Confidence in Beliefs and Rational Decision Making

Brian Hill*
GREGHEC, CNRS & HEC Paris†

March 19, 2018

Abstract

The standard, Bayesian account of rational belief and decision is often argued to be unable to cope properly with severe uncertainty, of the sort ubiquitous in some areas of policy making. This paper tackles the question of what should replace it as a guide for rational decision making. It defends a recent proposal, which reserves a role for the decision maker's confidence in beliefs. Beyond being able to cope with severe uncertainty, the account has strong normative credentials on the main fronts typically evoked as relevant for rational belief and decision. It fares particularly well, we argue, in comparison to other prominent non-Bayesian models in the literature.

*I wish to thank Richard Bradley, two anonymous referees, and seminar and workshop participants in Amsterdam (ILLC), Bayreuth, Bochum, London (LSE), Munich (Center for Mathematical Philosophy), Paris (Coping with Uncertainty workshop; MDOD), Pisa (SNS), Pittsburgh (Center for Philosophy of Science), Poznan (28th ECOR), Toulouse (IAST) for stimulating discussion and helpful feedback. The author gratefully acknowledges support from the French National Research Agency (ANR) project DUSUCA (ANR-14-CE29-0003-01).

†1 rue de la Libération, 78351 Jouy-en-Josas, France. E-mail: hill@hec.fr. URL: www.hec.fr/hill

1 Introduction

What constitutes rationality for belief and decision? A variety of domains, from epistemology to economics, from decision theory to decision analysis, standardly look to Classic Bayesianism for the answer. Founded on the idea that beliefs admit gradations of strength between the extremes of categorical acceptance or rejection of a proposition—Bayesians often speak of grades of uncertainty, degrees of belief, subjective probability or credences—this position can be summarized in three intertwined tenets, concerning belief, decision making, and learning respectively. This paper focusses on the first two:

1. **A thesis about rational belief:** Gradations of belief strength are represented by the assignment of a single number (between 0 to 1) to each proposition or event. These numbers satisfy the laws of probability.

2. **A thesis about rational decision:** The chosen action in any decision is that which maximises the *expected utility* or *desirability* on the basis of the agent's graded beliefs.

Bayesianism owes its status as the benchmark account of rational belief and decision making largely to its purported *coherence with normative intuitions*. Some justify the expected utility rule as directly capturing or following from some *normatively appealing pre-formal intuition or principle concerning how choices should be made*; Weirich (2001, Ch 3), for instance, purports to derive it directly from a 'principle of pros and cons'. A more popular route defends the account on the basis of the *normative plausibility of its implications for the choices that are made*. Dutch Book arguments are of this sort: by purportedly showing that only those with probabilistic beliefs will never accept a set of bets yielding a sure loss (or 'Dutch Book') they harness the spontaneous normative attractiveness of this behavioral consequence in support of the Bayesian position (see for example de Finetti 1937; Hájek 2008). The axiomatisations common in economic decision theory can be put to similar use: they establish a set of properties of preferences—the 'axioms'—that characterise decision makers who can be represented as adhering to the Bayesian tenets, and hence allow one to argue for the latter by appealing to the normative intuitiveness of the former (Ramsey, 1931; Savage, 1954; Gilboa, 2009; Gilboa et al., 2010; Cozic and Hill, 2015).

Another important advantage of the Bayesian account relates to its *scope*: in particular, it purportedly applies to groups as well as individuals. Despite the difficulties in connecting individual and group attitudes and decisions (Mongin, 1995), few deny the attractiveness of a single account of rationality applying at both levels. Much of Bayesianism's capacity to do this in practice is due to its *conceptual clarity*: it supports a neat separation of doxastic attitudes—beliefs, uncertainty judgements—which are entirely summarized by the probability measure, and conative attitudes—desires, values, tastes—which are fully captured by the utility function.[1] In many social contexts, one collection of people supplies the judgements about knowledge and uncertainty, whilst another determines the relevant values: in policy decisions about the environment, energy investments, drug safety and many other domains, it seems desirable for the experts to deliver the facts and policy makers (or some other representatives of society) to provide the values. The neat separation of the uncertainty or belief element from the value or taste one allows a Bayesian decision procedure to support such practice. Moreover, it allows the possibility of the value-free communication of beliefs that this practice requires: without it, any judgement that fully summarises the beliefs of an expert will concern not just the facts, but will inevitably be 'contaminated' by value judgements.

Despite these qualities, the Bayesian hegemony as a normative account of belief and decision making has been increasingly challenged, both by philosophers (Levi, 1974, 1986; Joyce, 2011a; Bradley, 2009) and economists (Gilboa et al., 2009; Gilboa and Marinacci, 2013), as well as in fields such as decision analysis (Lempert and Collins, 2007; Cox, 2012). In a word, the suggestion is that it suffers from significant *limitations in its domain of application*: there is an important class of 'severe uncertainty' situations where it is not appropriate. A typical example concerns an event about which information or evidence is scant, and contrasts it with one where it is plentiful. To take a case in the style of Ellsberg (1961), consider two urns each containing only black and white balls: for one of the urns (the unsampled urn), that is all you know; for the other (the sampled urn), you have observed 1 million draws (with replacement), half of which were black. Bayesianism enjoins you to have a precise degree of belief about the colour of the next ball drawn, for each urn—say, $\frac{1}{2}$ in it being black for both urns. Note that, given the contrast in the amount of evidence supporting these judgements, it is natural to be more sure of the degree of belief concerning the sampled urn than the unsampled one.[2] However Bayesianism ignores such differences

[1] 'Doxastic', from the Greek *doxa* ('opinion'), is the term used to qualify attitudes that have the character of beliefs, and 'conative', from the Latin *conari* ('to endeavour'), denotes attitudes related to desire or volition.

[2] Bayesianism has been argued to reflect something akin to this difference in the resilience of the proba-

when it comes to decision, as can be seen when comparing your attitudes to choosing in the two cases: would you prefer to bet on the colour of the next ball drawn from the sampled urn or the unsampled one? Under the Bayesian account, since the degrees of belief concerning the events are the same, you must be indifferent between the bets, despite the differences in how sure you are in the relevant degrees of belief. By contrast, if, as many people do, you prefer betting on the sampled urn, then it seems that you are taking such a factor into account in your decision. Indeed, these preferences, which typically violate the Bayesian 'axioms' (Ellsberg, 1961), have been argued to be perfectly reasonable from a normative perspective on such grounds (Levi, 1986; Gilboa et al., 2009).

Whilst artificial, the moral of this example extends to more realistic and significant decisions. Compare two patients: for one, all the tests support the doctor's degree of belief of $\frac{2}{3}$ that he has a particular disease which calls for a specific invasive treatment; for the other, the evidence is contradictory, but the doctor's best-guess judgement for her having the disease is again $\frac{2}{3}$. As above, Bayesianism requires that the same treatment be recommended in both cases; but would it be unreasonable for the doctor to be more cautious in his recommendations for the second patient? Compare our world with climate change to a counterfactual one where there is none: in the former, climate science cannot justify precise probabilistic judgements for future regional climate patterns; in the latter, statistics on past climate would provide a much greater deal of precision. Many infrastructure decisions—say, whether to build flood or drought defenses—depend on such climate forecasts, and Bayesianism dictates that the decisions should be taken in the same way in both worlds. In particular, it recommends the same policies in both worlds whenever the best-guess probabilities coincide. But would it not be more reasonable to take how unsure we are about regional climate forecasts into account when making policy decisions in the face of climate change, as recommended by some risk analysts (for example, Lempert and Collins 2007; Cox 2012)?

Health and climate decisions are arguably among those where normative guidance is most needed. Bayesianism's inability to render the widely-shared intuition that how sure we are in the decision-relevant judgements may reasonably have consequences for choice thus counts as a critical weakness. Is there a better account of rational belief and decision to be had?

bility judgements in the face of new information (Skyrms, 1977). This claim, which pertains to learning or belief formation, does not affect the central point made here concerning decision, namely that such differences are denied any role in choice.

Certainly, there is no lack of models that purport to capture the specificity of the Ellsberg examples: the economic literature on decision theory has spawned a plethora of 'ambiguity' models motivated by them, not to mention work in philosophy and statistics on 'imprecise probabilities'. However, there has been no comprehensive comparative discussion of their strengths and weaknesses as normative accounts. But it is not enough for an account to accommodate the behaviour in Ellsberg-style examples: it should *also* retain as many as possible of the attractive characteristics of the Bayesian benchmark. We need non-Bayesian alternatives with strong normative credentials across the board.

This paper will defend the account of belief and decision developed in Hill (2013a, 2016) on these grounds. At its heart is the notion of confidence: not the confidence in the truth of a proposition—which Bayesian degrees of belief are supposed to capture—but rather one's *confidence in one's beliefs* themselves. To avoid confusion and clumsiness, we tie down the multifarious term 'confidence' for the purposes of this paper and use it to speak of one's attitude of being more or less sure of one's beliefs. As such, it is a doxastic attitude—part of an agent's state of belief. Following standard terminology, we shall use the terms 'belief', 'degree of belief' or 'credence' for the dimension (degree of endorsement of a proposition) considered by standard Bayesianism. One way of formulating the central thesis is that rational individuals' states of belief—their doxastic states—do not necessary comprise only their beliefs, but include their confidence in their beliefs.

The previous examples suggest that confidence in beliefs has a role in decision making: they are all cases where one's behaviour seems to be sensitive to how sure, or confident, one is of the relevant beliefs. Indeed, the proposal comprises an approach to rational decision making that incorporates confidence, according to the following *prima facie* reasonable maxim: the higher the stakes involved in the decision, the more confidence is required in a belief for it to play a role. This paper will focus on defending and evaluating the proposal as an account of rational belief and decision; whilst there is much to be said about the role of confidence in belief formation—its relationship to evidence, for instance—we shall not be concerned with such issues here.

We shall first present the confidence approach (Section 2), before turning to a detailed evaluation of its normative credentials (Section 3). In Section 4, we compare it on this front with some other recent proposals. Whilst the paper mainly focusses on the normative question, Section 5 briefly discusses some prescriptive issues, relating to the tractability of the approach for applications.

2 Confidence in beliefs and decision: the proposal

The approach defended here comprises an account of beliefs (and confidence in them), and of their role in decision.[3] It draws upon, but differs significantly from, a popular current approach, often called 'imprecise probabilities' in philosophy or 'multiple priors' in economics. To ease exposition, as well as to elucidate the relationship with the existing literature, we shall present and discuss it in comparison with this latter approach.

2.1 A model of confidence in beliefs

According to the imprecise probabilities representation (defended by Levi 1986; Joyce 2011b, for example), an individual's state of belief is represented not by a single probability (or credence) measure but by a set $\mathcal{C}$ of such measures.[4] As pointed out by Joyce (2011b), such sets can be thought of as a formal representation of the agent's doxastic situation. So, for example, an agent will have a higher degree of belief for a proposition A than B if $p(A) \geq p(B)$ for all probability measures p in the set $\mathcal{C}$. Similarly, her degree of belief in A will be greater than (respectively equal to) $\frac{1}{2}$ if $p(A) \geq \frac{1}{2}$ (resp. $p(A) = \frac{1}{2}$) for all p in $\mathcal{C}$. Let us call statements about degrees of belief or credences—such as 'A has a higher degree of belief than B', 'A has a higher degree of belief than $\frac{1}{2}$', 'A is probabilistically independent of B' and so on—*credal statements*, or *credal judgements*. It is well-known that the set of probability measures involved in the imprecise probability representation can be 'lifted' to the level of credal statements (for instance Halpern, 2003, Ch. 7). Each set of probability measures $\mathcal{C}$ generates a collection of credal statements: those that hold for all of the probability

[3] In adopting the distinction between beliefs and decision, which is standard in the philosophy and economics literatures (see, for instance, Joyce 1999, 2011a; Gilboa 2009; Bradley 2016), we by no means wish to take a position on the relationship between the two. In particular, the discussion here is independent of whether beliefs are taken to be 'defined' or 'revealed' from preferences—as often assumed in the economics or parts of the statistics (Cozic and Hill, 2015)—or rather are conceptually primitive. All that is assumed is that there is a meaningful distinction, in particular between the representation and role of beliefs in the determination of preferences and the preferences themselves, which also depend on the decision maker's desires or values. For further discussion of the behavioural consequences of the account, see Section 3.2.

[4] Whilst the statistical literature on 'imprecise probabilities' is vast, and comprises several mathematical models (see for instance Walley 2000; Augustin et al. 2014), the set of probability measures model is doubtless the most prominent in philosophical discussion. Henceforth we use the term 'imprecise probabilities' to refer to this model. We discuss several other models sometimes placed under the 'imprecise probabilities' label, such as Dempster-Shafer belief functions, in Section 4. In our presentation, we also largely ignore technical details, involving for instance continuity issues, which are tangential to the main points made.

measures in the set. These are the credal statements to which the agent represented by $\mathcal{C}$ adheres. Sometimes this is presented in terms of a committee metaphor. Considering the probability measures in the set to be members of a committee, the accepted credal statements are those held unanimously—that is, by all members of the committee.

Given this, the imprecise probability representation has an immediate interpretation in terms of confidence in beliefs. An agent adheres to—and is confident in—each credal statement that holds for all probability measures in the set; she does not adhere to—and hence has no confidence in—the other credal statements. As a representation of confidence in beliefs, imprecise probabilities are evidently unsatisfactory, for they treat confidence as an all-or-nothing affair: either you hold a credal judgement with full confidence, or you do not hold it, and have no confidence at all. It does not allow for grades of confidence, of the sort seen above. For instance, it cannot represent an agent who, in the urn example, holds the credence of $\frac{1}{2}$ for drawing black for each of the urns, but who is more confident in the judgement for one of the urns than for the other.

To capture such confidence comparisons, the proposal is to replace the single set of probability measures by a *nested family* of such sets: that is, a family where each member is contained in or contains each other member (Hill, 2013a). Such a nested family is called a *confidence ranking*. The sets in the family correspond to levels of confidence, with larger sets corresponding to higher levels (see Figure 1). As noted above, each set generates a collection of credal statements: these represent the credal judgements the agent holds to the corresponding level of confidence. For larger sets in the family, corresponding to higher confidence levels, the generated collections of credal statements are smaller, and so fewer credal judgements are held by the agent at higher confidence levels (as one would expect).

Just as sets of probability measures correspond to collections of credal statements, confidence rankings induce an order on credal statements, which captures the relative confidence that the agent has in them. Credal statements that hold for all probability measures in larger sets are held with higher confidence than those that hold only in smaller sets.[5] So, for example, if the $p(A) = \frac{1}{2}$ for all probability measures in a small set in the confidence ranking, but not in larger sets, but $p(B) = \frac{1}{2}$ for all probability functions in a larger set in the confidence ranking, then this captures an agent who is more confident in her assessment of $\frac{1}{2}$ for her degree of belief in B than in her assess-

[5]And these are held with higher confidence than those statements that hold in none of the sets—which themselves correspond to statements that the agent does not adhere to.

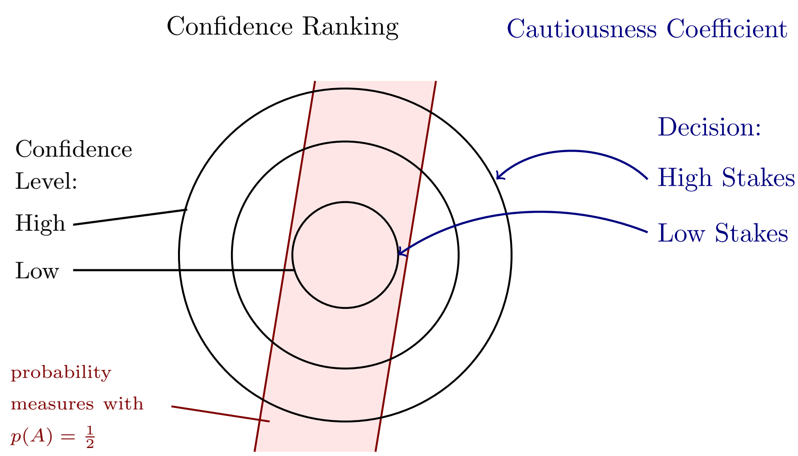

Figure 1: Representation of confidence in beliefs (black) and relation to decision (blue)

ment of $\frac{1}{2}$ for her degree of belief in A (see Figure 1 for a graphical representation of the confidence in judgements concerning A). Taking as A the event that the next ball drawn from the unsampled urn is black, and similarly for B and the sampled urn, confidence rankings can thus faithfully render the differing confidence levels in the previous example. In terms of the committee metaphor, confidence rankings invite one to think of a group with a hierarchical structure—at the centre, there are the leading scientists (say, members of the Academies), then there is a larger collection including all full professors, then a level with all active researchers, and so on up to the set of all members of the scientific community. A credal statement unanimously held by all leading scientists is adhered to by the community, but perhaps only with limited confidence, whereas one which is unanimously held by all members has high confidence.

Whilst we adopt the terminology used by Hill (2013a), it is but one of a family of representations based on similar ideas. To our knowledge, the first was proposed by Gärdenfors and Sahlin (1982), who use a real-valued measure of 'epistemic reliability' over the space of probability measures. The confidence ranking discussed above can be obtained from such a measure by 'throwing away' the numbers and keeping just the order over probability measures.[6] Nau (1992) develops a notion of 'confidence-weighted probabilities', under which each probability statement is indexed by a real-valued confidence number; again, the confidence ranking contains just the ordinal information,

[6]In this sense, the confidence ranking is ordinal whilst the epistemic reliability measure is cardinal (see also Section 5). As Gärdenfors and Sahlin (1982) note, they only require the order established by their epistemic reliability measure in their paper.

but not the cardinal information (numbers) involved. Indeed, the confidence ranking is related to ordinal representations in the literature on belief revision: where some models there (see Gärdenfors 1988; Grove 1988 for example) amount to orders on the set of states of the world, the confidence ranking is essentially an order on the space of probability measures. The aforementioned authors do not necessarily share the same account of how confidence is related to decision, a question to which we now turn.

2.2 Confidence in belief and decision making

As mentioned, we shall defend this account of belief in tandem with a story about decision. The examples in the Introduction attest to the importance of confidence in beliefs for choice. People's confidence in their belief may play a role in their decision making, and rightly so. But what sort of role should it play? The account we defend is based on the following maxim:

Maxim the higher the stakes involved in the decision, the more confidence is required in a belief for it to play a role.

This appears to be a sensible way of relating two aspects of a decision: its importance (or the stakes involved in it), and the beliefs one relies on to take it. It shall be discussed in more detail below. For the moment, notice that it directly motivates the following formal framework for decision.

We first assume that to each decision or option the decision maker is faced with, she can associate a level of confidence appropriate for it. As noted above, the confidence levels correspond to sets in the confidence ranking: so assigning a confidence level to a decision amounts to assigning a set in the confidence ranking. Moreover, the maxim requires that the assignment is made on the basis of the stakes involved: more important decisions—or those involving larger stakes—call for more confidence, and are thus associated to higher confidence levels, which correspond to larger sets in the confidence ranking. In summary then, we take a function D that assigns a set in the confidence ranking to each decision, such that decisions with higher stakes are sent to larger sets (see Figure 1). Such a function is called a *cautiousness coefficient*. As shall be discussed in Section 3.3, the cautiousness coefficient can be understood as a reflection of certain of the decision maker's attitudes, in much the same way as the utility function in standard Bayesianism is often interpreted as a representation of her desires.

The suggested rendition of the aforementioned maxim is simple: to evaluate an option, use the set of probability measures in the confidence ranking that corresponds to the decision at hand according to the cautiousness coefficient. Why? This amounts to using the credal judgements held to the corresponding level of confidence. But this is the level picked out (by the cautiousness coefficient) as being appropriate for the decision at hand, on the basis of the stakes involved. So using this set of probability measures basically means that the agent only relies on beliefs that she holds with enough confidence given the stakes involved in the decision. This procedure is thus faithful to the maxim.

The proposal does not amount to a single decision rule as much as a family of rules. Indeed, it just picks out a set of probability measures—or an 'imprecise probability measure'—but does not specify how to choose on the basis of it. Several decision rules for imprecise probabilities have been proposed in the aforementioned literatures; each one of these, when inserted into the framework, will result in a corresponding confidence-based decision rule. For example, using the maximin-EU decision rule (also called Γ-Maximin in robust statistics; Gilboa and Schmeidler 1989; Berger 1985), which looks at the lowest expected utility calculated across the set of probability measures, naturally yields a rule which evaluates an act f according to:

$$\min_{p \in D(f)} EU_p f \tag{1}$$

where $EU_p f$ is the expected utility of f calculated with probability p and utility U,[7] and D is a cautiousness coefficient assigning to every act a confidence level. Alternatively, if one uses the unanimity rule (or maximality; Bewley 2002; Walley 1991), then act f will be chosen over g if and only if:

$$EU_p f > EU_p g \qquad \text{for all } p \in D((f, g)) \tag{2}$$

where D is a cautiousness coefficient assigning to every binary choice (pairs of acts) a confidence level.[8]

As the two examples illustrate, models in the family may also differ on their treatment of the stakes associated to a decision. (1) implicitly assumes the stakes to be

[7] Since the focus here is on belief, we follow standard Bayesianism in assuming throughout the paper a precise utility or desirability function as a representation of desires over outcomes.

[8] There are different versions of this rule depending on the sort of dominance required (for example, strict or weak order in (2)); such details are orthogonal to the present discussion.

assigned to each option separately, and so the cautiousness coefficient is defined on them; (2) takes stakes to be assigned to the choice (ie. sets of options on offer), and so involves a cautiousness coefficient defined on those. For further technical details and discussions of these models, their relationship, and stakes, readers are referred to Hill (2013a) (for (1)) and Hill (2016) (for (2)).[9]

These two examples also illustrate how confidence-based decision rules are basically extensions of (corresponding) imprecise probability or ambiguity decision rules. For example, the standard maximin-EU rule is just like (1) except that $D(f)$ in the minimum is replaced by a fixed set $\mathcal{C}$, and similarly for the standard unanimity rule and (2). So the confidence-based family of rules can account for any choice patterns that imprecise probabilities can. For instance, in the previous example of betting on sampled or unsampled urns, just as the maxmin-EU model can account for a preference for betting on the sampled urn, so can the corresponding confidence model, (1). At any reasonable confidence level, the decision maker will endorse the credal statement that the probability of getting black from the sampled urn is $\frac{1}{2}$; by contrast, whenever the confidence level is high enough, she may not hold such a precise judgement on the probability of getting black from the unsampled urn, instead restricting herself to intervals, such as $[\frac{1}{4}, \frac{3}{4}]$. When the stakes are high enough to merit such a confidence level, she will use $\frac{1}{2}$ to evaluate the act of betting on the sampled urn, whilst look at the minimal expected utility over $[\frac{1}{4}, \frac{3}{4}]$ in evaluating the bet on the unsampled urn. Since the latter value is lower than the former, she will prefer to bet on the sampled urn. Similar points hold for the confidence-based unanimity rule (2) (see Hill, 2016).[10]

So the account is essentially a generalisation of standard approaches for sets of probabilities or imprecise probabilities. Is anything gained by this generalisation?

[9] In their axiomatic analyses, the cited papers assume the appropriate notion of stakes as given. A subsequent paper (Hill, 2015) dispenses with this assumption. We sidestep such technicalities and present a simplified version of the approach here, which is in line with the 2015 paper. See the cited papers for further discussion, and the 2016 paper on the interdefinability between stakes on options and stakes on choices.

[10] As for the representation of confidence discussed in the previous section, the account of decision here is related to others in the literature. Although Gärdenfors and Sahlin (1982) do not propose a formal model of how the confidence level is related to the decision at hand (and hence lack the notion of cautiousness coefficient), (1) is close to the sort of decision procedure they discuss. The model proposed by Nau (1992) is roughly a reduced form of a special case of (2) (see Hill, 2016), which lacks the distinction between confidence ranking and cautiousness coefficient.

3 Why confidence? An appraisal

To evaluate the approach just set out, we will consider how it fares on the points typically raised in favour of Bayesianism. Recall the main ones from the Introduction. Two concern an account's coherence with normative intuitions—be it with some normatively appealing *pre-formal intuition* captured by the rule or the attractiveness of its *implications for choice*. A further one concerns its scope, and whether it can fruitfully apply to both individuals and groups; on this front, an account's *conceptual clarity*—in particular whether it supports a neat separation of doxastic and conative attitudes—is crucial. We now consider these in turn, comparing, where relevant, with the 'imprecise probability' approach mentioned previously. (The relationship to other non-Bayesian approaches will be discussed in the next section.)

3.1 Pre-formal intuition

A decision procedure built on reasonable and easily explainable normative principles or intuitions would *ceteris paribus* seem preferable to one that is not, and some have argued for certain decision rules on such grounds. To the extent that the confidence-based proposal was built on a reasonable non-formal maxim, it should be no surprise that it can be defended on this front.

Firstly, the underlying maxim—the higher the stakes, the more confidence is required of a belief for it to play a role in the decision—might itself be defendable on independent grounds. It calls for a level of adequacy between the decision to be taken and the means—in particular the beliefs—mobilised in the taking of the decision. As such, it can be thought of as a consequence of the following more general principle:

Appropriateness the tools employed in the execution of task should be appropriate for the task at hand.

Considering one's beliefs as (among) the tools, and the decision as the task, the upshot would be a demand for some appropriateness of the former for the latter. Of course, any reasonable account of decision involves some form of appropriateness, in particular in the 'domain' of the beliefs. Beliefs about the weather tomorrow are irrelevant to (and inappropriate for use in) decisions about the investment of one's fortune. To employ a tool analogy, this would be like noting that a screwdriver is the wrong tool for taking down a dividing wall—what is needed is a hammer. However, the current proposal goes further, looking not only at the appropriateness in terms of the domain, but also in terms of the 'intensity'. A medium-sized hammer is the appropriate tool for breaking up a bookcase, a big hammer is appropriate for taking down a dividing wall, and a wrecking ball is appropriate for demolishing a building. Using too big a hammer or too small a hammer would be foolish (though perhaps not as foolish as using a screwdriver). Likewise, demanding excessive confidence in (the relevant) beliefs to use them in the most trivial decisions appears unnecessarily pedantic, just as it may seem irresponsible to rely entirely on hunches (beliefs in which one has little confidence), if avoidable, in decisions where many lives are at stake. Note that, whilst the confidence-based account captures this dimension of appropriateness of beliefs for decision, many others in the literature do not. For instance, the Bayesian approach mobilises all relevant beliefs—all the information concerning probability judgements about relevant events—in the expected utility formula, apparently giving no heed to such appropriateness considerations.

This principle, insofar as it concerns the intensity dimension of appropriateness, ties into a long tradition in philosophy, going back at least as far as Aristotle's views on virtues, which emphasises the importance of avoiding extremes in favour of the 'mean'. Indeed, the demand for some adequacy of the confidence level required of beliefs to the decision at hand reflects a sense of proportion that is often related to virtue in general, and rationality in particular.

Moreover, any intuition that can be claimed by the general maxim is inherited, in perhaps a more concrete form, by (reasonable) members of the proposed family of decision rules. Take the confidence-based maximin-EU model (1). Under this rule, when the act under evaluation involves higher stakes, the designated confidence level is higher, the decision maker relies on fewer beliefs (the set $D(f)$ is bigger), and so the evaluation is more pessimistic or cautious (the range of expected utility values over the larger set of probability measures is larger, so the minimum is lower). By contrast, when the stakes are low or the decision maker is particularly confident in the relevant beliefs, the set of expected utility values is smaller, and the evaluation is less pessimistic. So the rule embodies the following principle, which can be thought of as a special case of the general maxim above: choose boldly when one has sufficient confidence for the decision at hand; choose cautiously if not.

To our knowledge, non-expected utility decision rules in the literature are rarely defended by relating them to normatively appealing pre-formal principles such as this. Certainly, the standard maximin-EU rule does not faithfully reflect a maxim of this sort: it uses the same set of probability measures irrespective of the stakes—and so advises the same degree of boldness or caution. Indeed, this rule is sometimes criticised for being too cautious, insofar as it only looks at the worst case.[11] So, to go back to the urn example from the Introduction, if the set $\mathcal{C}_0 = \{p : 0 \leq p(black) \leq 1\}$ is used to evaluate a bet on black with \$1 billion at stake, then it is also used to evaluate a bet on black with \$1 at stake. In the latter case, at least, this may seem too cautious. The confidence-based refinement (1) provides some relief from this criticism: the caution exhibited is sensitive to both the decision maker's (lack of) confidence in the relevant beliefs and the importance of the decision. For instance, whilst $\mathcal{C}_0$ may be used when billions of dollars are at stake, a smaller set—even a single probability measure—could be used where there are only a few dollars at stake. So the decision maker displays less caution in the latter decisions compared to the former. Nothing suggests that the basic point that confidence-based models are (pre-theoretically) more normatively reasonable than their imprecise probability counterparts does not extend beyond the case of the maximin-EU rule.

3.2 Implications for choice

A highly influential family of arguments seek to justify the Bayesian account of belief and decision on the basis of the normative plausibility of its consequences for choice. For instance, classic Representation Theorems (Savage, 1954; Anscombe and Aumann, 1963; Gilboa, 2009; Gilboa and Marinacci, 2013) bring out these consequences in a form of a set of 'axioms'—properties of preferences—that hold of all and only decision makers whose behaviour is consistent with the Bayesian tenets. To the extent that these axioms can be argued to characterise rational behaviour, they support the normative pretentions of the underlying Bayesian account.

Existing research into the confidence-based family includes several Representation Theorems (Hill, 2013a, 2016) that play a similar role of bringing out the behavioural consequences of the confidence-based approach. These are formulated in a common, albeit technical setup in the economic literature on decision theory; evaluation of the axioms thus requires some explanation of the framework. However the general morals of these results can be brought out, perhaps more distinctly, in the much simpler context of the standard Dutch Book Argument. Whilst controversial, this is sometimes held as a typical pragmatic argument in favour of the Bayesian representation of belief.[12]

[11]The extent to which this criticism is fair may depend on how one interprets the set of probability measures in the rule; see for example Gilboa (Ch 18, 2009).

[12]On the relationship between Representation Theorems and Dutch Book Arguments, see Gilboa (2009), for example. Note that our aim is not to enter into the debate into the validity of Dutch Book Arguments,

The standard argument goes as follows. For each event A, consider the bet, with stakes S, yielding $\$S$ if A and $\$0$ if not. You are asked to price every such bet: that is, give the monetary value $\$qS$ for which you would be indifferent between buying and selling the bet. q (or $q(A)$ when the event is not evident from the context) is called the *betting quotient* for the event A. The argument invokes a characterisation of probability measures in terms of properties of betting quotients, sometimes known as the Dutch Book Theorem. It states that the values $q(A)$ satisfy the laws of the probability calculus if and only if you are not vulnerable to a Dutch Book—a sets of bets that, taken together, lead to a sure loss.[13] But, the thought goes, accepting bets that lead to a sure loss has to be irrational. Interpreting the betting quotients as your degrees of belief, this, the argument goes, establishes that they should be probabilities.

Does this mean that anyone who diverges from the Bayesian tenets—and in particular anyone who adopts the confidence-based approach—leaves himself open to Dutch Books? It turns out that the answer is no, because the argument involves some auxiliary assumptions, which are debatable. First of all, it rests on an assumption of *Buy-sell coincidence*: that the highest price for which you are willing to buy a bet is equal to the lowest price for which you are willing to sell it. But there is no reason why there should necessarily be a 'knife-edge' price at which you are willing to both buy and sell a given bet. A decision maker who knows that there is at least 10 black balls and 20 white balls in an urn containing 100 balls may be willing to buy a $\$1$ million bet on the next ball drawn being black for $\$0.1$ million but not more, and she may be willing to sell this bet for $\$0.8$ million but not less. It is not clear why this is irrational, or indeed why rationality should dictate that she specify a price between $\$0.1$ million and $\$0.8$ million at which she would be willing to both buy and sell the bet.

In the light of this, it would seem reasonable to specify *two* values for each gamble: $\underline{q}(A)$, where $\$\underline{q}(A)S$ is the most you would be willing to buy a bet on A with stakes S for, and $\overline{q}(A)$, where $\$\overline{q}(A)S$ is the least you would be willing to sell a bet on A with stakes S for. It is well-known that the standard Dutch Book Theorem no longer holds under such a weakening: satisfying the laws of probability is no longer the only

but to bring out the differences—in terms of behaviour—between the confidence approach and others. We thus accept for the sake of the exercise all the standard assumptions made in the Dutch Book framework, including linearity and state-independence of utility, and act-independence of states. Moreover, despite the relationship to no-arbitrage arguments in the finance literature, we follow standard philosophical treatments in ignoring the market dimension in the exposition here. For similar reasons, we adopt a simple presentation, ignoring technical details, some of which are provided in the Appendix.

[13] Some authors distinguish one direction of the implication (which they call the Dutch Book Theorem) from the other (the Converse Dutch Book Theorem); see for example Hájek (2008).

way to guarantee avoiding Dutch Books. Setting one's betting quotients according to (standard) imprecise probabilities also ensures Dutch Book invulnerability (Smith, 1961; Walley, 1991).

However, this is one way among many (Walley, 1991, Ch 2 & 3): invulnerability to Dutch Books does not force one's betting quotients to be set according to imprecise probabilities. To obtain a characterisation of imprecise probabilities, further conditions are required. (For readers interested in the technical details, the Appendix states one such characterisation for general gambles; see Walley 1991, Ch 2 & 3 for a thorough treatment.) Since precise probabilities are a special case, all of these conditions are also satisfied by the Bayesian approach. One such condition is *Stakes-Independence*: that the betting quotient is independent of the stakes.

However, like Buy-sell coincidence, the rational credentials of this principle are far from obvious. There is no reason to expect you to price bets the same way irrespective of the stakes involved. In the previous example of an urn containing at least 10 black balls out of 100, the decision maker may well be willing to pay much more than $\$0.10$ to buy a bet on the next ball drawn from the urn being black when only $\$1$ is at stake. This would suggest that the betting quotients relevant for buying or selling bets may depend on the stakes involved. Certainly, such dependence does not appear to be irrational. Moreover, it naturally seems to go in a particular direction: when the stakes are higher, the decision maker may reasonably refuse to buy or sell at betting quotients that she would have accepted at lower stakes.[14]

For an event A and stakes S, let $\$\underline{q}_S(A)S$ be the most you would be willing to buy a bet on A with stakes S for. Stakes-Independence demands that, for any stakes S and T, $\underline{q}_S(A) = \underline{q}_T(A)$. Whilst this is too strong, the observation above suggests that $\underline{q}_S(A) \leq \underline{q}_T(A)$ when the stakes S are higher than T: that is, any betting quotient accepted at higher stakes is accepted at lower stakes, but not necessarily vice versa. (Similar points hold for selling bets.)

If one takes a characterisation of imprecise probabilities and weakens Stakes-Independence in this way, one obtains a characterisation of the confidence-based approach; see the Appendix for details. That is, swapping Stakes-Independence for this form of stakes-dependence implies, in the presence of the other conditions yielding imprecise probabilities, that betting quotients are effectively derived from a confidence ranking and a cautiousness coefficient. For example, the betting quotient $\underline{q}_S(A)$ for a

[14] Armendt (2010), in the context of a discussion of stakes-sensitivity of beliefs, also questions Stakes-Independence, whilst holding on to Buy-Sell Coincidence.

bet on A at stakes S will be the worst-case probability for A over the set of probabilities measures in the confidence ranking corresponding to stakes S (according to the cautiousness coefficient).

Note first of all that this suggests a 'rule-of-thumb' way of understanding confidence in beliefs. For all its faults, the interpretation of degrees of belief in terms of betting quotients at the heart of the Dutch Book Argument gives a useful grasp on the concept, which can help guide intuition. The previous discussion suggests a similar 'proxy' for confidence: the confidence in a degree of belief is reflected in the stakes to which one is willing to let that degree of belief guide one's betting behaviour. As such, the introduction of confidence in belief appears a natural addition to degrees of belief: beyond the odds one gets (reflecting degrees of belief), there is the issue of how much one is willing to bet on those odds (reflecting confidence in those beliefs).[15]

More importantly, as mentioned above, such behavioural characterisations can be used to guage the account's normative credentials. As for the original Dutch Book Theorem, the characterisation tells us that, to the extent that the conditions involved can be argued to be rational, they provide support for the confidence-based approach. In particular, it clarifies that the behavioural differences between the Bayesian, imprecise probability and confidence approaches are not to be found in the vulnerability to Dutch Books: they are all invulnerable to them (see Appendix for details). Rather, the normative 'battleground' is pinpointed to two conditions: Buy-sell coincidence and Stakes-Independence. Denying that these constitute rational obligations leads to the confidence approach, whereas accepting one or both yields more standard accounts.

This moral generalises beyond the simple Dutch Book framework, as evidenced by the aforementioned representation results for cases (1) and (2) of the confidence family (Hill, 2013a, 2016, Thms 1). They confirm that a first choice-based difference between confidence-based models and standard Bayesian expected utility theory is common with imprecise probabilities. The behavioural difference between the confidence-based and imprecise probability approaches fundamentally boils down to the issue of stakes independence. The latter, but not the former, assume that preferences are, in an appropriate sense, independent of stakes.[16]

[15]Of course, this is only a rough proxy: just as the standard rendition of degrees of beliefs as betting odds neglects the specificities of the utility function, thinking of confidence in terms of stakes ignores the role of the cautiousness coefficient.

[16]The cited results make it clear that the stakes independence at issue cannot be captured by some property of the utility function—a point that may not come out clearly in the Dutch Book framework, given the assumption of linear utility.

The first difference is the subject of a long-standing debate, focussing mainly on whether non-Bayesian models are embarassed in dynamic or sequential choice situations. Roughly, accommodating the behaviour in the severe-uncertainty examples discussed in the Introduction requires that one relinquish either an axiom (or choice property) called completeness (the equivalent of Buy-Sell Coincidence in the previous discussion) or the independence axiom (or sure-thing principle). Different dynamic arguments have been proposed against the violation of each of these axioms. Whilst there is no space here to enter into the details, two remarks are in order. Firstly, non-Bayesian replies proposed to date either hold onto independence and defend violations of completeness (as recommended by Seidenfeld 1988; Bradley and Steele 2016, for instance), or retain completeness at the price of independence (a more common route in the ambiguity literature, see Machina and Siniscalchi, 2013). Both of these reactions are available to the defender of the confidence approach: (2) is an example of a rule retaining independence but dropping completeness (Hill, 2016), whereas (1) holds onto completeness at the price of independence (Hill, 2013a). Secondly, whilst some have tried to refute the dynamic arguments or their menace for non-Bayesian approaches (Bradley and Steele, 2013, 2016; Hill, 2013b), it suffices that Bayesianism's limitations in the sorts of severe-uncertainty situations discussed in the Introduction outweigh any advantage it might have as regards dynamic choice. Several have argued that this is indeed the case (Gilboa et al., 2009; Siniscalchi, 2009). As suggested at the outset, we adopt such a view here, and refer the interested reader to the cited papers for further discussion of these dynamic arguments.

The second difference—the weakening of stakes independence—is what sets the current proposal apart from other non-Bayesian accounts. And it is far from unreasonable. On the contrary, stakes independence appears to be overly restrictive as a normative condition, and as discussed previously, may be reasonably violated in some cases. So, for anyone who thinks that Bayesianism's troubles with severe uncertainty outweigh its purported dynamic advantages, such as proponents of imprecise probability, there are no choice-based reasons not to shift to the confidence-based approach.

3.3 Conceptual clarity

One attractive feature of a prospective account of rational belief and decision is that it apply to both individuals and groups. Since in group settings doxastic and conative attitudes—beliefs and values or tastes—may be under the remit of different actors, this basically requires a neat separation of these two sorts of attitude. Under the standard interpretation, the Bayesian model delivers such a separation: the state of belief is entirely summarized by the probability measure, while the desires, values or tastes are fully captured by the utility function. Moreover, this is not a mere artefact of the expected utility formula: in some areas of economics it is formalised in ‘comparative statics’ results which show, more or less, that modifications of the utility function lead to changes in the ‘taste’ aspects of choices.[17]

Compared to Bayesianism, the confidence-based approach involves two novel elements: the confidence ranking and the cautiousness coefficient.[18] As explained in Section 2.1, the confidence ranking captures the decision maker’s state of belief, incorporating in particular her confidence in her beliefs. As for the cautiousness coefficient, it can be understood as a representation of her attitude to choosing in face of limited confidence. This interpretation is suggested by its role in the model. It involves a judgement as to the appropriate confidence level for the decision at hand, and hence reflects the extent to which the decision maker is willing to rely on beliefs held with limited confidence in such a decision. Suppose Ann and Bob have the same confidence ranking, and are each evaluating the bet on black from the unsampled urn in the Introduction with stakes of \$1 billion. Suppose that Ann’s cautiousness coefficient assigns this decision to the set $\mathcal{C}_0 = \{p : 0 \leq p(black) \leq 1\}$ in their confidence ranking, whereas Bob’s assigns it to the smaller set $\mathcal{C}_1 = \{p : 0.25 \leq p(black) \leq 0.75\}$. Since $\mathcal{C}_1$ is in Ann’s confidence ranking, it represents beliefs that she holds (eg. she holds a credence for black greater than or equal to 0.25); however, she feels uncomfortable relying on beliefs held with that level of confidence in such a high-stakes decision. Bob, by contrast, is less averse to mobilising beliefs held with this much confidence in decisions of such importance. If you will, he is readier to take the ‘epistemic risk’ of relying on beliefs held with limited confidence when the stakes are so high. Ann and Bob differ in their attitudes, or tastes, for choosing on the basis of beliefs held with limited confidence.

The important point is that the cautiousness coefficient is conative in character: it reflects a taste or value judgement, rather than something of the order of a belief. The model thus neatly separates the doxastic element—fully captured by the confidence ranking—from conative attitudes—reflected entirely by the utility function and

[17] A paradigmatic example is the standard analysis showing that (under expected utility) differences in risk aversion correspond to specific comparisons in the utility function (Arrow, 1971; Pratt, 1964), which is often taken to confirm that it fully captures attitudes to risk.

[18] The final element in the models is the utility function, which, as standard, can be interpreted as reflecting the decision maker’s desires for outcomes, and hence deserves no further discussion here.

the cautiousness coefficient.

This is corroborated by the sort of 'comparative statics' considerations common in the economic literature on decision. This literature has developed a preference-based notion of relative attitude to uncertainty that can compare the extent to which one decision maker is more averse to options involving uncertainty than another.[19] For example, under a typical notion of this sort, Ann is more uncertainty averse than Bob if whenever she chooses an uncertain option over one that involves no uncertainty, then so does Bob.[20] Such notions are generally intended to be the equivalent for uncertainty of the standard economic notion of comparative risk aversion (Pratt, 1964; Arrow, 1971) and, as such, reflect decision makers' tastes for bearing uncertainty. By looking at what comparisons in terms of such notions correspond to at the level of the primitives of the model, one can draw conclusions about which primitives reflect this sort of taste. Under the confidence approach, they correspond to differences in the cautiousness coefficient, corroborating the interpretation of it as reflecting a taste (see Hill, 2013a, Thm 2 and Hill, 2016, Cor 1).

Such a clean separation of doxastic and conative attitudes turns out to be fairly rare in the non-Bayesian world. In particular, decision rules built on imprecise probabilities generally lack it, as can be illustrated on the maximin-EU rule. Recall that under this model, an act f is evaluated according to:

$$\min_{p \in \mathcal{C}} EU_p f \tag{3}$$

where $\mathcal{C}$ is a set of probability measures (and the rest of the notation is as specified in Section 2.2). A tempting, and perhaps even popular interpretation of $\mathcal{C}$ is as representing the decision maker's state of belief: after all, it seems to be the equivalent in this model of the Bayesian probability, which is supposed to represent beliefs. However, this interpretation does not fit well with the sorts of 'comparative statics' exercises alluded to above. In particular, under the maximin-EU model, relative uncertainty aversion—a taste notion—corresponds to differences in the set of probability measures $\mathcal{C}$: if Ann is more uncertainty averse than Bob, then $\mathcal{C}_{Ann}$ contains $\mathcal{C}_{Bob}$ (Ghirardato and Marinacci, 2002, Thm 17). So how are we to understand the set of probability mea-

[19]We use the term 'uncertainty' here in the economists' sense, covering cases where probabilities are not given, as opposed to situations of risk, where they are.

[20]We give the general sense of the notion; the precise statement distinguishes between risk and uncertainty (see previous foonote), and corrects for differences in utilities between the decision makers that are compared. The reader is referred to Ghirardato and Marinacci (2002); Gilboa and Marinacci (2013) for such technical details.

sures in this model: as capturing the decision maker's beliefs or her tastes for bearing uncertainty?[21]

In the economic literature, one generally draws the conclusion that there is no clean interpretation of the set $\mathcal{C}$: it reflects aspects of both belief and uncertainty attitude (see Klibanoff et al., 2005, Sect 3 & 5.1, for example). Consider the unsampled urn from the Introduction. On the basis of the 'objective' information available, any composition of the urn is possible; so the information is summarised by a set of probability measures $\mathcal{C}_0 = \{p : 0 \leq p(black) \leq 1\}$. What are we to say about a decision maker who chooses in this situation according to the maximin-EU rule, but with $\mathcal{C}_1 = \{p : 0.25 \leq p(black) \leq 0.75\}$ instead of $\mathcal{C}_0$? Does she have further beliefs, beyond the available information, that allow her to restrict the set of probability measures? Or does the restriction of this set reflect a greater tolerance of uncertainty—or less cautious attitude—on her part? The basic point is that the use of imprecise probabilities in the context of the maximin-EU model is not rich enough to decide this question—or, indeed, to represent the difference between these two possibilities. To that extent, it fails to support a clear interpretation of the set of probability measures.[22]

Although such comparative statics considerations have received relatively little traction in the philosophical literature, they can be seen as indicative of deeper, interrelated problems, concerning belief communication and incorporation of evidence. For instance, since the set of probability measures *can* reflect the decision maker's attitude to uncertainty, how are we sure, when an agent reports a set of probability measures in good faith for use to guide choice in the context of such a rule, that she is not inadvertantly letting her tastes for uncertainty contaminate her report, and the subsequent choice? Such an issue has been raised in the literature on (experimental) elicitation of imprecise probabilities (Smithson, 2014; Yaniv and Foster, 1995, 1997). Reporting probability intervals requires subjects to trade-off between the accuracy of the estimate

[21]Note that, under the revealed preference results for the maximin-EU model (Gilboa and Schmeidler, 1989), the representing set of priors is (essentially) unique, suggesting that the issue of separation is distinct from that of the uniqueness of the ingredients of the representation.

[22]Whilst we have just discussed the maximin-EU rule, these considerations (and those below) appear to generalise to other decision rules for imprecise probabilities, such as the standard version of the unanimity rule (Section 2). For some rules, the situation is further complicated by issues with the representation and its uniqueness, as appears to be the case for the Hurwicz or α-maximin-EU rule, which evaluates an act by the (α-)mixture of the minimum and maximum expected utilities over a set (Gilboa and Marinacci, 2013). However, refinements of imprecise probability decision models that explain how the set $\mathcal{C}$ 'results' from beliefs and uncertainty attitudes might be able to exhibit the desired separation (a potential example is Gajdos et al., 2008).

and its informativeness, so the interpretation of any intervals elicited depends on how subjects make these trade-offs. To the extent that they may involve value judgements (as to whether it is better to be more precise but wrong, or not, for the decision in hand), this is basically a consequence of the lack of a clear separation between doxastic and conative attitudes.

In practice, the sorts of trade-offs just mentioned are often related to the incorporation of evidence. One thought could be that imprecise-probability decision makers adopt (and report) the set of all measures that are consistent with their evidence: since this set is 'objectively' defined, there is no risk of infiltration of values. However, such a set is obviously too large in many situations: for instance, in the case of the *sampled* urn in the Introduction, with 1 million observations, this is the set of all probability measures except those giving probability one to black or to white (Walley, 1991). So they have to cut down the set of probability measures they report or use in the maximin-EU rule. However, this can involve weighing up not only the strength of the evidence but also how cautious one wants to be (which is reflected in the size of the set), and the imprecise probability framework provides no tools for separating the purely doxastic considerations from those involving value judgements. As noted, this is particularly problematic for the use of these models to guide public decision making, insofar as it jeopardises value-free communication of beliefs.

In summary, the confidence framework offers a clear story about its central elements.[23] On the one hand, there are beliefs and confidence in them, represented by the confidence ranking. On the other hand, there are tastes for, or value judgements concerning choosing on the basis of limited confidence.[24] Whilst lacking for some popular non-Bayesian approaches, and in particular imprecise probabilities, a clear separation of this sort is central for public decision making: in application of the confidence approach to such decisions, one should look to the experts to provide the confidence ranking, and to the policy maker to fix the cautiousness coefficient.

The upshot is that the confidence-based approach can not only cope comfortably with the severe-uncertainty situations where Bayesianism struggles, but also fairs well on the normative fronts typically raised in its favour. To summarize: the approach is based on a reasonable pre-formal intuition, its hallmark in terms of implications for

[23]Note that this does not hold for the accounts of confidence in belief cited in Sections 2.1 and 2.2 that lack the distinction between the confidence ranking and the cautiousness coefficient.

[24]It should come as no surprise that, compared to the Bayesian expected utility model, there is a new conative element: as is well-known (Gilboa, 2009; Gilboa and Marinacci, 2013), the Bayesian model is uncertainty neutral, whereas other decision rules may allow for differing attitudes to, or tastes for, uncertainty.

choice—dependence on stakes—is far from a sign of irrationality, and it supports a clean separation of beliefs and desires (or tastes). This provides a strong case in favour of the approach as an adequate account of rational belief and decision. Indeed, the discussion suggests that it has better normative credentials than a leading non-Bayesian approach, that of imprecise probabilities. We now briefly consider some other major non-Bayesian proposals.

4 Some other non-Bayesian approaches

There is a large non-Bayesian literature on belief and decision, and we cannot hope to treat it in full. Here we briefly compare the confidence approach to some major accounts other than imprecise probabilities, concentrating in particular on those with some motivation in the Ellsberg-type examples, and which are not entirely focussed on descriptive (rather than normative) questions.

One strand of the literature retains the assumption that the belief concerning an event can be fully summarised in a (single) real number, but denies that they must satisfy the laws of probability. Belief functions (Dempster, 1967; Shafer, 1976) are examples of such representations that have been widely studied in statistics and philosophy. Whilst these functions have been motivated drawing on considerations pertaining to learning and evidence,[25] they are known to be equivalent to a special class of sets of probability measures, so the points made above regarding the use of imprecise probabilities to guide decision carry over to them. More generally, they are special cases of the non-additive probabilities (or, to use mathematical terminology, capacities) studied in economics (Schmeidler, 1989). As is well-known, the main decision rule involving such functions that does not violate a dominance principle and several other standard axioms is the Choquet Expected Utility rule (Schmeidler, 1989; Gilboa and Marinacci, 2013). That is, an act f is evaluated according to:[26]

$$\sum_{x_i} \nu(\{s : U(f(s)) \geq x_i\}) \, [x_i - x_{i+1}] \tag{4}$$

[25] As stated in the Introduction, we do not consider the issue of learning here, and focus uniquely on belief and decision aspects in this discussion.

[26] For ease of exposition, we assume throughout that everything (states, outcomes, supports of probability measures etc) is finite, and so use sums in the place of integrals. Note that, when ν is a belief function (or more generally a convex capacity), (4) is equivalent to the maximin-EU rule over a derived set of probability measures (Schmeidler, 1989; Gilboa, 2009).

where ν is the capacity. (U is the utility function; x_i are the utility values of outcomes, organised in decreasing order.) It has proved difficult to give a solid pre-formal normative intuition or justification for the use of this rule to guide choice under uncertainty. As concerns implications for choice, it involves a weakening of expected utility comparable to that yielding the maximin-EU rule (Gilboa, 2009; Gilboa and Marinacci, 2013), though not necessarily involving aversion to uncertainty. For our purposes the essential point is that, like imprecise probability rules, it assumes stakes independence (Hill, 2013a). Moreover, on the conceptual front, results similar to those cited above for the maximin-EU rule (Section 3.3) suggest that there is no clean separation of beliefs and tastes: ν, which is often presented as a representation of the state of belief, also reflects uncertainty attitude (Ghirardato and Marinacci, 2002, Thm 17). In summary, as concerns its normative credentials for rational decision making, the non-additive probability approach does not clearly do better than the imprecise probability one discussed previously.

A large family of recent approaches use second-order representations on the space of probability measures, in a way akin to ours, and are sometimes interpreted in terms of confidence. Examples in the decision-theoretic literature include the variational preferences model (Maccheroni et al., 2006) and the so-called confidence model (Chateauneuf and Faro, 2009).[27] The former is closely related to the literature on robustness in macroeconomics: one of the models developed by Hansen and Sargent (2001) is a special case. The latter is motivated by and technically related to the literature on fuzzy sets. Despite differences in the details, each employs a real-valued function on the space of probability measures, which is sometimes interpreted as representing confidence (Marinacci, 2015; Chateauneuf and Faro, 2009). Moreover, as is clear from the sorts of comparative statics results alluded to previously, these models do not cleanly separate doxastic and conative attitudes: the real-valued functions in question, which are the only elements in these models that could reflect beliefs, capture uncertainty attitudes (Maccheroni et al., 2006, Prop 8; Chateauneuf and Faro, 2009, Prop 8). As concerns their choice-theoretical properties, they are relatively mild weakenings of the maximin-EU decision rule (3), though we are aware of no defense of their specific weakenings on grounds of rationality. They are motivated by the relationship to the robustness literature in macroeconomics and engineering, or the notion of fuzzy sets

[27]For completeness: the former evaluates an act f by $\min_{p\in\Delta}\left(EU_p f + c(p)\right)$ where c is a real-valued function on the space of probability measures Δ, and the latter evaluates it by $\min_{p\in L_{\alpha_0}\phi} \frac{1}{\phi(p)} EU_p f$ where ϕ is a $[0,1]$-valued function on Δ and $L_{\alpha_0}\phi$ is a set of probability measures depending on ϕ and a number α_0.

respectively, but, to our knowledge, no other pre-formal normative intuition has been proposed for these rules.

Perhaps the most popular 'second-order' approach represents the state of belief by a second-order probability over first-order probability measures (over events). Such second-order probabilities have been discussed in the philosophical literature by Skyrms (1980), for example. The most natural decision rule involving such a representation applies expected utility at both stages, evaluating an act f by:

$$\sum (EU_p f)\, \mu(p) \tag{5}$$

where μ is the second-order probability (and the sum is taken over all first-order probabilities to which it gives non-zero weight). However, this representation is easily seen to be equivalent to the standard expected utility representation with the 'reduced' probability $\sum p\mu(p)$; hence it does no better than the Bayesian theory at accommodating the uncertainty-sensitive behaviour mentioned in the Introduction.[28] Recently, researchers in economics have proposed the following variant:

$$\sum \phi\,(EU_p f)\, \mu(p) \tag{6}$$

where ϕ is a real-valued function on utility values (in much the same way that the utility function U is real-valued function on outcomes). This *smooth ambiguity* representation, most forcefully defended by Klibanoff et al. (2005)[29] and increasingly popular in economic modelling, can accommodate Ellsberg behaviour when ϕ is non-linear. They emphasise that this model admits a separation of beliefs from uncertainty attitudes: the second-order probability μ can be understood as a representation of the decision maker's state of belief, whereas the transformation function ϕ represents her attitudes to uncertainty. This interpretation is backed up by the sort of comparative statics considerations discussed above (Klibanoff et al., 2005, Sect 3).

As concerns the approach's normative credentials, a central question is clear from (6): if the decision maker can form precise second-order probabilities, which she can 'reduce' to precise first-order probabilities, then why doesn't she just use those—or

[28] Moreover, as noted by Gilboa and Marinacci (2013), it seems to ignore the difficulty in providing precise probabilities in, for example, climate decisions (Bradley and Steele, 2015).

[29] Related approaches have been proposed by Nau (2006); Ergin and Gul (2009); Seo (2009) with early work by Segal (1987).

equivalently (5)—to choose? This issue translates, in one of the most elegant characterisations of this sort of model, into a violation of the axiom of reduction of lotteries (Seo, 2009, Cor 5.2), which essentially says that the decision maker should treat a 20% chance of having a 50% chance of winning a prize the same as having a 10% chance of winning. It has been suggested that such violations, which ressemble inabilities to properly multiply probabilities, may undermine the rational credentials of the approach.

Perhaps the most explicit reply to these objections, and the only one we are aware of, is proposed by Marinacci (2015). He defends the use of representations of the form (6) for decision analysis, relying on the distinction between 'physical uncertainty'—essentially the randomness in the relevant mechanisms or processes—and 'epistemic uncertainty' (or 'model uncertainty')—reflecting the decision maker's lack of knowledge about the underlying mechanism. He suggests an interpretation in which the first-order probabilities in (6) correspond to physical uncertainty and the second-order probabilities (the μ) capture epistemic uncertainty. The idea is that these correspond to different 'sources of uncertainty' and that it is legitimate to have 'different attitudes toward the two uncertainty sources' (ibid. p1052). This is precisely what representation (6) does: U represents the attitude to physical uncertainty and $\phi \circ U$ the attitude to epistemic uncertainty. Marinacci explains this clearly showing that, when all of the uncertainty is physical, U is the relevant utility used by the model, whereas when all of the uncertainty is epistemic, $\phi \circ U$ is used instead. In particular, the same probability distribution will lead to different evaluations under this model according to whether it captures physical uncertainty (with no epistemic uncertainty) or epistemic uncertainty (with no physical uncertainty). As he puts it:

> 'different confidence in such [probability] judgements (whatever feature of a source causes it) translate as different degrees of aversion to uncertainty across sources, and so in different von Neumann-Morgenstern utility functions [U and $\phi \circ U$].' (Marinacci, 2015, p1052)

But the formal translation of the relationship between sources of uncertainty, confidence and uncertainty attitudes into decision rule (6) risks undermining one of its most vaunted qualities. Confidence (in this context, at least) is a doxastic attitude: more confidence means that you are more sure about your beliefs. However, the suggestion seems to be that differences of confidence 'translate' or correspond to something partially reflected in the transformation function ϕ—which was only supposed to capture uncertainty attitudes, that is tastes for bearing uncertainty. Formally rendering a doxastic judgement in what was supposed to be a conative element of the model seems to jeopardise its purported separation of beliefs and tastes.

Indeed, this defense seems to face a dilemma. Either the separation of beliefs and tastes argued for by Klibanoff et al. (2005) holds, so the state of belief is fully captured by the (second-order) probability distribution—but this risks undermining the normative defense of the use of the transformation function ϕ proposed above. Consider the aforementioned example where the same probability distribution is evaluated differently according to whether it represents physical or model uncertainty. Under the separation of beliefs and tastes, this probability fully captures the decision maker's state of belief—and hence her confidence in beliefs—concerning the physical and epistemic sources of uncertainty respectively. But since the probability distribution is the same for both sources, her confidence is the same in both cases, and thus there would seem to be no justification for different uncertainty attitudes, contrary to the claim in the quote above.

The other horn of the dilemma endorses the defense proposed above, thus admitting that aspects of a decision maker's belief state, in particular her confidence in probability judgements, are captured by the transformation function rather than the probability distribution. But this appears to clash with the separation of beliefs and tastes. To illustrate, compare three situations faced by a policy maker: (1) scientific experts provide a single probability distribution, in which they are very confident, reflecting entirely physical uncertainty (with no epistemic uncertainty); (2) the experts provide the same distribution, in which they are very confident, but it is entirely epistemic uncertainty (with no physical uncertainty); (3) the experts are not confident in any distribution as capturing the epistemic uncertainty, but when pushed for a precise distribution (as the model demands) provide the same one as in (2) (fully epistemic uncertainty, no physical uncertainty).[30] Representation (6) allows the policy maker to decide differently in situations (1) and (2), according to her ϕ. She is supposed to be able to justify such a difference on the basis of differing confidence in the probability judgements. But this justification is hard to square with the expectation that the *experts* are the best judges of how much confidence there is, and the fact that *they* are equally confident in (1) and (2). Moreover, the representation implies that the policy maker should take the same decision in situations (2) and (3), since the probability distribution reported and the source

[30]For instance: in (1), there is an objectively chancy mechanism determining the quantity of interest, which is fully understood, while in (2) and (3), the underlying process is fully deterministic (and predictable), but the scientists have (more or less severe) uncertainty about its properties.

are the same (epistemic uncertainty in both cases). Following the previous reasoning, this would seem to suggest that the confidence in the probability judgements is the same between the two situations. But here the experts do judge there to be difference in confidence, although the model, in asking them only for a probability distribution, does not give them the means to express this difference. Under this horn of the dilemma, the model leaves the judgement on a doxastic issue (confidence in probability judgements) to the actor who should be determining the values: an indication that attitudes might not be properly separated. As noted previously, and as illustrated by this example, this may lead to problems in the application of the model in public decision making. There thus seems to be a tension between the proposed defense of the rational credentials of the smooth ambiguity model and its promise of providing a clear separation of beliefs and attitudes to uncertainty.

The idea of 'source dependence' behind this interpretation of (6) is common to several approaches in economics and psychology,[31] and the central point seems to apply to this literature more generally. Most 'source dependent' models share two characteristics that are central to the preceding dilemma: the assumption of precise probabilistic beliefs *within* each source; and the assumption of potentially differing conative attitudes *towards* sources, which are crucial in accounting for Ellsberg-type examples (such as those in the Introduction). As argued above, any defense of the rationality of differing attitudes towards sources on the basis of confidence in beliefs jeopardises the clean separation of beliefs and tastes in the primitives of the model. This will have to be taken into account when evaluating the normative credentials of such accounts.

Bradley (2015) proposes a different interpretation of (6) that makes no reference to source dependence. There, the first-order probabilities are objective chances and the second-order probabilities represent beliefs,[32] but there is no confidence or uncer-

[31]The notion of source dependence is often traced to experimental work by Tversky and Fox (1995); Fox and Tversky (1995), and plays a central role in current prominent approaches (Abdellaoui et al., 2011; Wakker, 2010). As for the other accounts discussed here, the focus is entirely on the normative question, leaving aside considerations pertaining to the approaches' descriptive relevance.

[32]He is applying the (philosophical) distinction between objective chances and subjective probabilities. It is unclear to what extent it coincides with the distinction between physical and epistemic (or model) uncertainty, as it is used in practice. In particular, it is crucial for Bradley's position that objective chances are 'features of the world'. By contrast, a typical application of (6) to climate change, for example, takes as a proxy for 'physical uncertainty' probability distributions of the relevant climate variable drawn from the literature (Millner et al., 2012; Marinacci, 2015), but given the known inexactness of the climate models and the Bayesian methods, including prior probabilities (sometimes provided by experts), used to provide such distributions, it is unclear that they should necessarily be interpreted as objective 'features of the world'.

tainty attitude: ϕ represents the decision maker's attitude towards objective chances. As Bradley points out, his account is fully Bayesian concerning uncertainty: it diverges with the standard account only on the case of risk (in particular, attitudes to objective chances). As such, it does not treat the general issue under discussion here: that of belief representation and decision in the absence of readily available precise probabilities.[33]

The size of the literature prohibits an extensive review, and we can only encourage further evaluation and development of models in the light of the criteria considered. Nevertheless, this brief discussion of several prominent approaches suggests some tentative conclusions. First of all, the confidence-based account is relatively rare in claiming a clean separation of doxastic and conative attitudes, the smooth ambiguity model doubtless being the main existing proposal in the literature to be associated with this property. However, the previous considerations suggest that further work is required on the normative foundations of that model: there seems to be a deep tension between its claims of normative plausibility and separation of attitudes. Hence, the confidence-based account would seem to be the only approach to date to possess a pre-formal normatively plausible intuition and a clean separation of attitudes, as well as reasonable implications for choice relative to other non-Bayesian approaches. This only bolsters the case for it as an adequate account of rational belief and decision.

5 On Tractability

Whilst this paper is dedicated to the normative question, it is perhaps worth mentioning the related prescriptive issue. An important driver of the use of a model of beliefs and decision is *tractability*: not so much whether it provides a reasonable guide to rational choice, but rather how easy it is to actually implement in real-life cases, such as policy decisions. Whilst there is nothing better than actual application to bring out the strengths and weaknesses of the approach defended here, some comments on this topic are perhaps in order.

The first concerns how to find the optimal choice in complex decisions. Since it piggybacks on existing models, one would expect existing methods and techniques to extend to the confidence approach. For instance, it is common to use specific parametri-

[33]Though Bradley (2015) shows that non-neutral attitudes to chances can account for the standard Ellsberg behaviour without calling into question the Bayesian position on beliefs, he does not suggest that it can be fruitfully applied to the other cases mentioned in the Introduction, such as climate decisions.

sations of the set of priors in the standard maximin-EU model (3), under which there exist techniques for calculating optima in decision problems. Examples include sets of priors which are 'balls' centered on a given measure, such as the ϵ-contaminations popular in robust statistics (Berger, 1985) or the 'entropy' balls used in the robustness literature in macroeconomics (Hansen and Sargent, 2001).[34] Such parametrisations—and thus the optimisation techniques relying on them—can be easily extended to the confidence framework. It suffices to take as confidence ranking the set of all balls, of differing radii, centered on a given probability measure. Moreover, the confidence approach provides a story on how to fix the radius of the ball—the main free parameter in the standard accounts—via a specific value judgement reflected in the cautiousness coefficient.

A second issue, which is particularly relevant for the motivating examples where current science and statistics do not provide (reasonably justfied) precise probabilities, is that of the *elicitation* of the beliefs required by the model, from an expert for instance.[35] Of course, simpler representations of belief states generally require less information from the agent, and so are usually easier to elicit. The representation of the belief state under the current proposal—the confidence ranking—certainly seems more complicated than the Bayesian representation (by a probability measure), or the standard imprecise probability representation (a set of probability measures). However, it is, in a certain sense, the 'least complicated step up' from the latter, insofar as it is *ordinal* at the second-order level—it only involves an order on the space of probability measures (Hill, 2013a, Prop 2). So, to elicit a confidence ranking from an agent, it is sufficient to collect her *qualitative* confidence comparisons between credal judgements, ie. comparisons of the sort: I am more confident in the credence for A being greater than 0.5 than in the credence for B being less than 0.3.

By contrast, under the other second-order representations mentioned in Section 4, the numbers count: the representations are *cardinal* at the second-order level. So more information is required to pin down an agent's belief state under this representation: not just *whether* she is more confident in one judgement than another, but *how much*

[34] Hansen and Sargent (2001) show that for a class of decision problems, maximin-EU with such balls yields the same decisions as a subclass of variational preferences (Section 4), which themselves correspond to a special case of the second-order probability model in the style of (6) (Strzalecki, 2011). So techniques developed for any of these models can be mobilised to solve optimisation problems under corresponding versions of the confidence model.

[35] The importance of expert elicitation has been emphasised in several of the domains mentioned as motivation (see for example Morgan, 2014).

more confident she is. These *quantitive* comparisons—for example, I am confident to degree 0.7 in the credence for A being greater than 0.5 but only confident to degree 0.6 in the credence for B being less than 0.3—are significantly more difficult to extract from agents. So these models are more demanding on an expert who is to provide the doxastic judgements for use in guiding decision.

So whilst not the simplest representation of beliefs, the proposed confidence representation is at least at the simple end of the spectrum: it is ordinal. Indeed, it is the only non-Bayesian approach we are aware of that both provides a clean separation between doxastic and conative attitudes and is ordinal at the second-order level. This suggests that, in principle at least, it may be more applicable in situations where opinions need to be elicited from experts.

6 Conclusion

Decisions under severe uncertainty are becoming increasingly relevant. The Bayesian benchmark for rational belief and decision fails to provide a reasonable guide in such cases; this paper looks at the issue of what, if anything, should replace it. An adequate account should not only cope with severe uncertainty, but it should have strong normative credentials across the board. We defend a particular approach on these grounds, founded on the intuition that one's confidence in one's beliefs has a role to play in decision making. The confidence-based framework is argued to possess a normatively plausible pre-formal intuition, to have relatively reasonable consequences for choice, and to clearly separate the roles of beliefs on the one hand, and desires, values or tastes on the other. It appears to be unique in the existent literature to possess all these qualities. Moreover, the framework defended involves a simpler representation of beliefs than some other recent approaches, which may prove useful for the elicitation of opinions from experts.

Appendix: Behavioural Characterisations

In this appendix, we state simple behavioural characterisations of the confidence-based and imprecise probability approaches, which underlie the discussion in Section 3.2. We add this material to keep the paper self-contained and permit a simple comparison: the technical material is either drawn almost directly from the literature, or uses techniques developed elsewhere.

We adopt the following fairly standard setup. Consider a set of states Ω; for simplicity of exposition, we assume it to be finite. Gambles, or random values, are real-valued functions on Ω. A bet on an event A with stakes S is a gamble paying out S when $\omega \in A$ and 0 if not. (Addition of gambles and multiplication by real numbers is defined pointwise, as standard.) For a gamble X and a probability measure p on Ω, $E_p(X) = \sum_{\omega\in\Omega} X(\omega)p(\omega)$ is the expectation of X with respect to p. For a gamble X, the stakes involved in X are given by its maximum absolute value $S_X = \max_{\omega\in\Omega} |X(\omega)|$. Unit gambles are gambles with stakes of 1. For any gamble X, let $\overline{X}$ be the associated ('normalised') unit gamble: $\overline{X}(\omega) = \frac{X(\omega)}{S_X}$ for all $\omega \in \Omega$. For any gamble X and positive real number S, the gamble X with stakes S is given by $X^S(\omega) = S\overline{X}(\omega)$ for all $\omega \in \Omega$; when the stakes $S = S_X$, specific mention of them is omitted. As in the text, $\underline{q_S}(X)$ is the *lower betting quotient at stakes* S, where $\underline{q_S}(X)S$ is the highest amount for which the agent is willing to buy the gamble X with stakes S; similarly, $\overline{q_S}(X)$ is the *upper betting quotient at stakes* S, where $\overline{q_S}(X)S$ is the lowest amount for which the agent is willing to sell the gamble X with stakes S. Note that $\overline{q_S}(X)$ is definable from $\underline{q_S}(X)$ in the standard way: $\overline{q_S}(X) = -\underline{q_S}(-X)$. So the clauses below on selling gambles are unnecessary, but added for completeness.

We say that a set of betting quotients $\underline{q}_S, \overline{q}_S$ is *derived from a confidence ranking and a cautiousness coefficient* if there exist a confidence ranking Ξ and a cautiousness coefficient D assigning the set $D(S)$ to any gamble with stakes S such that $\underline{q}_S(X) = \min_{p\in D(S)} E_p(\overline{X})$ and $\overline{q}_S(X) = \max_{p\in D(S)} E_p(\overline{X})$ for every gamble $\overline{X}$ and stakes level (positive real number) S. In particular, the betting quotient for a bet on an event A at stakes S is the worst case probability that a decision maker using the confidence ranking deems possible for this event, at the level of confidence corresponding to that level of stakes (according to the cautiousness coefficient). Similarly, a set of betting quotients is *derived from a set of probability measures* if there exists set of probability measures $\mathcal{C}$ such that $\underline{q}_S(X) = \min_{p\in\mathcal{C}} E_p(\overline{X})$ and $\overline{q}_S(X) = \max_{p\in\mathcal{C}} E_p(\overline{X})$ for every gamble X and stakes level (positive real number) S. It is *derived from a* *probability measure* if there exists a probability measure p such that if there exists set of probability measures $\mathcal{C}$ such that $\underline{q}_S(X) = E_p(\overline{X}) = \overline{q}_S(X)$ for every gamble X and stakes level (positive real number) S.

We now formally state several conditions. The first is the standard Dutch Book invulnerability condition; the next three were discussed in detail in Section 3.2.

Dutch Book Invulnerability If the agent is willing to buy gambles $X_1, \ldots X_n$ at prices $p_{1}, \ldots, p_n$ respectively, then $\max \sum_{i=1}^{n} (X_i - p_i) \geq 0$.

Buy-sell coincidence For every gamble X and stakes level S, $\underline{q_S}(X) = \overline{q_S}(X)$.

Stakes-Independence For any positive S, T, the agent is willing to buy the gamble X with stakes S for qS if and only if she is willing to buy the gamble X with stakes T for qT. (And similarly for selling gambles.)

Stakes-Dependence If the agent is willing to buy the gamble X with stakes S for qS, then for any $T \leq S$ she is willing to buy the gamble X with stakes T for qT. (And similarly for selling gambles.)

The following two conditions are drawn, with some slight modifications, from Walley (1991).[36]

Accepting Sure Gains There is a price $p \geq \min_{\omega \in \Omega} X(\omega)$ for which the agent is willing to buy X.

Packaging If the agent is willing to buy a gamble X_1 with stakes S for a price of $q_1 S$ and he is willing to buy the gamble X_2 with stakes S for a price of $q_2 S$, then he is willing buy to the gamble $X_1 + X_2$ with stakes S for a price of $\frac{q_1 S_{X_1} + q_2 S_{X_2}}{S_{X_1+X_2}} S$. (And similarly for selling gambles.)

For comparison, here are behavioural characterisations of the confidence, imprecise probability and Bayesian approaches in this framework.

Characterisations A set of betting quotients $\underline{q}_S, \overline{q}_S$:

1. satisfies Accepting Sure Gains, Packaging, Stakes-Independence and Buy-sell coincidence if and only if it satisifies Dutch Book Invulnerability and Buy-sell coincidence if and only if it is derived from a probability measure.

[36] They are related to conditions discussed in the philosophical literature: the former appears to be a weak form of 'Czech Book Invulnerability' as discussed by Hajek (2008); the latter is a version of the Package Principle which has received some attention in the philosophical literature (Schick, 1986; Hajek, 2008).

2. satisfies Accepting Sure Gains, Packaging and Stakes-Independence if and only if it is derived from a set of probability measures. Moreover, in this case, Dutch Book Invulnerability is satisfied.

3. satisfies Accepting Sure Gains, Packaging and Stakes-Dependence if and only if it is derived from a confidence ranking and a cautiousness coefficient. Moreover, in this case, Dutch Book Invulnerability is satisfied.

These characterisations are either just a reminder of known results for precise and imprecise probabilities (in particular Walley (1991, §§3.3.3 & 3.2.2)), or can be simply proved by combining these results with techniques developed in a more refined setup in Hill (2013a, 2016).

For completeness, we sketch the proof of the least well-known characterisation, 3. For every gamble X, let $\underline{P}_S(X) = \underline{q}_S(X)S_X$; $\underline{P}_S$ gives the highest buying price for each gamble, considered "as if" it had stakes S. We show that, for every S, $\underline{P}_S$ is a coherent lower prevision in the sense of Walley (1991, §2.3.3): that is, it satisfies three conditions that he calls accepting sure gains, positive homogeneity and superlinearity. Fix an arbitrary stakes level (positive real number) S. By Accepting Sure Gains, $\underline{P}_S(X) \geq \frac{S_X}{S} \min X^S = \min X$ for all gambles X(accepting sure gains). By definition, $\underline{P}_S(\lambda X) = \underline{q_S}(X)\lambda S_X = \lambda \underline{P}_S(X)$ (positive homogeneity). Finally, Packaging holds if and only if $\underline{P}_S(X+Y) = \underline{q}_S(X+Y)S_{X+Y} \geq \frac{\underline{q_S}(X)S_X + \underline{q_S}(Y)S_Y}{S_{X+Y}} S_{X+Y} = \underline{P}_S(X) + \underline{P}_S(Y)$ (superlinearity). So by Walley (1991, §§3.3.3 & 3.2.2), for each stakes level S, there exists a set of probability measures $\mathcal{C}_S$ such that $\underline{P}_S(X) = \min_{p \in \mathcal{C}_S} E_p(X)$ for all gambles X. Moreover, there is a unique maximal such set for each S; let $\mathcal{C}_S$ be the maximal such set. By the aforementioned properties of these sets, for each probability measure p, $p \in \mathcal{C}_S$ if and only if $E_p(X) \geq \underline{P}_S(X)$ for all gambles X. However, by Stakes-Dependence, for any S, T with $S \leq T$, $\underline{P}_S(X) = \underline{q_S}(X)S_X \geq \underline{q_T}(X)S_X = \underline{P}_T(X)$. Thus for every $p \in \mathcal{C}_S$, $E_p(X) \geq \underline{P}_S(X) \geq \underline{P}_T(X)$, and hence $p \in \mathcal{C}_T$; so $\mathcal{C}_S \subseteq \mathcal{C}_T$. Let $\Xi = \{\mathcal{C}_S : S > 0\}$; this is a nested family of sets of probability measures, and hence a confidence ranking, in the sense of Section 2.1. Let D be the function on gambles assigning the set $\mathcal{C}_{S_X}$ to gamble X. Note that D assigns sets to gambles uniquely on the basis of their stakes, and moreover, for any pair of gambles X, Y, $D(X) \supseteq D(Y)$ whenever $S_X \geq S_Y$; so it is a well-defined cautiousness coefficient. We henceforth use $D(S)$ for the set assigned to gambles of stakes S. By construction, $\underline{q}_S(X) = \min_{p \in D(S)} E_p(\overline{X})$, establishing the existence of the required confidence ranking and cautiousness coefficient.

Conversely, suppose that the specified confidence ranking and cautiousness coefficient exist. For $T \leq S$, since $D(T) \subseteq D(S)$, $\min_{p \in D(S)} E_p(\overline{X}) \leq \min_{p \in D(T)} E_p(\overline{X})$, so Stakes-Dependence holds. Moreover, by Walley (1991, §§3.3.3 & 3.2.2), for any S, $\underline{P}_S$ satisfies superlinearity, whence Packaging holds. Since $\min_{p \in D(S_X)} E_p(\overline{X}) \geq \min_{\omega \in \Omega} X(\omega)$, Accepting Sure Gains holds. Finally, since the confidence ranking is nested, there exists $p \in \bigcap_{\mathcal{C} \in \Xi} \mathcal{C}$. For any such p and any gamble X, $E_p(X) \geq \underline{q_{S_X}}(X)S_X$, whence by Walley (1991, §3.3.3), Dutch Book Invulnerability holds.

2. is just a reformulation of Walley's definition of coherent lower previsions and results concerning them (1991, §§2.3.3 & 3.3.3). This can be seen simply by noting that Stakes-Independence holds if and only if $\underline{P}_S = \underline{P}_T$ for all stakes levels S, T, so Walley's lower prevision $\underline{P}$, defined by $\underline{P}(X) = \underline{P}_{S_X}(X)$ for all gambles X, satisfies his three conditions. The representation by a single set of probability measures follows immediately. 1. is just a reminder of standard results for precise probabilities (recalled in Walley (1991, §§2.3.6 & 2.8)), including the classic Dutch Book Theorem.

References

Abdellaoui, M., Baillon, A., Placido, L., and Wakker, P. P. (2011). The rich domain of uncertainty: Source functions and their experimental implementation. *The American Economic Review*, 101(2):695–723.

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

Armendt, B. (2010). Stakes and beliefs. *Philosophical Studies*, 147(1):71–87.

Arrow, K. J. (1971). *Essays in the Theory of Risk Bearing*. Markham Publishing., Chicago.

Augustin, T., Coolen, F. P., de Cooman, G., and Troffaes, M. C. (2014). *Introduction to imprecise probabilities*. John Wiley & Sons.

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*. Springer, New York.

Bewley, T. F. (2002). Knightian decision theory. Part I. *Decisions in Economics and Finance*, 25(2):79–110.

Bradley, R. (2009). Revising incomplete attitudes. *Synthese*, 171(2):235–256.

Bradley, R. (2015). Ellsberg's Paradox and the Value of Chances. *Economics and Philosophy*, pages 1–18.

Bradley, R. (2016). *Decision theory with a human face*. Cambridge, UK: Cambridge University Press.

Bradley, R. and Steele, K. (2015). Making climate decisions. *Philosophy Compass*, 10(11):799–810.

Bradley, S. and Steele, K. (2013). Uncertainty, Learning, and the "Problem" of Dilation. *Erkenntnis*, pages 1–17.

Bradley, S. and Steele, K. (2016). Can free evidence be bad? Value of information for the imprecise probabilist. *Philosophy of Science*, 83(1):1–28.

Chateauneuf, A. and Faro, J. H. (2009). Ambiguity through confidence functions. *J. Math. Econ.*, 45:535–558.

Cox, L. A. T. (2012). Confronting Deep Uncertainties in Risk Analysis. *Risk Analysis*, 32(10):1607–1629.

Cozic, M. and Hill, B. (2015). Representation theorems and the semantics of decision-theoretic concepts. *Journal of Economic Methodology*, 22(3):292–311.

de Finetti, B. (1937). La Prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré*, 7:1–68.

Dempster, A. P. (1967). Upper and Lower Probabilities Induced by a Multivalued Mapping. *Annals of Mathematical Statistics*, 38:325–339.

Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. *Quart. J. Econ.*, 75(4):643–669.

Ergin, H. and Gul, F. (2009). A theory of subjective compound lotteries. *J. Econ. Theory*, 144(3):899–929.

Fox, C. R. and Tversky, A. (1995). Ambiguity aversion and comparative ignorance. *The Quarterly Journal of Economics*, 110(3):585–603.

Gajdos, T., Hayashi, T., Tallon, J.-M., and Vergnaud, J.-C. (2008). Attitude toward imprecise information. *J. Econ. Theory*, 140(1):27–65.

Gärdenfors, P. (1988). *Knowledge in Flux : Modeling the Dynamics of Epistemic States*. MIT Press, Cambridge, MA.

Gärdenfors, P. and Sahlin, N.-E. (1982). Unreliable probabilities, risk taking, and decision making. *Synthese*, 53(3):361–386.

Ghirardato, P. and Marinacci, M. (2002). Ambiguity Made Precise: A Comparative Foundation. *J. Econ. Theory*, 102(2):251–289.

Gilboa, I. (2009). *Theory of Decision Under Uncertainty*. Econometric Society Monographs. Cambridge University Press.

Gilboa, I., Maccheroni, F., Marinacci, M., and Schmeidler, D. (2010). Objective and Subjective Rationality in a Multiple Prior Model. *Econometrica*, 78(2):755–770.

Gilboa, I. and Marinacci, M. (2013). Ambiguity and the Bayesian Paradigm. In *Advances in Economics and Econometrics: Theory and Applications, Tenth World Congress of the Econometric Society*.

Gilboa, I., Postlewaite, A., and Schmeidler, D. (2009). Is it always rational to satisfy Savage's axioms? *Economics and Philosophy*, 25(3):285–296.

Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *J. Math. Econ.*, 18(2):141–153.

Grove, A. (1988). Two Modelings for Theory Change. *Journal of Philosophical Logic*, 17:157–170.

Hajek, A. (2008). Arguments for-or against-Probabilism? *The British Journal for the Philosophy of Science*, 59(4):793–819.

Hájek, A. (2008). Dutch book arguments. *The handbook of rational and social choice*, pages 173–196.

Halpern, J. Y. (2003). *Reasoning about Uncertainty*. MIT Press, Cambridge.

Hansen, L. P. and Sargent, T. J. (2001). Robust Control and Model Uncertainty. *The American Economic Review*, 91(2):60–66.

Hill, B. (2013a). Confidence and decision. *Games and Economic Behavior*, 82:675–692.

Hill, B. (2013b). Dynamic Consistency and Ambiguity: A Reappraisal.

Hill, B. (2015). Confidence as a Source of Deferral. Technical report, HEC Paris Research Paper No. ECO/SCD-2014-1060.

Hill, B. (2016). Incomplete preferences and confidence. *Journal of Mathematical Economics*, 65:83–103.

Joyce, J. M. (1999). *The Foundations of Causal Decision Theory*. CUP, Cambridge.

Joyce, J. M. (2011a). A Defense of Imprecise Credences in Inference and Decision Making. In Gendler, T. S. and Hawthorne, J., editors, *Oxford Studies in Epistemology*, volume 4. OUP.

Joyce, J. M. (2011b). A Defense of Imprecise Credences in Inference and Decision Making. In Gendler, T. S. and Hawthorne, J., editors, *Oxford Studies in Epistemology*, volume 4. OUP.

Klibanoff, P., Marinacci, M., and Mukerji, S. (2005). A Smooth Model of Decision Making under Ambiguity. *Econometrica*, 73(6):1849–1892.

Lempert, R. J. and Collins, M. T. (2007). Managing the Risk of Uncertain Threshold Responses: Comparison of Robust, Optimum, and Precautionary Approaches. *Risk Analysis*, 27(4):1009–1026.

Levi, I. (1974). On Indeterminate Probabilities. *Journal of Philosophy*, 71:391–418.

Levi, I. (1986). *Hard Choices. Decision making under unresolved conflict*. Cambridge University Press, Cambridge.

Maccheroni, F., Marinacci, M., and Rustichini, A. (2006). Ambiguity Aversion, Robustness, and the Variational Representation of Preferences. *Econometrica*, 74(6):1447–1498.

Machina, M. J. and Siniscalchi, M. (2013). Ambiguity and ambiguity aversion. *Economics of Risk and Uncertainty, Handbook in Economics, North Hollad, Machina and Viscusi eds*.

Marinacci, M. (2015). Model uncertainty. *Journal of the European Economic Association*, 13(6):1022–1100.

Millner, A., Dietz, S., and Heal, G. (2012). Scientific Ambiguity and Climate Policy. *Environmental and Resource Economics*, 55(1):21–46.

Mongin, P. (1995). Consistent Bayesian Aggregation. *Journal of Economic Theory*, 66(2):313–351.

Morgan, M. G. (2014). Use (and abuse) of expert elicitation in support of decision making for public policy. *Proceedings of the National Academy of Sciences*, 111(20):7176–7184.

Nau, R. F. (1992). Indeterminate Probabilities on Finite Sets. *The Annals of Statistics*, 20(4):1737–1767.

Nau, R. F. (2006). Uncertainty Aversion with Second-Order Utilities and Probabilities. *Management Science*, 52:136–145.

Pratt, J. W. (1964). Risk Aversion in the Small and in the Large. *Econometrica*, 32:122–136.

Ramsey, F. P. (1931). Truth and Probability. In *The Foundations of Mathematics and Other Logical Essays*. Harcourt, Brace and Co., New York.

Savage, L. J. (1954). *The Foundations of Statistics*. Dover, New York.

Schick, F. (1986). Dutch bookies and money pumps. *The Journal of Philosophy*, 83(2):112–119.

Schmeidler, D. (1989). Subjective Probability and Expected Utility without Additivity. *Econometrica*, 57(3):571–587.

Segal, U. (1987). The Ellsberg Paradox and Risk Aversion: An Anticipated Utility Approach. *International Economic Review*, 28(1):175–202.

Seidenfeld, T. (1988). Decision Theory Without "Independence" or Without "Ordering". *Economics and Philosophy*, 4:267–290.

Seo, K. (2009). Ambiguity and Second-Order Belief. *Econometrica*, 77(1963):1575–1605.

Shafer, G. (1976). *A Mathematical Theory of Evidence*. Princeton University Press, Princeton.

Siniscalchi, M. (2009). Two out of three ain't bad: a comment on "The Ambiguity Aversion Literature: A Critical Assessment". *Economics and Philosophy*, 25(Special Issue 03):335–356.

Skyrms, B. (1977). Resiliency, Propensities, and Causal Necessity. *The Journal of Philosophy*, 74(11):704–713.

Skyrms, B. (1980). Higher order degrees of belief. In Mellor, D. H., editor, *Prospects for pragmatism. Essays in memory of F. P. Ramsey*, pages 17–25. Cambridge University Press.

Smith, C. (1961). Consistency in Statistical Inference and Decision. *Journal of the Royal Statistical Society. Series B*, 23(1):1–37.

Smithson, M. (2014). Elicitation. In Augustin, T., Coolen, F. P., de Cooman, G., and Troffaes, M. C., editors, *Introduction to imprecise probabilities*, pages 318–328. John Wiley & Sons.

Strzalecki, T. (2011). Axiomatic foundations of multiplier preferences. *Econometrica*, 79(1):47–73.

Tversky, A. and Fox, C. R. (1995). Weighing risk and uncertainty. *Psychological Review*, 102:269–283.

Wakker, P. P. (2010). *Prospect Theory For Risk and Ambiguity*. Cambridge University Press, Cambridge.

Walley, P. (1991). *Statistical reasoning with imprecise probabilities*. Chapman and Hall, London.

Walley, P. (2000). Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning*, 24(2-3):125–148.

Weirich, P. (2001). *Decision space: Multidimensional utility analysis*. Cambridge University Press.

Yaniv, I. and Foster, D. P. (1995). Graininess of judgment under uncertainty: An accuracy-informativeness trade-off. *Journal of Experimental Psychology: General*, 124(4):424.

Yaniv, I. and Foster, D. P. (1997). Precision and accuracy of judgmental estimation. *Journal of behavioral decision making*, 10(1):21–32.