is the better improvement on Principle III (Kyburg 1983, pp. 259–61). In Kyburg's response (p. 271) he concedes that Interval Dominance should be improved on, but finds the violation of Union Consistency to be undesirable, and thus opts for $\mathcal{M}_{\geq_{\text{Dom}}}$. Seidenfeld (2004) points out that $\mathcal{L}$ violates Union consistency and also All-or-nothing.

E-admissibility also just seems to get it wrong in certain intuitive cases.

Example 6: You are betting on a coin of unknown bias. It will be tossed twice. The following bets are available to you.

- o: Win £1 $-\varepsilon$ if the coin lands the same way twice, lose ε otherwise
- q: Win £1 $+\varepsilon$ if the coin lands on different sides each time, win ε otherwise
- r: Win £1 if the first toss lands heads, 0 otherwise
- s: Win £1 if the first toss lands tails, 0 otherwise

The intuition here is that whatever value **pr** takes, o does almost as well as the better of r, s, and normally better than q as well. However, o is not E-admissible. It's always *close* to whatever act maximises – something not true of any of the other acts – but it never actually maximises for any **pr**. No credal committee member thinks o is the best, but they all think it's pretty good. It seems reasonable that a choice rule should not rule out o, and arguably *should* rule out q. $\mathcal{L}$ on the contrary rules out o and does not rule out q.

[20] Seidenfeld mistakenly refers to this as Sen's beta property – All-or-nothing – which $\mathcal{L}$ also violates.

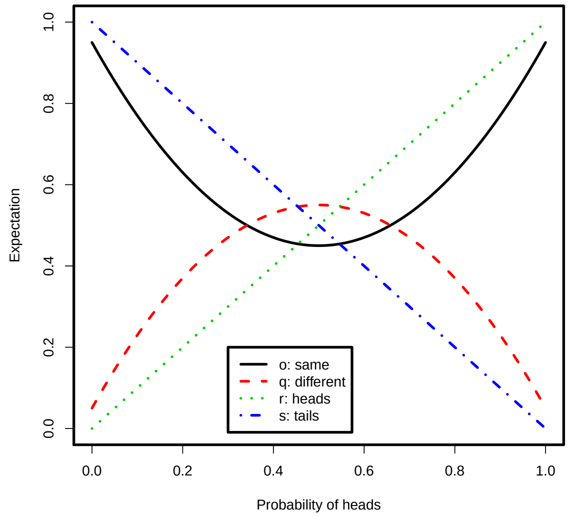

Figure 6: Graph of Example 6

Despite being more discriminating, E-admissibility does not seem like an improvement on non-domination. It seems to get things wrong in some intuitive cases, and it violates additional plausible properties for choice rules.

7. More discriminating rules

Both of the above choice rules are fairly undiscriminating. Let's try to find some choice rules that give us smaller choice sets. Why not also try to find choice rules that admit of the stronger interpretation? In the precise case, we can evaluate acts in terms of their expectation. This gives us a way to linearly order the acts. The following rules will all involve a component that basically evaluates acts on a linear scale in the same way that $\mathrm{E}_{\mathbf{pr}}$ works for precise choice. Recall that in the precise context, there is a way of assigning a value to each act (its expectation) and then choice is simply a matter of maximising that quantity. One method of imprecise choice would be to find some analogous quantity that works in the imprecise case. That is, we could look for something that can straightforwardly play the role that $\mathrm{E}_{\mathbf{pr}}$ plays in the precise context.

7.1. Valuing acts

A first attempt at valuing acts in the imprecise case would be to look at $\underline{\mathcal{E}}$. Is maximising $\underline{\mathcal{E}}(\varphi)$ a good decision rule? We can describe this rule as $\mathcal{M}_{\underline{\mathcal{E}}}$.[21] This rule is

[21] This is a little bit of an abuse of notation. The function $\underline{\mathcal{E}}$ determines a relation $\succeq_{\underline{\mathcal{E}}}$ as described earlier. It is this relation that this choice rule is maximal for.

sometimes described as "gamma-maximin" (Seidenfeld 2004). It is also the rule that Gärdenfors and Sahlin (1982) advocate.

$\mathcal{M}_{\underline{\mathcal{E}}}$ does not satisfy non-domination. That is, $\mathcal{M}_{\underline{\mathcal{E}}}$ sometimes contains acts that are dominated, as Example 1 shows. The above problem isn't just a problem for $\mathcal{M}_{\underline{\mathcal{E}}}$, but for any rules that focus only on the set of expectations. For example, instead of maximising $\underline{\mathcal{E}}$, consider maximising $\mathcal{H}_p(\varphi) = p\,\underline{\mathcal{E}}(\varphi) + (1-p)\,\overline{\mathcal{E}}(\varphi)$ for some real number p between 0 and 1. This is an "imprecise analogue" of the *Hurwicz criterion* for choice under complete ignorance (Hurwicz 1951; Milnor 1951). This is actually a whole class of different decision rules depending on choice of p. If $p = 1$ then we recover maximise minimum expectation ($\mathcal{M}_{\underline{\mathcal{E}}}$). If a precise p value seems arbitrary, perhaps consider looking for acts that do well for many different values of p. Bandyopadhyay (1994) suggests a rule that, effectively, amounts to preferring φ to ψ just in case φ is better according to all values of p. Sadly, none of these rules can avoid making (weakly) dominated acts permissible: none of these rules get things right in Example 1. That is, since $\underline{\mathcal{E}}(a) = \underline{\mathcal{E}}(b)$ and $\overline{\mathcal{E}}(a) = \overline{\mathcal{E}}(b)$, any rule that values acts as some function of these values must treat the two bets the same. What we want is a choice function that rules out b.

As well as violating the rationally compelling Non-domination principle, the $\mathcal{M}_{\underline{\mathcal{E}}}$ rule also violates STP. Consider Example 2: $\mathcal{M}_{\underline{\mathcal{E}}}$ chooses d over c in the odd branch. But when you mix with the even branch, c ends up looking better. That is, the payouts of c and d for the "mixed" decision problem are "5 if H, 3 otherwise" and "1 if H, 4 otherwise" respectively.

7.2. A composite rule

Since the problem with $\mathcal{M}_{\underline{\mathcal{E}}}$ (and similar rules) is that it allows weakly dominated acts to be choiceworthy, why not just compose it with $\mathcal{M}_{\succeq_{\mathrm{Dom}}}$ to make a better rule? Consider $\mathcal{M}_{\underline{\mathcal{E}}} \circ \mathcal{M}_{\succeq_{\mathrm{Dom}}}$: this is the rule that maximises minimum expectation among the acts that are undominated. This rule obivously satisfies Non-domination. It still fails STP, however.

It has a further problem. Consider a variant of Example 1 (the one heads versus ten heads game) where we replace act b with act b' which is like b but with a small bonus ε in all states. This act is undominated but still arguably a bad idea. Our composite rule not only fails to rule out b': it actually makes b' the *only* permissible choice in this game. This seems wrong.

7.3. Equivocal expectation

Let's consider another kind of valuation of the acts. Let's take our representor, $\mathcal{P}$, and find some privileged probability function within this set and maximise expectation with respect to that. This is essentially the method that Smets and Kennes (1994) suggest. It is also, in effect, what Objective Bayesians like Williamson (2010) endorse. Smets and Kennes call the privileged function the *pignistic probability*, for Williamson the privileged function is the maximally equivocal probability function. Call $\Downarrow\mathcal{P}$ the

maximally equivocal probability.[22] In what follows I have less to say about this sort of approach. It seems there are epistemological problems with deciding how to "average" your probability judgements, but once that is achieved, the decision theory is wholly orthodox. This decision rule, which we can term $\mathcal{M}_{\mathrm{E}_{\Downarrow\mathcal{P}}}$ satisfies pretty much all the conditions we have discussed (except Continuity). However, it gets things wrong in Example 6. Not only does it fail to rule in betting on the coin's landing the same on the two tosses (o), it actually makes betting on the coin's landing differently on two tosses (q) the *only* permissible option. Note that almost every credal committee member thinks o is better than q, but $\mathcal{M}_{\mathrm{E}_{\Downarrow\mathcal{P}}}$ – the tyranny of the equivocal – fails to take this fact into account. If we want to "average" the committee members' opinions in some way, it seems we should not be averaging their evaluations of the events. This gets us the wrong answer. What we want to do is average their evaluations of the *acts*. What this would actually involve, I am not sure.

7.4. Levi's rule

Isaac Levi champions using choosing among the E-admissible acts, and to maximise minimum expectation among them. That is, he champions $\mathcal{M}_{\underline{\mathcal{E}}} \circ \mathcal{L}(A)$. He argues that only those acts that have some chance of being the best are worthy of consideration. He argues for the "maximise minimum expectation" part by analogy to what happens in precise cases of equal expectation. He considers a bet like the following: I offer you a bet where you win £1 if the fair die lands even, but you lose £1 if it lands odd. Should you accept this bet, or refuse it? Both acts (accept, refuse) have the same expectation – £0 – so how do you choose between them? Levi suggests that in this situation you should maximise minimum gain. He says that the reason to refuse the bet is:

> not that refusal is better in the sense that it has higher expected utility than accepting the gamble. The options come out equal on this kind of appraisal. Refusing the gamble is "better" however, with respect to the security against loss it furnishes. Levi (1974, p. 411)

He suggests the same reasoning works in the imprecise case. We should use "security" as a tiebreaker, and thus that among the E-admissible acts we should choose with $\mathcal{M}_{\underline{\mathcal{E}}}$.

We have seen that both $\mathcal{M}_{\underline{\mathcal{E}}}$ and $\mathcal{L}$ have problems as decision rules. Combining them in the way Levi suggests leads to further problems. Levi's rule violates Contraction Consistency, as Seidenfeld (2004) points out.

> Example 7: Consider the choice between t, u and the choice between t, u, v.
>
> - t: £10 if H, nothing otherwise
> - u: £3 if H, £3 otherwise

[22] The method of choosing a privileged function in Smets and Kennes (1994) is different, but for the current case of complete ignorance of a coin's bias, it coincides with $\Downarrow\mathcal{P}$.

- v: £–1 if H, £8 otherwise

In a choice between t and u, it is u that does best by $\mathcal{M}_{\underline{\mathcal{E}}}$. However, adding v means that u is no longer E-admissible and of t and v, t does better. This is on top of Levi's

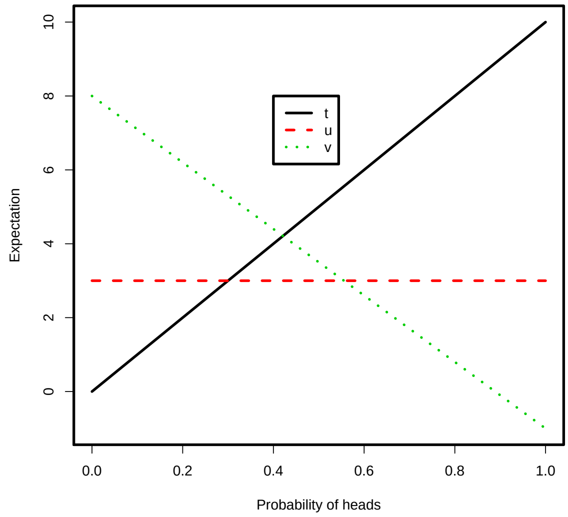

Figure 7: Graph of Example 7

rule's violation of All-or-nothing and Union Consistency.

Recall Example 6: where a coin will be tossed twice. Levi's rule gets things wrong here. Not only does E-admissibility rule out what seems like the best act (o), but then maximising minimum expectation picks out q, which seems like the *worst* act of the four. This gives a further argument against E-admissibility: if you are looking to have a permissive choice rule and use some sort of "tie-break" like $\mathcal{M}_{\underline{\mathcal{E}}}$, then using $\mathcal{L}$ gets you into more trouble than $\mathcal{M}_{\geq_{\mathrm{Dom}}}$ does.

In short, Levi's two-part choice rule composes two choice functions that have problems to generate a rule that has even more problems than the two parts have individually.

8. Regret based rules

In Example 6, the act o seems to do pretty well. No member of the representor takes it to be the best act, but it is always *close* to being top. Can we find a method of valuing acts that takes into account this intuition? We certainly can! Consider:

$$\mathcal{R}(\varphi) = -\max_{\mathbf{pr}\in\mathcal{P}} \left\{ \max_{\psi\in A} \left\{ \mathrm{E}_{\mathbf{pr}}(\psi) \right\} - \mathrm{E}_{\mathbf{pr}}(\varphi) \right\} \tag{6}$$

What this says it that we evaluate an act in terms of its "maximum regret": φ's regret is how far below the best act φ is for a particular $\mathbf{pr} \in \mathcal{P}$. So we take the maximum

regret over the probabilities in the representor. That is, for each probability in your representor, subtract the expectation of act φ from the "top act" for that probability. The maximum of those values over all probabilities in your representor is your maximum regret for act φ. You should act to minimise that. This is a quantity we want to *minimise*. This is why $\mathcal{R}$ is defined with a minus sign in front: this makes it line up with our general strategy of trying to find values we want to *maximise*. So let's consider the choice rule given by $\mathcal{M}_{\mathcal{R}}$. This is an "imprecise analogue" of Savage's *Minimax regret* rule for decision under complete ignorance (Luce and Raiffa 1989, pp. 280–2).

Unfortunately, this rule violates STP.[23] It also violates Contraction consistency. Given that the assessment of each act depends on maxima over the set of acts, it shouldn't be surprising that adding or removing acts can affect how the acts are evaluated.

Example 8: Consider the following acts, and consider choosing when z is an option, and when it isn't.

- x: Gain 10 if H, nothing otherwise
- y: Gain 8 if H, 1 otherwise
- z: Gain 4 if H, 3 otherwise

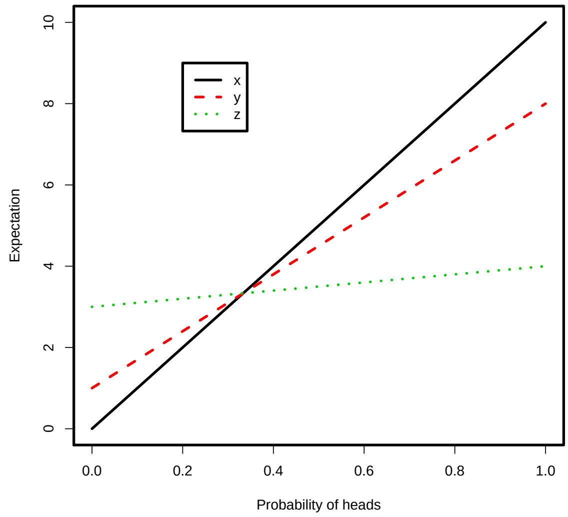

Figure 8: Graph of Example 8

Now, in a choice between x and y, x wins, since it is at most 1 from the best act (when $\mathbf{pr}(H) = 0$) while y drops to 2 less than the top act (when $\mathbf{pr}(H) = 1$). All

[23] For example, in Example 2, both acts have the same regret in the even branch, but c has lower regret in the mixed game. So a minor modification of this game would demonstrate a proper violation.

this changes when z becomes an option. Now, when $\mathbf{pr}(H) = 0$, x is 3 from the top act while y's maximum distance from the top stays at 2. So the addition of z has switched the evaluation of the two acts round! In short, $\mathcal{M}_{\mathcal{R}}(\{x, y, z\}) = \{y\}$ but $\mathcal{M}_{\mathcal{R}}(\{x, y\}) = \{x\}$. This violates Contraction Consistency. Example 5 shows that $\mathcal{M}_{\mathcal{R}}$ violates Union Consistency as well. This is so since $\mathcal{M}_{\mathcal{R}}(\{l, n\}) = \{l\}$ and $\mathcal{M}_{\mathcal{R}}(\{m, n\}) = \{m\}$, but in a choice between all three, only $\{n\}$ is choiceworthy.[24]

Note that the $\mathcal{M}_{\mathcal{R}}$ rule never picks a dominated act.[25] It does, however, violate a lot of plausible conditions on reasonable choice. But there is *something* going right in this rule. It seems to get the right answers, even in many of our problematic examples. For example, in the same/different problem (Example 6) it chooses the intuitively right act. It also chooses a in the unmodified as well as the modified version of Example 1. It also satisfies continuity. So it seems to have some advantages over $\mathcal{M}_{\succeq_{\mathrm{Dom}}}$. And it seems like it is a bad thing to choose an act if it can only do slightly better than some other, and could be a lot worse. That is, the motivation behind $\mathcal{M}_{\mathcal{R}}$ seems to point in the right direction.

Let's try to find a restriction on choice that is in the spirit of $\mathcal{M}_{\mathcal{R}}$, but that doesn't have all its bad consequences. Consider the following relation:

$$\varphi \succeq_{\mathrm{PWR}} \psi \text{ iff } \max_{\mathbf{pr}}\{\mathrm{E}_{\mathbf{pr}}(\varphi) - \mathrm{E}_{\mathbf{pr}}(\psi)\} \geq \max_{\mathbf{pr}}\{\mathrm{E}_{\mathbf{pr}}(\psi) - \mathrm{E}_{\mathbf{pr}}(\varphi)\}$$

This is basically the relation that holds between φ and ψ when φ's regret is less than ψ's when the two are compared pairwise. Because the comparison is done just over pairs of acts, it won't violate the various kinds of contraction consistency properties in the same way that $\mathcal{M}_{\mathcal{R}}$ does. We will see later that it does still violate these properties.

Can we demand that $\mathcal{C}$ satisfy this relation? Unfortunately not, since the relation isn't in general acyclic. For example consider acts that satisfy the properties in Table 2 This means that $\succeq_{\mathrm{PWR}}$ would impose inconsistent demands on $\mathcal{C}$. In short, $\mathcal{M}_{\succeq_{\mathrm{PWR}}}$ can

	a_1	a_2	a_3
$\mathrm{E}_{\mathbf{pr}_1}$	10	1	8
$\mathrm{E}_{\mathbf{pr}_2}$	1	9	1
$\mathrm{E}_{\mathbf{pr}_3}$	1	2	4

Table 2: Cyclic Pairwise Regret

be empty.

Despite its failings, there is a feeling that this rule is getting something right. Satisfying this relation in Example 6 – betting on whether two tosses of a coin will be the same or different – gets you the right choice of act (o). In example Example 8 (where $\mathcal{M}_{\mathcal{R}}$ failed), pairwise regret points to a consistent choice: whether or not z is included,

[24] Note that this is a very different profile of advice from that given by $\mathcal{L}$, which had n as choiceworthy in both smaller sets, but not in the choice between all three.

[25] The dominating act always has lower regret, for every $\mathbf{pr}$, so the dominated act is never maximal with respect to $\mathcal{R}$.

x is the most preferred act. Given that $\mathcal{M}_{\succeq_{\text{PWR}}}$ can be empty, it's not immediately clear how to build a decision rule out of this relation. Perhaps we should demand only that when the maximal set is non-empty, that set should be the choice set. In the same way that we want $\mathcal{C}(A) = ID(A)$ whenever $ID(A) \neq \emptyset$, we want $\mathcal{C}(A) = \mathcal{M}_{\succeq_{\text{PWR}}}(A)$ whenever this is nonempty.

Consider the following choice rule:

$$\mathcal{S}(A) = \begin{cases} \mathcal{M}_{\succeq_{\text{PWR}}}(A) \text{ if } \mathcal{M}_{\succeq_{\text{PWR}}}(A) \neq \emptyset \\ \bigcup_{\varphi \in A}\{\mathcal{S}(A \setminus \{\varphi\})\} \text{ otherwise} \end{cases}$$

The maximal set is never empty for two element sets of acts since you can't have cycles on pairs. This guarantees that this recursive definition never yields an empty set. What it effectively says is "if there is no maximal element, try all the ways of removing one act and look for maximal elements there." Since it is cycles that cause there to be no maximal elements, this procedure will look at all the ways you could remove an element from the cycle, and thus induce there to be a maximal element. Every element of a cycle will end up in the choice set. This is shown more carefully in the appendix (Theorem 5). This rule is also a subset of $\mathcal{M}_{\succeq_{\text{Dom}}}$ (Theorem 6), this means that it satisfies Non-domination. It also satisfies Union Consistency but not Convexity.

Unfortunately, this rule violates Contraction Consistency. Consider a cycle $a_1 \succeq_{\text{PWR}} a_2 \succeq_{\text{PWR}} a_3 \succeq_{\text{PWR}} a_1$. Now consider $\mathcal{S}(\{a_1\})$ and $\mathcal{S}(\{a_2, a_3\})$. The first of these is $\{a_1\}$ and the second is $\{a_2\}$. However, $\mathcal{S}(\{a_1\} \cup \{a_2, a_3\}) = \{a_1, a_2, a_3\}$. Despite not being chosen in either smaller set, a_3 is among those chosen in the larger set. This is because the larger set contains a cycle and thus the whole cycle is part of the choice set. In the smaller sets, there are no cycles and so only the maximal elements make it into the choice set.

If all the acts available are "linear in $\mathbf{pr}(X)$" for some X, then $\mathcal{M}_{\succeq_{\text{PWR}}}$ is non-empty and the rule satisfies Contraction Consistency. By "linear in $\mathbf{pr}(X)$" I mean $\mathrm{E}_{\mathbf{pr}}(f) = \alpha\,\mathbf{pr}(X) + \beta$ for α, β real numbers. If the graph of expectation against probability is a straight line. This holds of bets on H. Note that this doesn't hold when we take more complex bets. For example in Example 6, you were betting on a coin's landing the same way twice in a row. This is not linear in $\mathbf{pr}(H)$. It is linear in $\mathbf{pr}(HH \vee TT)$, but the other acts on offer are not linear in this event. There may be a more general set of cases where $\mathcal{M}_{\succeq_{\text{PWR}}}$ is guaranteed to be non-empty. Given how similar Table 2 looks to tables of Condorcet cycles in social choice, I conjecture that some sort of single-peakedness property suffices to have the pairwise regret relation to be acyclic (see Gaertner 2009, pp. 43–9). In the simplified circumstances of choosing between a set of well behaved bets that are all simple bets on the same proposition, then $\mathcal{S}$ is a good choice rule since it is more discriminating than $\mathcal{M}_{\succeq_{\text{Dom}}}$ but satisfies many of the same properties. In the more general case however, the possibility of cycles means that the rule does not do so well. $\mathcal{S}$ violates Contraction Consistency less often that $\mathcal{M}_{\mathcal{R}}$ does. There is another practical problem with this rule. Given that it relies on

pairwise comparisons between the acts, and each pairwise comparison requires comparing maxima ranging over probability values, it is quite computationally intensive.

9. Conclusion

We have explored a number of different kinds of choice rule. None is entirely satisfactory. So how should we act? I think we can at least take Non-domination as a requirement on rational choice. So $\mathcal{M}_{\geq_{\text{Dom}}}$ serves to rule out some bad acts. This means that $\varphi \in \mathcal{M}_{\geq_{\text{Dom}}}(A)$ is acting as a necessary *but not sufficient* condition on imprecise choice. A variety of options for going beyond this – to attempt to find sufficient conditions for rational choice – have failed. All the more discriminating rules we have looked at seem to violate one or more intuitively compelling property of rational choice.

We can understand $\mathcal{M}_{\geq_{\text{Dom}}}(A)$ as a weak kind of choice set. That is, it is reasonable to rule out all the acts that $\mathcal{M}_{\geq_{\text{Dom}}}$ rules out. But it seems like some acts that make it into $\mathcal{M}_{\geq_{\text{Dom}}}$ that we would not consider to be reasonable choices (for example q in Example 6). The various attempts to come up with a choice rule that can be given a stronger interpretation have failed. That is, every attempt to construct a choice rule that positively endorses all the acts in the choice set have come up short. ID is such a rule, but it is often empty. I would argue that $\mathcal{S}$ is such a rule, but it suffers from problems with our constraints on rational choice.[26]

In summary, rules like $\mathcal{M}_{\underline{\mathcal{E}}}$ and $\mathcal{M}_{\mathcal{H}_p}$ violate Non-domination and so are not good rules. They also violate Sure Thing Principle. $\mathcal{L}$ violates Union Consistency which might be considered a problem. In any case $\mathcal{L}$ just gets things wrong in cases like the same/different coins case: it rules out the arguably best act. Levi's rule of using $\underline{\mathcal{E}}$ to break ties among elements of $\mathcal{L}$ is doubly bad: it violates Contraction Consistency and Sure Thing Principle *and* it is doubly wrong in the same/different case: as well as ruling out the best act, it makes only the intuitively worst act permissible. $\mathcal{M}_{\mathcal{R}}$ also violates Contraction Consistency and Sure Thing Principle, but has the advantage of seeming to accord with the intuitive judgements about our examples. $\mathcal{S}$ seems to improve on $\mathcal{M}_{\mathcal{R}}$ at the cost of becoming less discriminating. $\mathcal{M}_{\Downarrow\mathcal{P}}$ satisfies pretty much all the conditions we have discussed, but fails to accord with intuitive judgements in the examples. This rule also seems not to be in the spirit of imprecise decision. In short, $\mathcal{M}_{\geq_{\text{Dom}}}$ seems hard to improve on: every proposed improvement, every more discriminating choice rule, has some flaw or other.

What I take myself to have shown here is that we can make some progress on the problem of imprecise choice. It is not the case that when your credences become imprecise, all constraint on choice falls away. In many cases of "moderate" imprecision, the above constraints on choice (in particular Non-domination) will be enough to fix your choice. That is, consider Example 1 again. This was a choice between bets that win if the coin lands heads once or ten times. Since the "ten heads" bet is dominated, your choice is determined in this case.

[26]What I mean is that, if it weren't for the problems it has with Contraction Consistency and the like, being in $\mathcal{S}(A)$ would be a positive endorsement of that act.

When your credences are imprecise, then it's difficult to know how you should act. So it seems strange to criticise imprecise choice for being true to this fact about decision making and uncertainty (as Williamson 2010, pp. 68–72 does). What imprecise choice can do is focus you on what you would need to know in order to be able to make a choice. For example, consider Example 5. In a choice between l and m, neither is obviously better than the other. But, if you were to learn that $\mathbf{pr}(X) > 0.5$, then l would be better than m. So imprecise choice, and its refusal to make discriminations not supported by the evidence, can focus attention on what evidence would help make the decision. This understanding of imprecise choice, therefore, can help us make choices by helping us to understand what evidence we should collect in order to facilitate the decision.

This is obviously an incomplete discussion of imprecise choice. There may be properties of choice rules I have not explored, and there may be choice rules I have not considered. But I hope to have shown that there is interesting work to be done here, and that progress can be made. The progress I have made here consists in arguing that E-admissibility is not an improvement on non-domination; in listing at least some of the properties that we want imprecise choice rules to satisfy; and in urging a more subtle understanding of what rationality requires in cases of severe uncertainty.

Acknowledgements

This research was made possible by the Alexander von Humboldt Foundation. Many thanks to Richard Bradley, Martin Rechenauer, Cedric Paternotte and Sebastian Lutz for comments. Thanks to Jim Joyce and Jon Williamson for comments on a previous version of this work.

A. Properties of relations

At various points we will want to talk about relations among objects. These relations will have various properties that we would like to have names for.

Transitive If $\varphi \succeq \psi$ and $\psi \succeq \rho$ then $\varphi \succeq \rho$

Complete $\varphi \succeq \psi$ or $\psi \succeq \varphi$ for all φ, ψ

Reflexive $\varphi \succeq \varphi$ for all φ

Irreflexive $\neg(\varphi \succeq \varphi)$ for all φ

Symmetric If $\varphi \succeq \psi$ then $\psi \succeq \varphi$

Acyclic For all $\varphi_1, \varphi_2, \ldots, \varphi_n$ we have $\varphi_1 \succ \varphi_2, \varphi_2 \succ \varphi_3, \ldots, \varphi_{n-1} \succ \varphi_n$ implies $\neg(\varphi_n \succ \varphi_1)$

"$\succeq$" will be understood as a reflexive relation, while "$\succ, \sim$" will be the irreflexive and symmetric parts of it.[27] That is, we have the following two definitions.

- $\varphi \succ \psi$ iff $\varphi \succeq \psi$ and $\neg(\psi \succeq \varphi)$
- $\varphi \sim \psi$ iff $\varphi \succeq \psi$ and $\psi \succeq \varphi$

For reflexive $\succeq$ and $\succ, \sim$:

- If $\varphi \succ \psi$ then $\neg(\psi \succ \varphi)$
- $\neg(\varphi \succ \varphi)$

B. Proofs

Theorem 1 *If $\mathcal{C}$ satisfies $\succeq$ and $\mathcal{C}(A)$ is nonempty for nonempty A, then $\mathcal{C}$ pairwise satisifes $\succeq$.*

Proof Assume $\varphi \succ \psi$ and $\mathcal{C}$ satisfies $\succeq$ and is nonempty. Then $\psi \notin \mathcal{C}(\{\varphi, \psi\})$. $\mathcal{C}(\{\varphi, \psi\})$ is a subset of $\{\varphi, \psi\}$, does not contain ψ and is nonempty. Therefore $\mathcal{C}(\{\varphi, \psi\}) = \{\varphi\}$.

Assume $\varphi \succeq \psi$ and $\mathcal{C}$ satisfies $\succeq$ and is nonempty. Now, either $\varphi \succ \psi$ and the above argument shows that $\mathcal{C}(\{\varphi, \psi\}) = \{\varphi\}$, or $\varphi \sim \psi$. Therefore, since $\mathcal{C}$ satisfies $\succeq$, $\varphi \in \mathcal{C}(\{\varphi, \psi\})$ if and only if $\psi \in \mathcal{C}(\{\varphi, \psi\})$. Since $\mathcal{C}$ can't be empty, and must be a subset of $\{\varphi, \psi\}$, $\mathcal{C}(\{\varphi, \psi\}) = \{\varphi, \psi\}$. In either case, $\varphi \in \mathcal{C}(\{\varphi, \psi\})$ as required. ∎

Theorem 2 *(i) $\mathcal{M}_{\succeq}$ is a choice function and (ii) $\mathcal{M}_{\succeq}$ pairwise satisfies $\succeq$. (iii) If $\succeq$ is acyclic on A where A is finite then $\mathcal{M}_{\succeq}(A)$ is non-empty. (iv) Furthermore, if $\succeq$ is transitive, then $\mathcal{M}_{\succeq}$ satisfies $\succeq$.*

[27] Some of the relations in what follows will have subscripts: $\succeq_X$. The relations $\succ_X, \sim_X$ relate to $\succeq_X$ in the obvious way.

Proof (i) $\mathcal{M}_{\geq}(A) \subseteq A$ by definition. It is equally obvious that $\mathcal{M}_{\geq}(\mathcal{M}_{\geq}(A)) = \mathcal{M}_{\geq}(A)$.

(ii) We need to show that if $a \geq b$ then $a \in \mathcal{M}_{\geq}(\{a, b\})$. The only way a could fail to be in $\mathcal{M}_{\geq}(\{a, b\})$ is if $b > a$. But this is ruled out by definition of $>$. If $a > b$ then $a \geq b$, so by the above, we have that $a \in \mathcal{M}_{\geq}(\{a, b\})$, and by definition, $b \notin \mathcal{M}_{\geq}(\{a, b\})$.

(iii) Let $\geq$ be acyclic on some finite A. By a result similar to Lemma 1*l of Sen (1970, p. 16), acyclicity is sufficient to ensure non-emptiness of $\mathcal{M}_{\geq}(A)$.

(iv) If $a > b$ then $b \notin \mathcal{M}_{\geq}(A)$ by definition. Finally, assume for contradiction that $a \sim b$ and $a \in \mathcal{M}_{\geq}(A)$ but $b \notin \mathcal{M}_{\geq}(A)$. This means there exists some $c > b$. But $b \geq a$ so by transitivity[28] $c > a$, contradicting $a \in \mathcal{M}_{\geq}(A)$. ■

Theorem 3 *If $\mathcal{C}$ satisfies $\geq$ then $\mathcal{C}(A) \subseteq \mathcal{M}_{\geq}(A)$ for all A.*

Proof Let $a \in \mathcal{C}(A)$. Assume for contradiction that there is some $b \in A$ such that $b > a$. If there were such a b, then a would not have been in $\mathcal{C}(A)$ by definition of "satisfies". Thus $\neg \exists b \in A, b > a$. This is exactly the condition required for inclusion in $\mathcal{M}_{\geq}$. ■

For the next theorem we will need a little bit more notation. We will use $\varphi \bowtie \psi$ to mean $\neg \varphi \geq \psi$ and $\neg \psi \geq \varphi$. That is, $\varphi \bowtie \psi$ if and only if the two acts are incomparable. We will also need this fact about $\bowtie$.

Lemma 1 *For transitive $\geq$: if $\varphi \sim \psi$ and $\psi \bowtie \rho$ then $\varphi \bowtie \rho$*

Proof Assume $\varphi \sim \psi \bowtie \rho$. Assume for contradiction that $\varphi \geq \rho$. Then $\psi \sim \varphi \geq \rho$ which implies $\psi \geq \rho$ which contradicts our assumptions.[29] Likewise for $\rho \geq \varphi$. Thus $\varphi \bowtie \rho$. ■

Theorem 4 *When* $\mathrm{Opt}_{\geq}(A) \neq \emptyset$, *and $\geq$ is transitive then* $\mathrm{Opt}_{\geq}(A) = \mathcal{M}_{\geq}(A)$.

Proof We first show that $\mathrm{Opt}_{\geq}(A) \subseteq \mathcal{M}_{\geq}(A)$. We then show that if φ is maximal but not optimal, then no act is optimal.

Assume $\varphi \in \mathrm{Opt}_{\geq}$. Assume for contradiction that there is some ψ such that $\psi > \varphi$. Therefore $\neg \varphi \geq \psi$, which contradicts our assumption. Thus $\neg \exists \psi \in A, \psi > \varphi$. This is exactly the criterion for inclusion in $\mathcal{M}_{\geq}(A)$.

Assume now that $\varphi \in \mathcal{M}_{\geq}(A)$ but, $\varphi \notin \mathrm{Opt}_{\geq}(A)$. For φ not to be optimal, this means there is some ψ such that $\neg \varphi \geq \psi$. φ is maximal, so φ and ψ must be incomparable. Assume there is some $\rho \in \mathrm{Opt}_{\geq}(A)$. So $\rho \geq \varphi$, but since φ is maximal, this must mean $\varphi \sim \rho$. $\rho \sim \varphi \bowtie \psi$, therefore $\rho \bowtie \psi$ by the above lemma. In particular $\neg \rho \geq \psi$ which contradicts our assumption. Therefore $\mathrm{Opt}_{\geq}(A)$ is empty. ■

Theorem 5 *$\mathcal{S}(A) \neq \emptyset$ for all finite $A \neq \emptyset$.*

[28] Strictly speaking, we don't really need transitivity here: we only need that $\psi \sim \varphi$ and $\rho > \psi$ imply $\rho > \varphi$.

[29] Strictly speaking we only need something slightly weaker than transitivity: if $\psi \sim \varphi$ and $\varphi \geq \rho$ then $\psi \geq \rho$.

Proof Let $|A|$ mean the size of the set A. If $|A| = 2$ then $\mathcal{M}_{\geq_{\text{PWR}}}(A) \neq \emptyset$. This is just because you can't get the sort of "cycle" behaviour on sets that small. For induction, assume $\mathcal{S}(A) \neq \emptyset$ for $|A| = n$. Now let $|B| = n+1$. The only difficult case is if $\mathcal{M}_{\geq_{\text{PWR}}}(B) = \emptyset$. Thus $\mathcal{S}(B) = \bigcup_{x \in B}\{\mathcal{S}(B \setminus \{x\})\}$. But $|B \setminus \{x\}| = n$ for all $x \in B$. Thus by our assumption, $\mathcal{S}(B \setminus \{x\}) \neq \emptyset$. The union of nonempty sets is nonempty, so $\mathcal{S}(B) \neq \emptyset$. ∎

I have been a little conservative in only aiming to prove the finite case here. I think the proof holds much more generally, and I hope it is clear how one would extend the proof. There may be some awkwardness caused by infinitely long cycles that may require something a little more sophisticated than $\mathcal{S}$ to get the proof through. The idea, at least, should be clear: make a rule where, if you hit a cycle, take the union of the ways you might break the cycle by dropping an element (or possibly a set of elements).

To show that $\mathcal{S}(A) \subseteq \mathcal{M}_{\geq_{\text{Dom}}}(A)$, I first need two short lemmas.

Lemma 2 *If $a \geq_{\text{Dom}} d$ then $a \geq_{\text{PWR}} d$.*

Proof Since $a \geq_{\text{Dom}} d$ we have that $\text{E}_{\mathbf{pr}}(a) - \text{E}_{\mathbf{pr}}(d) \geq 0$ for all $\mathbf{pr} \in \mathcal{P}$. Therefore $\text{E}_{\mathbf{pr}}(d) - \text{E}_{\mathbf{pr}}(a) \leq 0$. From this it follows that $a \geq_{\text{PWR}} d$. ∎

This shows that the $\geq_{\text{PWR}}$ relation is more discriminating than the $\geq_{\text{Dom}}$ relation. Next we need to show we have a *sort of* transitivity.

Lemma 3 *If $c \geq_{\text{PWR}} a$ and $a \geq_{\text{Dom}} d$ then $c \geq_{\text{PWR}} d$.*

Proof $a \geq_{\text{Dom}} d$ so $\text{E}_{\mathbf{pr}}(a) \geq \text{E}_{\mathbf{pr}}(d)$ for all $\mathbf{pr} \in \mathcal{P}$. Therefore:

$$\begin{aligned}
\max\{\text{E}_{\mathbf{pr}}(c) - \text{E}_{\mathbf{pr}}(d)\} &\geq \max\{\text{E}_{\mathbf{pr}}(c) - \text{E}_{\mathbf{pr}}(a)\} \\
&\geq \max\{\text{E}_{\mathbf{pr}}(a) - \text{E}_{\mathbf{pr}}(c)\} \\
&\geq \max\{\text{E}_{\mathbf{pr}}(d) - \text{E}_{\mathbf{pr}}(c)\}
\end{aligned}$$

Therefore $c \geq_{\text{PWR}} d$. ∎

From the above lemmas, we can show the following.

Theorem 6 $\mathcal{S}(A) \subseteq \mathcal{M}_{\geq_{\text{Dom}}}(A)$ *for all $A \subseteq \mathcal{A}$.*

Proof We need to prove that if d is dominated, then $d \notin \mathcal{S}(A)$. Say that d is dominated by a. If $\mathcal{M}_{\geq_{\text{PWR}}}(A) \neq \emptyset$ then, since $a \geq_{\text{PWR}} d$ by Lemma 2, d cannot be in $\mathcal{M}_{\geq_{\text{PWR}}}(A)$.

We now need to show that if $\mathcal{M}_{\geq_{\text{PWR}}}(A) = \emptyset$ we still have $d \notin \mathcal{S}(A)$. The only place we need to check is $\mathcal{S}(A \setminus \{a\})$ since in all other subsets the above result guarantees that d is not in $\mathcal{S}(A)$. If $\mathcal{M}_{\geq_{\text{PWR}}}(A) = \emptyset$ that means, in particular, that there exists some c such that $c \geq_{\text{PWR}} a$. If no such c existed, a would have been in $\mathcal{M}_{\geq_{\text{PWR}}}(A)$. c is in $A \setminus \{a\}$ and $c \geq_{\text{PWR}} d$ by Lemma 3. Therefore, d is not maximal in $A \setminus \{a\}$. ∎

References

Anscombe, F. J. and R. J. Aumann (1963). "A Definition of Subjective Probability". *The Annals of Mathematical Statistics* 34, pp. 199–205.

Bandyopadhyay, P. S. (1994). "In Search of a Pointless Decision Principle". *Philosophy of Science Association Proceedings*, pp. 260–269.

Binmore, K. (2008). *Rational Decisions*. Princeton University Press.

Bradley, R. (2007). "A Unified Bayesian Decision Theory". *Theory and Decision* 63, pp. 233–263.

——— (2009). "Revising Incomplete Attitudes". *Synthese* 171, pp. 235–256.

Bradley, S. and K. Steele (ms.). *Subjective Probabilities Need Not Be Sharp*.

Broome, J. (1991). *Weighing Goods*. Blackwell.

Camerer, C. and M. Weber (1992). "Recent Developments in Modeling Preferences: Uncertainty and Ambiguity". *Journal of Risk and Uncertainty* 5, pp. 325–370.

Chu, F. and J. Y. Halpern (2004). "Great expectations. Part II: Generalized expected utility as a universal decision rule". *Artificial intelligence* 159, pp. 207–230.

——— (2008). "Great expectations. Part I: On the customizability of General Expected Utility". *Theory and Decision* 64, pp. 1–36.

Cozman, F. (n.d.). *A Brief Introduction to the Theory of Sets of Probability Measures*. `http://www.poli.usp.br/p/fabio.cozman/Research/CredalSetsTutorial/quasi-bayesian.html`.

Elga, A. (2010). "Subjective Probabilities should be Sharp". *Philosophers' Imprint* 10.

Ellsberg, D. (1961). "Risk, ambiguity and the Savage axioms". *Quarterly Journal of Economics* 75, pp. 643–696.

Evren, Ö. and E. Ok (2011). "On the multi-utility representation of preference relations". *Journal of Mathematical Economics* 47, pp. 554–563.

Gaertner, W. (2009). *A Primer in Social Choice Theory*. Oxford University Press.

Gärdenfors, P. and N.-E. Sahlin (1982). "Unreliable probabilities, risk taking and decision making". *Synthese* 53, pp. 361–386.

Halpern, J. Y. (2003). *Reasoning about uncertainty*. MIT press.

——— (2006). *Using sets of probability measures to represent uncertainty*. `arXiv:cs/0608028v1`.

Hurwicz, L. (1951). *Optimality Criteria for Decision Making under Ignorance*. Cowles Commission Discussion Paper Statistics 370.

Joyce, J. M. (2005). "How Probabilities Reflect Evidence". *Philosophical Perspectives* 19, pp. 153–178.

——— (2011). "A Defense of Imprecise Credence". *Oxford Studies in Epistemology* 4. forthcoming.

Kaplan, M. (1996). *Decision Theory as Philosophy*. Cambridge University Press.

——— (2010). "In defense of modest probabilism". *Synthese* 176, pp. 41–55.

Kyburg, H. E. (1983). "Rational belief". *The Brain and Behavioural Sciences* 6, pp. 231–273.

Kyburg, H. E. and M. Pittarelli (1992). *Set-based Bayesianism*. Tech. rep. UR CSD;TR407. http://hdl.handle.net/1802/765. University of Rochester, Computer Science Department.
Levi, I. (1974). “On Indeterminate probabilities”. *Journal of Philosophy* 71, pp. 391–418.
——— (1985). “Imprecision and Indeterminacy in Probability Judgment”. *Philosophy of Science* 52, pp. 390–409.
——— (1986). *Hard choices: decision making under unresolved conflict*. Cambridge University Press.
Luce, R. D. and H. Raiffa (1989). *Games and Decisions*. Dover.
Milnor, J. (1951). *Games against Nature*. Tech. rep. RAND corporation.
Peterson, M. (2009). *An Introduction to Decision Theory*. Cambridge University Press.
Schwarz, W. (ms.). *Lost Memories and Useless Coins: Revisiting the Absentminded Driver*. https://github.com/wo/papers/blob/master/driver.pdf?raw=true.
Seidenfeld, T. (1988). “Decision theory without “independence” or without “ordering”. What’s the difference?” *Economics and Philosophy*, pp. 267–290.
——— (2004). “A contrast between two decision rules for use with (convex) sets of probabilities: Γ-maximin versus *E*-admissibility”. *Synthese* 140, pp. 69–88.
Sen, A. (1970). *Collective choice and social welfare*. Holden-Day.
——— (1977). “Social choice theory: A re-examination”. *Econometrica* 45, pp. 53–89.
Smets, P. and R. Kennes (1994). “The Transferable Belief Model”. *Artificial Intelligence* 66, pp. 191–234.
Steele, K. (2007). “Distinguishing indeterminate belief from “risk averse” preference”. *Synthese* 158, pp. 189–205.
——— (2010). “What are the minimal requirements of rational choice? Arguments from the sequential setting”. *Theory and Decision* 68.4, pp. 463–487.
Sturgeon, S. (2008). “Reason and the grain of belief”. *Noûs* 42, pp. 139–165.
Sugden, R. (1985). “Why be Consistent? A Critical Analysis of Consistency Requirements in Choice Theory”. *Economica* 52, pp. 167–183.
van Fraassen, B. (1990). “Figures in a Probability Landscape”. In: *Truth or Consequences*. Ed. by M. Dunn and K. Segerberg. Kluwer, pp. 345–356.
von Neumann, J. and O. Morgenstern ([1944] 2004). *Theory of Games and Economic Behaviour*. Sixtieth Anniversary Edition. Princeton University Press.
Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities*. Vol. 42. Monographs on Statistics and Applied Probability. Chapman and Hall.
——— (2000). “Towards a unified theory of imprecise probabilities”. *International Journal of Approximate Reasoning* 24, pp. 125–148.
Williamson, J. (2010). *In Defense of Objective Bayesianism*. Oxford University Press.