Synthese (2017) 194:3931–3954
DOI 10.1007/s11229-015-0782-5

S.I. : VAGUENESS AND PROBABILITY

Vague Credence

Aidan Lyon[1,2,3,4]

Received: 20 June 2014 / Accepted: 20 May 2015 / Published online: 5 July 2015

Abstract It is natural to think of precise probabilities as being special cases of imprecise probabilities, the special case being when one's lower and upper probabilities are equal. I argue, however, that it is better to think of the two models as representing two different aspects of our credences, which are often (if not always) vague to some degree. I show that by combining the two models into one model, and understanding that model as a model of vague credence, a natural interpretation arises that suggests a hypothesis concerning how we can improve the accuracy of aggregate credences. I present empirical results in support of this hypothesis. I also discuss how this modeling interpretation of imprecise probabilities bears upon a philosophical objection that has been raised against them, the so-called inductive learning problem.

Keywords Vagueness · Imprecise · Indeterminate · Credence · Subjective probability · Degree of belief · Aggregation

✉ Aidan Lyon
aidanlyon@gmail.com; alyon@umd.edu

1 Department of Philosophy, University of Maryland, Skinner Building, College Park, MD 20742, USA

2 Munich Centre for Mathematical Philosophy, Ludwig Maximilian University of Munich, Munich, Germany

3 Centre of Excellence for Biosecurity Risk Analysis, University of Melbourne, Melbourne, Australia

4 Department of Philosophy, The Australian National University, Canberra, Australia

1 Introduction

The use of sets of credence functions—or *credal sets*, as I shall call them—to analyse subjective uncertainty is all the rage these days. The idea is not a new one and has been studied in detail many times in the past.[1] However, the literature has recently seen a new wave of activity on the topic and is now booming with puzzles, paradoxes, new applications to old problems, and interesting new distinctions.[2]

Although their popularity has recently soared, there has been little corresponding discussion of how we should interpret credal sets or how their interpretation should bear upon their applications. To be sure, there have been important discussions of their interpretation in the past (e.g., Levi 1985; Walley 1991, §2.10). However, with the notable exceptions of Christensen (2004) and Rinard (2013) (both of which I will discuss in detail later), the implicit consensus in the recent literature can easily be seen as something of a literal interpretation of credal sets: that our credences (our "degree-of beliefs", "partial beliefs", etc.) have, or ought to have, the set-theoretical structure of credal sets. I doubt that this is the *actual* consensus of the literature; however, the rejection of this interpretation is rarely discussed, and as far as I am aware, its philosophical implications have not been noticed. So, in this paper, I shall argue for an alternative interpretation of credal sets, and discuss some of the implications of this interpretation.

The interpretation, put simply, is that credal sets are imperfect models of our credences and the norms that apply to them. Our actual credences tend to have a structure to them that differs from that of a set of credence functions, and this is mostly because our credences tend to be *vague*. And, importantly, our credences may be vague because they *ought* to be vague. Given this vague nature of our credences, credal sets should be seen as precise, and therefore necessarily imperfect, models of our credences.

I will argue that once we understand credal sets as merely imperfect models of our credences, we can obtain theoretical and practical benefits that would otherwise be easily missed. In particular, I show that by combining the two most-popular models of subjective uncertainty—viz., credence functions and credal sets—into one model, rather than seeing them as competing models, or as one being a generalisation of the other, we can elicit and aggregate credences in ways that lead to more accurate collective credences than we would otherwise obtain. And I will also explain how one particular philosophical objection that has been raised against the credal set model—the alleged impossibility of inductive learning—looks rather different once we understand credal sets as imperfect models.

I will start, in Sect. 2, by clearing the way with a few necessary distinctions and some background on credal sets. In Sect. 3, I will discuss the notion of vague credence and how this leads to the modeling interpretation of credal sets and credence functions. In Sect. 4, I develop another model—the *fuzzy credal set model*—and use it to explain

[1] See e.g., Koopman (1940), Kyburg (1961), Levi (1974), Kaplan (1983), Van Fraassen (1990), Walley (1991), Seidenfeld and Wasserman (1993).

[2] See e.g., Elga (2010), Joyce (2010), White (2010), Bradley and Steele (2012), Hájek and Smithson (2012), Seidenfeld et al. (2012), Rinard (2013), Pedersen and Wheeler (2014), Chandler (2014), Dallmann (2014), Moss (2014), Singer (2014). For a detailed overview of the literature see Bradley (2014).

how we should understand the original two models. Then, in Sect. 5, I discuss how this interpretation of these models suggests an empirical hypothesis concerning the accuracy of collective credences. In Sect. 6, I present evidence in support of that hypothesis. Finally, in Sect. 7, I discuss a particular philosophical consequence of the modeling interpretation of credal sets: its impact on the inductive learning problem for credal sets.

2 Credal sets

What is your credence that it will rain tomorrow? If you are like most people, you will find it difficult to answer that question with a single, precise number. If pressed, you could come up with some number, but it will be hard to justify why you gave that number, rather than some other nearby one. Suppose you said “80 %”. Why not 81 %?, or 79 %?, or 75 %?, or 85 %?, or... and so on. Typically, we find ourselves unable to justify reports of our credences as precise numbers.

A common remedy to this problem is to allow people to express the imprecision surrounding their credences. The usual way of doing this is to allow them to express their credence as a convex set of probabilities, rather than as a single probability.[3] The idea is that instead of forcing you to state a single number as your credence that it will rain tomorrow, you are allowed to express the imprecision associated with your credence in the form of an interval of numbers. For example, instead of being forced to say “80 %” you can say “80 $\pm$ 10 %”, or, equivalently, “[0.7, 0.9]”.

It is common to call these sets of probabilities *imprecise probabilities*. However, that name can be somewhat misleading, and if we use only that name we can easily lose an important distinction that (Levi 1985, 2000) has identified. Levi has argued that we need to be careful not to conflate what he calls *imprecise* probabilities with *indeterminate* probabilities. The distinction can be seen by considering two explanations for why we might be unable to specify a precise number as our credence in some proposition.

The first explanation states that our credences are always precise but we are just not always able to perfectly announce what they are.[4] This might be for a variety of reasons. We may not have perfect access to our all aspects of our mental states, including our credences, and so, because of a failure of introspection, we may be limited to estimating imprecisely what our credences are. Or, alternatively, we might not be perfectly articulate and thus we automatically round off some decimal points when we try to express what our credence in some proposition is. Or, we may simply be pressed for time, and so we might express an imprecise estimate as a compromise between saying nothing and saying something false. For example—*quick!*—answer this question in five seconds: how confident are you that 1000 fair–and–independent coin tosses will land “heads” 425 times? You may know what formula will allow you to calculate this probability precisely, but with only five seconds to answer, you will probably resign to giving an imprecise *estimate* of some sort—e.g., “less than

[3] There has been some debate over whether the sets in question should be convex or not (see e.g., Kyburg and Pittarelli 1996). For the sake of simplicity, I shall be assuming convexity, but nothing hinges upon this.

[4] This is somewhat analogous to epistemicism about vagueness (e.g., Williamson 1994).

1 %”. Depending on how we understand what credences are metaphysically (see Eriksson and Hájek 2007), we might understand such an estimate as an estimate of what your credence *is* (and not merely what it *ought* to be). If, for whatever reason, we are sometimes only able to announce estimates of our credences, and if this is the reason why we sometimes want to announce imprecise probabilities, then the term “imprecise” seems apt. This is analogous to our ability to report the value of some physical quantity, such as the mass of a marble (which we may assume is precise for the sake of argument): due to our imperfect measuring devices, there is always some imprecision in our physical measurements.

The second explanation states that our credences *themselves* are not precise. According to this explanation, the imprecision in someone’s expression of a credence is not necessarily due to their inability to perfectly know their own mental states or what have you; rather, it is due to the imprecise nature of the mental states *themselves*. To put it in picturesque terms, this second explanation says that our mental states are more like collections of clouds than they are like bags of marbles. Compared to a marble, the boundaries of a cloud are rather vague and indeterminate.[5] So if we were to attempt to measure a cloud’s mass, then—even if we had perfect cloud-measuring technology—there would still be some imprecision in that measurement. And so it goes, according to this second explanation, for our credences: our difficulty in finding a single number to express as our credence in a given proposition is due to the fact that there is an indeterminacy in the credence itself and the fact that no single number reflects that indeterminacy.

One important difference between these two explanations is that the first says that the imprecision in question is due to some failure of ours, while a motivation for the second explanation says quite the opposite: that the imprecision could be due to a *perfectly rational* response to one’s evidence. To see how this might be, consider an urn that contains 100 balls—30 of which are black, and the remaining 70 are red or blue in some unknown proportion. Suppose that I randomly select one of the balls from the urn. What is your credence that the ball I drew is black? It seems clear that your credence ought to be 0.3. Now, what is your credence that the ball I drew is blue? It seems much less clear what your credence ought to be. It should not be less than 0, nor should it be greater than 0.7, but apart from that, the evidence you have about the urn doesn’t seem to further constrain your credence. Some authors say you are therefore free to have a credence that is a probability between 0 and 0.7 (e.g., Jeffrey 1987; Elga 2010). Other authors balk at such arbitrariness and say that your credence ought to be *all* of the probabilities that are compatible with your evidence, namely the set [0, 0.7] (see e.g., Levi 1985; Kaplan 1996; Walley 1991; Sturgeon 2008; Joyce 2005, 2010).[6] According to these authors, your credence ought to be imprecise in this

[5] For convenience, I’m speaking as though objects themselves can be vague, but we need not be committed to ontological vagueness. We could instead talk about about vagueness at the level of semantics: e.g., ‘the cloud’ may not determinately refer to a unique set of molecules in the sky.

[6] One might claim that the evidence requires that your credence be a fraction of 100, since the number of blue balls is an *integer* between 0 and 70 and the total number of balls is 100. However, this is not true, because various chance-mixings of hypotheses are compatible with the evidence. For example, the urn’s constitution may have been determined by a fair coin toss that decided whether there would be 30 or 31 blue balls. Conditional on that hypothesis, one’s credence that the ball I drew is blue should be 30.5.

way because that is the only way to rationally respond to the given evidence. And this stems from a more general principle concerning the character of evidence and the character of a credence rationally based upon that evidence, namely: that they must match—call this the *character matching principle* (cf. Sturgeon 2008, p. 159). So, as a special case of this principle, if your evidence for a proposition is imprecise, then your credence in that proposition should also be imprecise.

Although the character matching principle sounds eminently plausible, I'm not aware of any arguments for or against it.[7] I won't attempt to provide any argument for the principle here, and in what follows I will simply assume that the principle is true. This will suffice for my purposes, since something like the principle seems to be endorsed by many authors who prefer to understand subjective uncertainty in terms of credal sets.[8]

Incidentally, it is worth noting that these two explanations (for why we sometimes struggle to announce precise numbers as our credences) are compatible with each other. For example, it may be that *sometimes* we really do have precise credences that we don't have full access to and at *other times* we have indeterminate credences (and perhaps rationally so). It could also be that sometimes we fail to have full access to our indeterminate credences. In those cases, there would be two sources of imprecision that make it difficult for us to express our credences as precise numbers.

For the rest of this paper, I will primarily be focused on cases where the imprecision of one's announced credences are due to the indeterminacy of those credences. Before moving on, there are few other terminological matters that I need to clarify. I shall use 'credence' to refer to the doxastic attitude that one has towards a proposition that often goes by the name of 'degree of belief' or 'partial belief'. One's credence in a proposition is, therefore, understood as part of one's doxastic state, which is in turn part of one's mental state. I will assume, for simplicity, that a complete specification of an agent's credences will be a specification of the agent's doxastic state (i.e., there is no more to an agent's doxastic state than the agent's credences).[9] The term 'credence function' will refer to a function from propositions to real numbers that might represent an agent's credences. The term 'credal set' will refer to a set of credence functions that might represent an agent's credences. If a credal set does represent an agent's credences, and therefore the agent's doxastic state, then the agent's credence in a particular proposition will be represented by a set of numbers—i.e., the set of numbers that the credence functions in the credal set assign to the proposition. For simplicity, I will assume that credal sets are convex,[10] so if a set of numbers represents an agent's credence, it will form an interval, which I'll call an *interval credence*. For example, my credence that a blue ball was drawn from the urn described earlier is the interval

7 Wheeler (2014) has argued against Sturgeon's (2008) use of the principle in connection with full belief and its relation to partial belief; however, those issues are tangential to the ones I am focused on here.

8 There are also several other arguments for why it is rational for one to have indeterminate credences. See Bradley (2014) for an overview of them.

9 Levi (2009) has argued that we also need to also specify what the agent *fully believes*. This detail will not matter for what I have to say in this paper.

10 This assumption has long been debated (see e.g., Kyburg and Pittarelli 1996). Strictly speaking, I do not need it in what follows, but it does greatly simplify the discussion, so I make it only for that reason.

credence [0, 0.7]. If a single number represents a credence, then I will say it is a *precise credence*. Finally, if an agent is perfectly rational, then the agent's credence function will be a probability function—i.e., it will satisfy some standard set of probability axioms (e.g., Kolmogorov 1933, Reńyi 1955)—or the agent's credal set will be a set of probability functions.

The use of credal sets to represent our doxastic states has a number of desirable benefits. One significant benefit is that since the credal set of a rational agent is a set of probability functions, there is a very natural method by which one ought to *update* one's credal set when one acquires new evidence: one simply applies Bayesian conditionalisation to each probability function in one's credal set. Given plausible "real life" assumptions, it can be shown that credal sets of rational agents behave the way we want them to. For example, if we are learning about a coin of unknown bias, we might start off with a very wide interval credence for the outcome "heads"—wide in the sense that the difference between the supremum and infimum of the interval is large. Then, as we conditionalise on the outcomes of repeated tosses of the coin, our interval credence becomes narrower [see e.g., Walley (1991) for a precise formulation of how this can work]. I included the qualification of plausible "real life" assumptions, since if these are not made, then credal sets do not appear to behave the way we want them to when new evidence is acquired. This is the problem of inductive learning, which I'll discuss in detail in Sect. 7.

It is important to note that one's interval credence does not necessarily always become narrower as one acquires more evidence. If a usually-trustworthy friend tells me that the urn (described before) contains 30 to 60 blue balls, then my interval credence may narrow from [0, 0.7] to [0.3, 0.6]. However, if another, even-more-trustworthy friend convinces me that my first friend was lying, then my interval credence will widen back to, or approximately back to, [0, 0.7].[11] Even though I have learned two propositions, one evidentially undermined the other with respect to the issue of the drawn ball being blue, and so I ended up where I started (or at least close to where I started). (There are also cases of *dilation*, where no matter what I learn, my interval widens. See Seidenfeld and Wasserman 1993; Joyce 2010; Pedersen and Wheeler 2014 for discussion).

We can measure two important properties that a rational agent's evidence may have regarding a particular proposition: its *balance* (Keynes 1921, p. 77) and its *specificity* (Joyce 2005, p. 154). The balance of an agent's evidence regarding a proposition is how much the agent's evidence favours that proposition. We can measure the balance of an agent's evidence for a proposition by the *location* of the agent's credence in that proposition on the [0, 1] interval. The closer the agent's precise credence is to 1, or the closer the agent's interval credence is to 1,[12] the more the agent's evidence, on balance, favours the proposition. The specificity of the agent's evidence can increase and decrease as the agent acquire's new evidence, in that the *precision* of agent's

[11] There is an issue here concerning how we can conditionalise on events that are assigned a credence of 0. I'm assuming that these difficulties can be avoided in one way or another—e.g., by taking conditional probabilities as primitive (see e.g., Hájek 2003).

[12] This can be measured in various ways—e.g., by the supremum of the distances between points in the interval and 1.

interval credence can increase and decrease, where precision is measured as $1-(u-l)$, and l is the lower bound of the agent's interval credence and u is the upper bound.[13] The more precise a rational agent's interval credence, the more specific their evidence is concerning the proposition in question.

Although there is much more that can be said about credal sets, we now have enough of the necessary background and basic ideas behind credal sets to be able proceed to the notion of vague credence.

3 Vague credence

Although the previous section followed a fairly standard presentation of credal sets, I hope that something in particular about the presentation struck you as odd. I made the point that our credences are sometimes imprecise or indeterminate and more like clouds than marbles and then I started talking about *sets* of credence functions and interval credences, with their *precise lower and upper bounds*. But just as clouds do not have precise boundaries, neither do many of our credences.

I presented the material in this way because that is how it is often presented. For example, Joyce writes:

> Bayesians are often accused of being committed to the existence of sharp numerical degrees of belief. This is not true. The idea that people have sharp degrees of belief is both psychologically implausible and epistemologically calamitous. Sophisticated versions of Bayesianism, as found in, e.g., (Levi 1980, pp. 85–91); (Kaplan 1996, pp. 27–31), have long recognized that few of our credences are anywhere near definite enough to be precisely quantified. A persons beliefs at a time t are not best represented by any one credence function, but by a *set* of such functions $\mathbf{c}_t$, what we are calling her *credal state* ['credal set' in my terminology]. (Joyce 2005, p. 156).

Similarly, van Fraassen writes:

> Our subjective probabilities are not usually very precise. Rain does seem more likely to me than not, but does it seem more than π times as likely as not? That may not be an answerable question. The standard remedy, as elaborated in the literature by Isaac Levi, Richard Jeffrey, Patrick Suppes, among others, is to represent opinion by a class of probability functions rather than a single one. [...]
>
> To allow for vagueness,[14] a person's state of opinion can thus be represented by a class of probability functions—call it the *representor* ['credal set' in my terminology]. His or her (vague) degrees of belief are meant to be captured by what the members of that class have in common. (Van Fraassen 2006, pp. 483–484).

---

13 Note that the notion of specificity is distinct from Keynes' notion of *weight*, which always increases as one's evidence increases.

14 An anonymous referee has pointed out that van Fraassen uses the term 'vague' to apply to sharply defined intervals. Although I think this is an infelicitous use of the term 'vague', van Fraassen still moves from the point that our credences are often not precise to representing them with credal sets.

There are other motivations for using credal sets (see Bradley 2014, §2), but a common one stems from the point that our credences are often not precise numbers. However, our credences are also often not precise *intervals* either. So, if our credences are often not precise intervals of numbers, why, then, is there a widespread focus on credal sets in the literature?

One possibility is that it is widely thought that our credences are sometimes intervals because they are rationally required to be so—as in the urn example from the previous section—and they are only vague because of our finite, human cognitive limitations. This idea sometimes seems to be lurking in the background in some discussions of credal sets. For example:

> There are a variety of problems with the idea that credences must be precise. As many commentators have observed (see, e.g., Kyburg 1983; Levi 1985; Kaplan 1996), numerically sharp degrees of belief are psychologically unrealistic. It is rare, outside casinos, to find opinions that are anywhere near definite or univocal enough to admit of quantification. [...] The best most people can do, however, is to specify some vague range.
>
> While psychological implausibility is one worry, a more decisive problem is that precise degrees of belief are the wrong response to the sorts of evidence that we typically receive. As argued in Joyce (2005), since the data we receive is often incomplete, imprecise or equivocal, the epistemically right response is often to have opinions that are similarly incomplete, imprecise or equivocal. (Joyce 2010, p. 283).

Joyce then goes on to introduce the standard credal sets model, where the primary motivation seems to be cases in which the evidence in question picks out a precise interval (as in the urn example from Sect. 2).

However, sometimes our evidence mandates that our credences do not correspond to intervals of numbers. This can be seen by considering a modified version of the urn example from the previous section. Suppose I had told you that there are *about* 30 black balls in the urn, with the remaining being either blue or red. What is your credence that the ball I drew is blue? For the same reasons that motivated the idea that one's credence should not be a precise number in the first version of this example, it seems now that one's credence should not be exactly [0, 0.7] (as it was before)—nor should it be any other interval. Rather, one's credence should be *vague*, in order to match the vague character of the evidence (Sturgeon 2008 makes a similar point).

Although there are cases of vague evidence requiring vague credence, it might be thought that these cases are relatively rare. The idea, then, would be that theory of credal sets has a large domain of applicability, and so that is why the intense study of credal sets is warranted. However, if we're counting philosophical thought experiments, there is a vague-urn example for every interval-urn example, so in that respect, cases of vague evidence requiring vague credence are not relatively rare. More importantly, though, it seems that cases of vague evidence requiring vague credence are actually *the norm*. Consider how much of our evidence comes from testimony and how prevalent vague statements are in our natural languages. Evidence from testimony will often require vague credences. For example, suppose that a friend has told you that Barak Obama is *tall* (and suppose you have no other evidence concerning Obama's height), then your credence that his height is between 6′ and 6′1″ ought to be vague and not some interval. Evidence that we acquire on our own might also be vague: e.g., I can notice that today is a *cool* day without knowing the precise temperature, I can notice that an apple on my desk is *red* without knowing its specific kind of redness, and so on. Therefore, it seems to be far more common that our evidence warrants vague credences rather than interval ones. From this perspective, rationally required credal sets seem to largely reside in the realm of philosophical thought experiments about urns with coloured balls and the like.

So, why the widespread focus on credal sets? The most plausible answer seems to be that most philosophers who study credal sets take them as *non-literal* representations of our doxastic states and thus treat them as *imperfect models* of our doxastic states. Fans of credal sets can happily agree that we rarely have, and rarely ought to have, interval credences as our credences while maintaining that credal sets are *better* models of our doxastic states than single credence functions. They are better models in much the same way that super-valuational logic is better than classical logic as an account of, say, how we should reason with vague predicates, and they are imperfect models in much the same way that super-valuational logic is ill-equipped to handle higher-order vagueness. Although my credence that it will rain tomorrow is definitely not the interval [0.7, 0.9], that interval is nevertheless a better model of my credence than the number 0.8 because it captures some of the vagueness of my credence that the single number 0.8 misses.

It is important to stress that the idea here is that credal sets can be better (but imperfect) models of *rationally required* credences, and not merely better models of the credences of us cognitively limited human beings. Hájek has argued that credal sets are better in this second respect, in a response to Pascal's wager:

> Pascals argument is addressed to *us*—mere humans. It is a fact about us that our belief states are irremediably *vague*: we cannot assign probability, precise to indefinitely many decimal places, to all propositions. [...]
>
> As a result, a human agent cannot really be modeled by a single precise probability function at all. A better model is provided by a *set* of such probability functions—what (Van Fraassen 1990), elaborating on proposals by Levi (1980) and Jeffrey (1983), calls a *representor*. (Hájek 2000, pp. 3–4). (Emphasis in original.)

Although credal sets might be more psychologically realistic for us mere humans, they are also more realistic for *what we should aspire to be*. We can be *rationally required* to have vague credences when we have vague evidence, and this rationality requirement can be better modeled by credal sets than by single credence functions.

Although I suspect that many fans of credal sets have a view like this in mind (some say it conversation), it is rarely noted explicitly and much of the discussion in the literature seems to implicitly assume a literal interpretation of credal sets (a point I will come back to in Sect. 7). One notable exception is David Christensen, who, in a footnote, has discussed the same modeling interpretation:

> [I]t seems to me fairly plausible that, if rational attitudes toward propositions may be spread out along ranges of degrees of confidence, those ranges themselves

> will have vague boundaries—there may well be some vagueness in which precise degrees of belief the evidence rules out. But this point is quite compatible with the ranges providing a vastly improved model of spread-out belief. Consider an analogy: We might represent Lake Champlain as stretching from latitude 43:32:15 N to 45:3:24 N. We would realize, of course, that there really aren't non-vague southernmost and northernmost points to the lake; lakes are objects that lack non-vague boundaries. But representing the lake as ranging between these two latitudes is sufficiently accurate, and vastly better than representing the lake's location by picking some single latitude in between them. Similarly, we might represent an ideally rational agent's attitude toward P in a certain evidential situation as ranging from 0.2 to 0.3. We may well do this while realizing that the lowermost and uppermost bounds on degrees of confidence allowed by the evidential situation are vague. But this representation may yet be very accurate, and a considerable improvement over representing the agent's attitude by a single degree of belief. (Christensen 2004, p. 149, n. 4).

In a review of Christensen's book, Patrick Maher has criticised Christensen's claim that an interval representation of one's attitude may be "very accurate", by pointing out there will be situations in which one's credence is so vague and that no interval will be a very accurate representation:

> [C]onsideration of actual cases suggests that [interval representation] isn't very accurate, at least if epistemically rational degrees of belief can be as vague as the corresponding inductive probabilities. For example, the inductive probability that there is intelligent life in the Andromeda galaxy, given my evidence or yours, has a relatively large indeterminacy in its boundaries, and the same is true in many other cases. The boundaries of epistemically rational degrees of belief are not as definite as the boundaries of Lake Champlain. (Maher 2006, p. 145).

Maher is surely correct that there are cases, perhaps like the Andromeda-galaxy one, that involve large levels of indeterminacy. However, this is beside the point. Christensen only claimed that interval credences *may* be very accurate representations of agents' attitudes—not that they *always* are.

Nevertheless, Maher appears to be on to something: credal sets have their limitations. In cases where an agent's evidence is extremely vague or indeterminate, there might not be a credal set that is a very accurate representation of what credence the agent ought to have. It seems to me that the correct conclusion to draw from this consideration is that we should therefore be open to other models that can offer further improvements in representation. However, Maher doesn't draw this conclusion. His preferred model is the standard one (based on individual credence functions) but with an important qualification: a credence function should be understood as an *explication* of an agent's doxastic state (*ibid*, p. 148) and not as a literal representation. Maher's argument for this conclusion is that "use of a single number is simpler, and likely to be more fruitful, than using a set of numbers" (*ibid*).

I think it is safe to interpret Christensen and Maher as agreeing that rational credences can be vague and as disagreeing over which model of vague credences is the better model. Christensen appears to think that credal sets can be considerably more accurate representations of vague credences whereas Maher appears to think that the additional complexity of the credal sets model, in comparison to the credence function model, is not worth any alleged increase in representational accuracy. This appears to be a classic model selection problem. We have two virtues—accuracy and simplicity—that seem to trade off with each other, and the question is which model of vague credence strikes the best compromise in that trade off.

In the next two sections, I will argue that a third, and even-more complicated, model can win this best-model competition—at least for a particular purpose, if not in general. However, before I get to that, it is worth reflecting on the main point of this section. I have argued that our credences are often vague and that, given the character matching principle, they often *ought* to be vague. Much of our evidence comes by way of testimony from others in the form of vague statements, and even evidence that we acquire ourselves may be vague. Our credences in these vague propositions might be precise—e.g., I'm 100 % confident that today is a cool day. But our credences in precise propositions that the vague ones evidentially bear upon will tend to be vague—e.g., I'm *pretty sure* that today's maximum temperature is less than 17 degrees Celsius.

Once we see that many rational credences are vague, I think it becomes clear that we ought to understand credence functions and credal sets as merely *imperfect models* of rational credences, rather than as literal representations of them. This fairly simple point has some interesting philosophical consequences that I'll discuss in Sect. 7. But before I get to those, I need to introduce a third model of vague credence.

4 Fuzzy credal sets

Drawing from the literature on vagueness, a natural model to study in a search for further improvements over the credal set model is a model based on *fuzzy sets* of credence functions. The basic idea is that instead of understanding each credence function as either a member or not a member of one's credal set, we understand each credence function as having a *degree of membership* in one's credal set.[15]

A fuzzy set is a pair (S, m) where S is a set and m is a *membership function* which maps elements s of S to the [0, 1] interval. If $m(s) = 1$, then s has maximal degree of membership in S, and if $m(s) = 0$, then s has no degree of membership in S and we can say that s is not in S. So instead of using a single credence function c or a set of credence functions that meet some constraint K, $C = \{c : c \in K\}$, we use a fuzzy set of credence functions that meets some constraint K, $FC = (\{c : c \in K\}, m)$ to represent one's doxastic state. I will call a fuzzy set of credence functions a *fuzzy*

[15] Gärdenfors and Sahlin (1982) have a very similar model, but they understand the function as a "degree of epistemic reliability". They explain that more reliable credence functions are those that are backed up by more information than other distributions (*ibid*, p. 366). This could be a mere terminological difference, but when the information in question is vague, it seems to me that there can be a difference in the "reliability" in two credence functions, even though they are "backed up" by the same amount of information. More importantly, though, is that Gärdenfors and Sahlin attribute no special importance to those credence functions that have the highest degree of "reliability" (p. 377), they only use their reliability measure to determine (with help from an additional risk aversion attitude) to determine a set of probabilities that are used in their decision procedure (p. 370). As I shall argue, this neglects an important aspect to our credences.

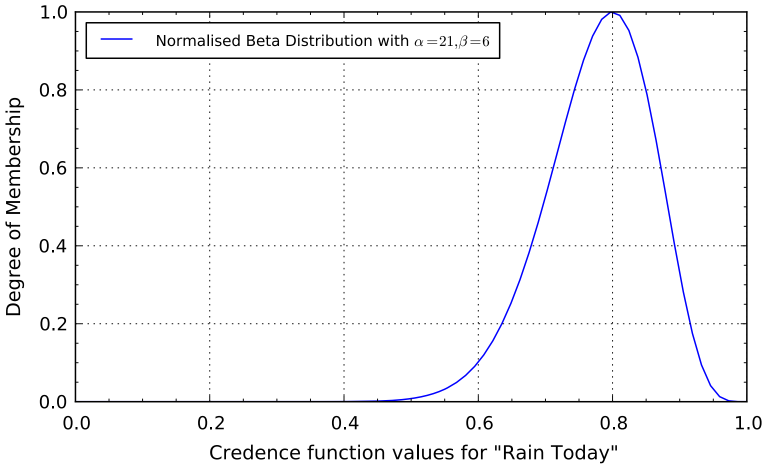

Fig. 1 Example membership function for a fuzzy credence

credal set and I will call the fuzzy set of assignments they give to a proposition a *fuzzy credence*.[16]

We now have three models of our doxastic states: credence functions, credal sets, and fuzzy credal sets. None of these models are perfect. The fuzzy credal set model is clearly not perfect because the membership function m is a precise, real-valued function, and it is implausible that we have anything like that in our heads—to speak loosely—nor is it plausible that we often *ought* to have anything like that in our heads.

However, fuzzy credal sets can better represent the vagueness of one's credences than credal sets can. Just as I can make up a number that roughly represents my credence of rain today—e.g., 0.8—or make up an interval that roughly represents this credence—e.g., [0.7, 0.9]—I can also make up a membership function that roughly represents this credence. It may take more cognitive effort for one to come up with a membership function, but that does not mean that it cannot be done. For example, the membership function that is a beta distribution with $\alpha = 21$ and $\beta = 6$ is a pretty good representation of my credence that it will rain today (see Fig. 1).

Indeed, I think it is a better representation than 0.8 or [0.7, 0.9] in that it respects the fact that my evidence does not sharply discriminate between any two nearby probabilities values—e.g., 0.69 and 0.7—such that one is part of my credence and one is not.[17] Of course, there is a huge degree of arbitrariness in choosing this particular membership function. For example, I can't say that my evidence distinguishes $\beta = 6$ from $\beta = 6.1$. However, it is nevertheless a more accurate representation of my credence than any arbitrary single number or any arbitrary interval.

[16] Note that a fuzzy credal set is not a *fuzzy probability*, which is a probability function defined over an algebra of fuzzy sets (see e.g., Zadeh 1968; Gudder 2000). Note also that this precise notion of a fuzzy credence is different from Sturgeon (2008) notion of *fuzzy confidence*, which has more to do with the notion of vague credence from the previous section.

[17] Since the support of m is [0, 1], there are no other arbitrary cut-off points.

Given that there are three models of vague credence now on the table, it is worth considering how the models relate to each other. It might be tempting to think of the credal set model as a generalisation of the credence function model and the fuzzy credal set model as a generalisation of the credal set model. The idea would be, then, that the credence function model only applies to rational agents whose credences are all precise, and the credal set model only applies to rational agents whose credences are intervals, and so on. But then the models will apply to relatively very few agents and will be of little interest to us, given that we *often* ought to have—and do have—genuinely vague credences.

So it is more useful to think of each model as imperfectly representing our doxastic states. I touched on this idea in the previous section in my discussion of the dispute between Christensen and Maher, but there the idea quickly turned into the issue of which model *better* approximates our doxastic states *in toto*. However, that is not the only way to think of the models. We could, instead, think of them as summarising different *properties* of our doxastic states. To make it easy to see how this idea works, let's suppose—just for the moment—that fuzzy credences *perfectly represent* our doxastic states. (This is surely not true, but it will help fix our ideas.) With that supposition in place, we can think about how credence functions and credal sets may approximate different properties of our credences by thinking about how they may approximate different properties of fuzzy credal sets. For example, a precise credence might provide a good summary of the *location* of a fuzzy credence—e.g., 0.8 is a reasonably good one-number summary of where my fuzzy credence in rain today is located in the [0, 1] interval (from Fig. 1). And an interval credence might provide a good summary of the *spread* of a fuzzy credence—e.g., [0.55,0.95] is a reasonably good interval summary of how my fuzzy credence in rain today is spread across the [0, 1] interval. We can also see why these two properties might be important to summarise, since the location of a rational fuzzy credence is connected to the *balance* of the evidence that supports it, and the spread of a rational fuzzy credence is connected to the *specificity* of the evidence that supports it (cf., Sect. 2).

The concepts of the location and spread of a fuzzy credence may seem vague, but they are no more vague than the corresponding concepts that are employed in plain-old statistics. Quantities such as *means*, *medians*, and *modes* can be what are known as *location parameters*, which measure the "location" of a probability distribution, and then there are quantities such as *variances* that can be what are known as *scale parameters*, which measure the "spread" of a probability distribution. Indeed, since the portion of the membership function of a fuzzy credal set that is restricted to a particular fuzzy credence is *formally speaking* a probability distribution (on the [0, 1] interval), we can lend ourselves the concepts of location and scale parameters from statistics and apply them with equal precision to fuzzy credences.[18]

[18] It is very important to note that this really is only a *formal* equivalence, and the intended interpretation of the membership function of a fuzzy credal set is not that it is some kind of higher-order probability. For example, the degree to which a credence function is a member of an agent's fuzzy credal set is not the agent's uncertainty as to whether that credence function perfectly represents, or ought to perfectly represent the agent's doxastic state. This is, in no way, higher-order probability theory.

One minor point needs to be clarified before we can drop the supposition that fuzzy credences perfectly represent credences. An interval credence not only summarises the spread of a fuzzy credence but it may also usefully summarise its location too. For example, the interval credence [0.55, 0.9] can be used to convey the information that the location of the fuzzy credence is inside that interval, and perhaps also that is close to the interval's midpoint. However, a precise credence provides a more precise summary of the fuzzy credence's location, and in some cases, it may provide a more accurate summary too. For example, suppose we measure a fuzzy credence's location by its mode and its spread by a *95 % membership interval*—i.e., an interval that contains 95 % of the membership function—that contains the mode. Then the interval credence [0.55, 0.9] will be a good summary of the spread of the fuzzy credence in Fig. 1, and if the mode is close to the midpoint of that interval, 0.725, then the interval would also be a good summary of the location. However, by definition, the precise credence of 0.8 would be a more accurate summary of the location (which we have supposed is the mode, which in this case is 0.8). So even though an interval credence may summarise both the spread and location, it may be worth *also* using a precise credence to provide a more accurate summary of the location.

Now let's take a step back and drop the supposition that fuzzy credences perfectly represent vague credences. If the concepts of location and spread can be applied to vague credences, then we can think of interval credences as summarising the spreads and locations of our vague credences and precise credences as sometimes more accurately summarising the locations of our vague credences. Of course, since we are now using precise concepts to summarise properties of vague objects (i.e., vague credences), measuring the accuracy of those summaries will be tricky. Accuracy will either have to be vague to some degree, or we might make it precise in some way—but then we would have to live with it not being entirely appropriate. This is a common problem that arises when we try to use precise terms to analyse vague objects or properties. However, it seems that we can often sensibly use precise terms to reason with vagueness without having a full analysis of their representational accuracy. The location of the peak of Mount Kosciuszko is indeterminate, but I can still give you a reasonable GPS summary of it—"36.455981, 148.263333". And we can understand that GPS summary as being more accurate than the boundary summary of "Kosciuszko National Park", which itself is a more-accurate summary of the indeterminate boundaries of Mount Kosciuszko. If we can do it for mountains, then why not credences?

Anyone who understands credal sets as imperfect models of rational credences owes some story about how some credal sets are better representations of our doxastic states than other credal sets. A fan of understanding credal sets as imperfect models has to be able to say something about why, say, the interval [0, 0.7] is a better representation of the credence that we ought to have in the vague urn example than, say, the interval [0, 0.5] or the point 0.35. We saw in Sect. 3 that Christensen said that interval credences can be considerable improvements over precise credence and that the intervals can be *very accurate*. And we saw Maher criticise Christensen by arguing that interval representations may fail to be very accurate when a credence has a large degree of indeterminacy in its boundaries. If we can make sense of any of this, then we can make sense of interval credences as summarising the spreads and locations of our vague credences and precise credences as sometimes more accurately summarising the locations of our vague credences.

How actual human beings decide to use, say, intervals to approximate their vague credences is an empirical matter that we may be able to study, and there are many possibilities for how they may go about doing this. For example, working in a very similar formal framework, Gärdenfors and Sahlin (1982) suggest that in decision-making contexts, the appropriate intervals will be determined, in part, by the stakes at hand (p. 370). Presumably, the idea is that if people have some degree of ambiguity aversion, then, as the stakes at hand increase, they will tend to reason (perhaps not entirely consciously) about the decision problem with wider intervals, which tend to produce safer decisions. Another possibility is that people can roughly intuit where some significant percentage of their vague credence is located, e.g., 90 %, and when suitably prompted, they announce intervals that approximately capture that amount of their vague credence. Yet another possibility is that people can intuit the inflection points of their vague credences, and so when suitably prompted, they announce intervals with endpoints close to the inflection points (which are surely vague).[19] To work out which, if any, of these ideas is realistic, we need some independent way of measuring an individual's vague credence. One possibility might be to get individuals to select fuzzy credences that feel best to them (perhaps using a software interface, much as I did when I chose the fuzzy credence in Fig. 1), and compare those thus-announced fuzzy credences with the individuals' announced precise and interval credences. I won't try to resolve these empirical and methodological issues here, because I wish to focus on a simpler—and more practical—issue, and it suffices to note that there do appear to be ways for us to get a scientific handle on vague credences and how humans employ them in their reasoning.

5 Aggregating credences

One might be skeptical that human beings can do anything like what I've been suggesting—e.g., use interval credences to summarise their vague credences. It might be thought that we are so far beyond any degree of psychological realism, that questions about fuzzy credences and such are of little interest. However, it turns out that people are willing to provide intervals credence when prompted and it appears that these intervals contain useful information that is not contained in their precise credences.

Drawing from the preceding sections, we can see what that information might be. Recall, from Sect. 2, that the specificity of one's evidence regarding a proposition can be measured by the precision of their interval credence in that proposition (cf. Joyce 2005). Suppose that someone announces an interval credence for P that is narrower than the interval credence they announce for Q. Then that would be a sign that the individual considers their evidence for P to be more specific than the evidence they have for Q.

We can turn this basic idea into one about how we can improve the accuracy of the aggregation of the credences of several individuals. If Ann's evidence concerning

[19] Thanks to Alan Hájek for suggesting this idea.

some proposition is more specific than Bob's, then, if we have no other information about the reliabilities of Ann and Bob, it seems to make sense to give more weight to Ann's credence than to Bob's. If evidential specificity tracks credence accuracy at all, then we should expect that giving more weight to Ann's credence than to Bob's will produce a more accurate collective credence than if we gave their credences equal weight. So, all else being equal, if Ann announces an interval credence for P that is more precise than the one that Bob assigns to P, then by giving Ann's credence more weight, we should obtain a more accurate collective credence than if we gave their credences equal weight.

To be more precise, suppose that we have N individuals, $i = 1, ..., N$, and suppose that we ask each of them for an interval credence in some proposition *and* we also ask them for a precise credence in that proposition. Let's denote their announced interval credence as $[l_i, u_i]$ (for "lower" and "upper") and their announced precise credence as b_i (for "best guess"). Then, from the preceding paragraph, we should expect that an aggregation method that gives more weight to the precise credences of individuals with more precise interval credences, such as:

$$C_1 = \frac{1}{\mathbf{N}} \sum_{i=1}^{N} b_i (1 - (u_i - l_i))$$

will tend to produce collective credences that are more accurate than a method that gives the precise credences equal weight:

$$C_0 = \frac{1}{N} \sum_{i=1}^{N} b_i,$$

where $\mathbf{N} = \sum_{i=1}^{N} 1 - (u_i - l_i)$.

One might wonder why we ought to bother asking people for precise credences in addition to their interval credences. Can't we just simply use the midpoints, m_i, of their intervals? If precise credences do not have any tendency to be more accurate summaries of the locations of people's vague credences, then C_1 should not tend to produce more accurate collective credences than a method that simply uses the midpoints of the intervals:

$$C_2 = \frac{1}{\mathbf{N}} \sum_{i=1}^{N} m_i (1 - (u_i - l_i)).$$

C_1 would tend to be no more accurate than C_2 if it is true that by asking people for precise credences in addition to their intervals, we simply get noisy announcements that are, at best, centered on their interval midpoints. However, if the announced precise credences depart from the interval midpoints by tracking the locations of people's vague credences, and if those locations have some tendency to be accurate, then we should expect C_1 to produce more accurate collective credences than C_2.

In the next section I will present empirical results that support these hypotheses.

6 Empirical results

6.1 Methods

350 participants were recruited using Amazon's Mechanical Turk service and were asked to answer 30 questions about future events. Participants were informed that the Human Intelligence Task (HIT) posted on Mechanical Turk was an invitation to be part of an experiment. They were then redirected to the website http://crowdpredict.it where they could agree to the consent form, get more background info, and answer the 30 questions. Once a participant had finished answering the 30 questions, they were then asked to fill out an optional demographics form and then they were given a completion code. Upon entering the completion code in the original HIT page on the Mechanical Turk site they were paid 1 USD for their participation in the experiment (as initially promised to them). All participants were US residents and had completed at least 100 HITs in the past with at least a 90 % approval rate.

The 30 questions covered a range of topics: from economic events, to geo-political events, to pop album sales, to the 86th Academy Awards. Timeframes for the events ranged from one month after the time of the experiment to one day after the experiment. For each question, e.g., "Will Egypt release or acquit currently-detained Al-Jazeera journalists by March 14, 2014?", participants were given some basic background information about the event and then were asked to specify their lower and upper probabilities and a (single) precise probability, which was elicited as a "best-guess" probability.

The aggregation methods defined in the previous section were evaluated in terms of both the absolute errors and the Brier scores of the answers that they produced. For each question, the answer of the each aggregation method, C_1 and C_2, was compared with the answer produced by C_0 in terms of the percentage improvement. For example, in terms of the Brier Score, Br:

$$\text{Percentage improvement of} \quad C_i \quad \text{over} \quad C_0 = \frac{Br(C_0) - Br(C_i)}{Br(C_0)} \times 100\,\%$$

And then for each aggregation method, its mean percentage improvement over C_0 was used to assess its performance. This was done for both the absolute error and the Brier score, since it is well-known that the results of aggregation methods can be sensitive to which accuracy measures are used to evaluate them (see e.g., Armstrong and Collopy 1992).

Results are reported in effect size with 95 % confidence intervals (95 CIs) and they were examined across demographic breakdowns to see if they were robust.

6.2 Results

C_1 had a mean percentage improvement in absolute error over C_0 of 4.01 %, 95 CI [2.21, 5.82]. C_1 also had a mean percentage improvement in Brier score over C_0 of 7.64 %, 95 CI [4.20, 11.07]. So on both measures of credence accuracy, weighting precise credences by interval credence precisions tended to produce collective credences that were more accurate than those produced by giving the precise credences equal weight.

C_2 had a mean percentage improvement in absolute error over C_0 of 2.59 %, 95 CI [0.54, 4.64]. C_2 also had a mean percentage improvement in Brier score over C_0 of 4.82 %, 95 CI [0.80, 8.84]. So on both measures of credence accuracy, weighting interval credence midpoints by interval credence precisions tended to produce collective credences that were more accurate than those produced by giving the interval credence midpoints equal weight; however the improvement of C_2 was less than that of C_1, as expected. The mean percentage increase in absolute error of C_2 over C_1 was 1.52 %, 95 CI [0.06, 2.99], and the mean percentage increase in Brier score of C_2 over C_1 was 3.22 %, 95 CI [0.22, 6.22].

Qualitatively similar results were obtained for a wide range of demographic breakdowns: men (172) and women (171) were similar to the entire sample; age groups thirty or younger (95), thirty to forty (102), forty to fifty (63), fifty to sixty (57), and sixty or older (33) were all similar to the entire sample; and those who identified as white (279) and those who identified as non-white (71) were also similar to the entire sample.[20] Moreover, similar results were also obtained from the top 10 % most-accurate individuals (35).

6.3 Discussion

Assuming the results from the previous section are representative of people in general, they show that there is value in eliciting interval credences, but they also show that we should not the abandon precise credence model in favour of the interval credence model. By eliciting both kinds of probabilities from individuals, we can obtain information that can be used to improve the accuracy of collective credences. Although, as Maher pointed out (Sect. 3), use of a single number is simpler than use of a set of numbers, it does not seem to be more fruitful. This not to say that the use of interval credences is more fruitful than the use of precise credences. Rather, both are fruitful in different ways, and what is most fruitful, is the use of *both* models. The superior performance of C_1, which combines both models, over C_0 which only uses the precise credence model, and C_2, which only uses the interval credence model, shows that it can be advantageous to use a combination of the models.

Although the participants probably had at least some basic familiarity with probability theory, there is no reason to expect that the participants had any prior familiarity with credence intervals. It is perhaps somewhat surprising that the results obtained for these participants. It would be interesting to see what happens with participants who are expected to be more familiar with credence intervals and with thinking about the specificity of their evidence.

The results also support the claim that our credences have more structure to them than that which either of the precise credence and interval credence models represent by themselves and that this additional structure is tracking the truth values of propositions.

[20] The sizes of the demographic breakdowns do not sum to 350, the total sample size, because some participants did not respond to all of the demographic questions, which were optional.

If we interpret the models in the way that I suggested in Sect. 4, this suggests that the spreads of our vague credences track the accuracies of the locations of our vague credences (a smaller spread indicating a more accurate location). The next step, then, is to test these ideas more directly, by eliciting fuzzy credences from individuals. Once we do this, we can begin to get a more complete view of how complicated people's credences are: Do they sometimes have multiple peaks? Are they sometimes discontinuous? And so on.

I have reported the results of one experiment, and it goes without saying that it should be repeated. Although there is evidence that suggests that participant pools recruited from Amazon's Mechanical Turk are representative of the larger population (see e.g., Paolacci et al. 2010; Berinsky et al. 2012), it is worth examining whether similar results can be obtained from different kinds of people— for example, experts in some particular field—and for different kinds of questions—for example, questions about the participant's field of expertise.

Putting these empirical questions to the side, I wish to turn to some philosophical consequences of the previous sections.

7 The inductive learning problem for credal sets

Let's assume that our credences tend to be vague, and rationally so, and that we ought to think of the three models of credences—credence functions, credal sets, and fuzzy credal sets—along the lines that I have described. What philosophical consequences follow from this modeling view? I think there are many. Such a view will surely affect our decision theory, how we evaluate the quality of our credences, and how we generalise various classical probabilistic concepts such as independence and calibration. However, I will close by discussing one consequence in particular: how this modeling view of credence impacts the so-called *inductive learning problem* for credal sets.

Many proponents of credal sets claim that if an agent has no evidence concerning the truth of some hypothesis H_1, then the agent's interval credence in H_1 should be maximally imprecise, [0, 1] (see e.g., Kaplan 1996; Joyce 2005). However, this leads to the so-called inductive learning problem. Let's suppose that H_1 is the hypothesis that all of the marbles in some particular urn are green, and suppose that you get some evidence E that a marble randomly selected from the urn was green. The inductive learning problem then arises from the following facts, which Rinard (2013) describes:

> [A]fter conditionalizing each function in your [credal set] on E, you have exactly the same spread of values for H_1 that you did before learning E—namely, [0, 1]. Moreover—and this is what is most problematic—your credence for H_1 will remain at [0, 1] no matter how many marbles are sampled from the urn and found to be green. (Rinard 2013, pp. 4–5).

In general, the problem is that if you start from total ignorance regarding some proposition, then no matter how much you subsequently learn, your credence in that proposition will forever remain stuck at [0, 1].

Rinard goes on to consider possible escape routes for the proponents of credal sets. For example, one might reply that one should, or could, start at some interval $[a, b]$

that is properly contained in [0, 1]. Such a reply could[21] solve the learning problem, but as Rinard points out, starting at any such interval will seem arbitrary. Indeed, this escape route will suffer from the kind of arbitrariness that motivated the credal set view in the first place (see Sect. 2).

One of the other possible escape routes that Rinard considers is the fuzzy credal set model that I developed in Sect. 4, although in her hands it looks a little different:[22]

> One alternative approach to color vagueness involves many-valued logics, on which each shade is assigned a precise numerical degree of redness. The analogous approach would be to represent a doxastic state with a fuzzy set of probability functions. Different functions are assigned different precise degrees of membership in the set. Unfortunately, the problem remains. What is the first shade with a degree of redness less than 1? There is no good answer. What is the lowest number r such that $Pr(H_1) = r$ according to some function whose degree of membership in the set is greater than 0? Again, there is no good answer. (Rinard 2013, p. 7).

The general point of the objection is that having any fuzzy credal set as one's credence that allows one to learn will re-introduce the arbitrariness that was to be avoided in the first place. The specific point of the objection, I think, is that for an agent with a fuzzy credal set to be able to learn, there must be a credence function in the fuzzy credal set such that:

- it assigns the lowest credence to H_1 out of all the credence functions in the fuzzy credal set, and
- the credence it assigns to H_1 is greater than 0.

Call this credence function Cr_{lowest} and suppose that the credence it assigns to H_1 is $Cr_{\text{lowest}}(H_1) = k$. If the agent starts in that kind of fuzzy credal state, then the question naturally arises: why didn't the agent start in a state with a lower positive value for $Cr_{\text{lowest}}(H_1)$, say, $k/2$? Any choice of $k > 0$ seems arbitrary. It seems that the only way to avoid the arbitrariness is for k to be 0, and so, the objection goes, fuzzy credal sets don't solve the learning problem either.

The above constraints for learning about H_1 on the fuzzy credence model seem to be derived from the idea that an agent's credence in H_1 is the set of points that have some positive degree of membership in the fuzzy credence. If not, then it's not clear why one would insist that $Cr_{\text{lowest}}(H_1) > 0$. However, according to the fuzzy credence model, literally interpreted, the agent's credence is not the set of points with positive membership; instead, it is the fuzzy credal set itself, which has the membership function as a property.

So let's consider what happens if we adopt an account of learning that is more appropriate for the fuzzy credence model. To simplify things a bit, let us again give a literal interpretation to the fuzzy credence model, and for an example, consider again

[21] Some additional assumptions need to be made—in particular, that the credal set contains no anti-inductive priors.

[22] I have replaced Rinard's M-Green hypothesis with her H_1 hypothesis from earlier. This does not affect the philosophical issues.

my fuzzy credence in the proposition that it will rain, R, from Fig. 1, which contains every credence point in the [0, 1] interval. Operating in the background—normatively, of course—is a fuzzy credal set with its credence functions. Suppose that I acquire some precise evidence E in favour of R. That means I need to conditionalise each credence function on E, and this will determine a new fuzzy credence in R. Let's suppose that my fuzzy credal set was such that for each prior credence point in the fuzzy credence of R, with the exception of the credence point 0, the posterior credence point resulting from the conditionalisation on E is closer to 1 than the prior credence point. Figuratively speaking, then, what happens upon learning E is that the curve in Fig. 1 gets shifted to the right (some points move more than others, so its shape will change too). As more pieces of evidence like E are acquired, the location of the fuzzy credence will move closer to 1 and the spread will shrink. Note that this will happen, even though there will forever be some credence points in the fuzzy set that assign a probability of 0 to R. The support of the membership function will forever be the entire [0, 1] interval, but it is clear that it would odd, at best, to interpret this as no learning happening. So we can say that learning happens on the fuzzy credence model if the location or spread of a fuzzy credence changes. (There may be other ways for learning to happen, but this will suffice for now).

Note that the learning problem has not yet been solved, because the learning problem is for an agent who starts with *total ignorance* regarding some proposition, and the fuzzy credence in Fig. 1 is surely not one of an agent with total ignorance regarding R. So we need to know what fuzzy credence a totally ignorant agent should start with. Whatever it is, it needs to be a fuzzy credence with a location and/or spread that can change when the agent conditionalises on new evidence. However, this constraint will rule out the obvious candidates: fuzzy credences with uniform membership functions over the [0, 1] interval. The location and spread of any such fuzzy credence will not change under any conditionalisations. And, importantly, any fuzzy credence that does satisfy this constraint will have some arbitrariness to it.

So, what can be done? The first thing to be done is to drop the literal interpretations of these models, which is something Rinard goes on to do (*ibid*, p. 8). An agent's credence need not always be some interval credence, nor need it always be some fuzzy interval credence. Again, these are just models for credences, which can be vague and indeterminate.[23] And what happens in cases of *total* ignorance is that one's evidence is *maximally* indeterminate—since one has *no* evidence. Even fuzzy credal sets cannot handle that level of indeterminacy. Our models will always have some kind of precision that some evidential states will lack, so we need to be aware of the limitations of our models. This lesson is also important for cases that are less extreme than total-ignorance cases. For example, recall Maher's point (from Sect. 3) that the credal set model seems to deliver a very inaccurate representation of what our credence ought to be in his example involving intelligent life in the Andromeda galaxy. It is even more inaccurate in cases where we have absolutely no evidence concerning the proposition in question.

[23] Recall, that this point applies to ideally rational agents—one's credence might be vague precisely because one is responding appropriately to one's vague evidence.

This still doesn't answer the question: what credence should a totally ignorant agent have? However, an answer is starting to suggest itself. If we follow the original thought that motivated credal sets in the first place to its limit; that is, if we follow the thought that an agent's credence should match the character of the agent's evidence, then, if the agent has *no* evidence for a proposition, the agent should have *no* credence in that proposition. No wonder our various models break down in such cases of total ignorance: we were trying to use them to come up with the best precise representations of credences that are rationally required to be non-existent. It is only once the agent has some evidence of some form that the agent's credence should also take some form, and only then can our models begin to represent that form. To put it another way, if the character matching principle is true, then there is no inductive learning problem for credal sets. This is because credal sets can only be an appropriate model once the credences of the agent in question have taken some kind of form, and the agent's credences should only take some kind of form once the agent's evidence has taken some kind of form, and that can only happen once the agent has *some* evidence, which contradicts the learning problem's supposition that the agent is totally ignorant and thus has *no* evidence.

When seen from the perspective of the modeling interpretation of credal sets (and credence functions and fuzzy credal sets) and the character matching principle, the learning problem really becomes a problem of figuring out how the character of an agent's credence in a proposition is to match the character of the agent's evidence for that proposition. I'm not sure what we can say about that, or even if there is anything complete and coherent that can be said (cf. Milgram's 2009, §6.4 matching of 'sort of' qualifiers). As I mentioned in Sect. 2, I've assumed the character matching principle without argument, and without further clarification. So there is work to be done in clarifying the character matching principle before we can solve the genuine learning problem. However, there is at least one pitfall that we should avoid when thinking about this problem: we should avoid falling for literal interpretations of our models.

8 Conclusion

It is appears to be common wisdom among fans of credal sets that the use of credal sets to represent our doxastic states is superior to the use of single credence functions (see e.g., Kaplan 1996). And some authors think the opposite: that single credence functions are better than credal sets (e.g., Maher 2006; White 2010). I have argued that neither is better than the other, and that we should use both credence functions and credal sets to model uncertainty—at least in some cases.

The main reason why is that these two representations are merely imperfect *models* of our (rationally required) doxastic states. Neither is a literal representation of our doxastic states, since our doxastic states tend to be far more vague—and yet structured—than either model suggests. I have argued that instead of favouring one model over another, we are better off employing them both. We should think of credence functions as summarising the locations of our vague doxastic states in the [0, 1] scale and credal sets as summarising the spreads and locations of our vague doxastic states in the [0, 1] scale. By combining both models, we can obtain fruits that neither model by itself can provide. For example, I have shown that we can improve the accuracy of collective credence judgements by combining the two models.

Instead of thinking that the models represent different kinds of doxastic states or that they offer competing theories of what our doxastic states are like (or should be like), we should think of the two models as complementary and merely picking out different aspects of the one kind of doxastic state that we tend to have. Sometimes our credences might be sharp numbers—e.g., if you come to know the chance of some event, and you have no other relevant evidence, then it seems your credence should be a single, sharp number. Sometimes our credences might be sharp intervals—e.g., if you come to know the chance of some event is in some interval or that the chance *is* some interval (e.g., Dardashti et al. 2014), and you have no other relevant evidence, then it seems your credence should also be that interval. However, more often than not, our evidence is vague, and so our credences ought to be vague, and so more often than not, credences functions and credal sets are imperfect but complementary models of our credences.

Finally, I discussed what happens in extreme cases where our credences are so vague and indeterminate that all of the precise models that I have discussed—credence functions, credal sets, and fuzzy credal sets—simply break down. These include cases of total ignorance, where an agent has no evidence whatsoever. The main conclusion was that once we understand the models as *models*, then the way they break down in these extreme cases doesn't seem so bad, and even makes some sense.

Acknowledgements The author would like to thank Seamus Bradley, Mark Burgman, Rachael Briggs, Alan Hájek, Michael Morreau, Daniel Nolan, and Reuben Stern for helpful discussion and feedback. This work was financially supported by the Alexander von Humboldt Foundation.

References

Armstrong, S., & Collopy, F. (1992). Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting*, *8*(1), 69–80.

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's mechanical turk. *Political Analysis*, *20*(3), 351–368.

Bradley, S. (2014). Imprecise probabilities. *Stanford Encyclopedia of Philosophy*. Stanford: Stanford University.

Bradley, S., & Steele, K. (2012). Uncertainty, learning, and the "problem" of dilation. *Erkenntnis*, *79*(6), 1287–1303.

Chandler, J. (2014). Subjective probabilities need not be sharp. *Erkenntnis*, *79*(6), 1273–1286.

Christensen, D. (2004). *Putting logic in its place*. Oxford: Oxford University Press.

Dallmann, J. M. (2014). A normatively adequate credal reductivism. *Synthese*, *191*(10), 2301–2313.

Dardashti, R., Glynn, L., Thébault, K., & Frisch, M. (2014). Unsharp humean chances in statistical physics: A reply to beisbart. In M. C. Galavotti et al. (Eds.), *New directions in the philosophy of science* (pp. 531–542). Dordrecht: Springer.

Elga, A. (2010). Subjective probabilities should be sharp. *Philosophers' Imprint*, *10*(5), 1–11.

Eriksson, L., & Hájek, A. (2007). What Are Degrees of Belief? *Studia Logica*, *86*(2), 183–213.

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

Gudder, S. (2000). What is fuzzy probability theory? *Foundations of Physics*, *30*(10), 1663–1678.

Hájek, A. (2000). Objecting vaguely to Pascal's Wager. *Philosophical Studies*, *98*(1), 1–14.

Hájek, A. (2003). What conditional probability could not be. *Synthese*, *137*(3), 273–323.

Hájek, A., & Smithson, M. (2012). Rationality and indeterminate probabilities. *Synthese*, *187*(1), 33–48.

Jeffrey, R. (1983). Bayesianism with a human face. *Testing Scientific Theories, Minnesota Studies in the Philosophy of Science*, *10*, 133–156.

Jeffrey, R. (1987). Indefinite probability judgment: A reply to Levi. *Philosophy of Science*, *54*, 586–591.

Joyce, J. M. (2005). How probabilities reflect evidence. *Philosophical Perspectives*, *19*(1), 153–178.

Joyce, J. M. (2010). A defense of imprecise credences in inference and decision making. *Philosophical Perspectives*, *24*(1), 281–323.

Kaplan, M. (1983). Decision theory as philosophy. *Philosophy of Science*, *50*(4), 549–577.

Kaplan, M. (1996). *Decision theory as philosophy*. Cambridge: Cambridge University Press.

Keynes, J. M. (1921). *A treatise on probability*. London: Macmillan.

Koopman, B. O. (1940). The bases of probability. *Bulletin of the American Mathematical Society*, *46*(10), 763–774.

Kyburg, H. (1983). *Epistemology and Inference*. Minneapolis: University of Minnesota Press.

Kyburg, H. E, Jr. (1961). *Probability and the logic of rational belief*. Middletown: Wesleyan University Press.

Kyburg, H. E, Jr., & Pittarelli, M. (1996). Set-based bayesianism. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, *26*(3), 324–339.

Levi, I. (1974). On indeterminate probabilities. *Journal of Philosophy*, *71*(13), 391–418.

Levi, I. (1980). *The enterprise of knowledge: an essay on knowledge, credal probability, and chance*. Cambridge, MA: MIT Press.

Levi, I. (1985). Imprecision and indeterminacy in probability judgment. *Philosophy of Science*, *52*(3), 390–409.

Levi, I. (2000). Imprecise and indeterminate probabilities. *Risk, Decision and Policy*, *5*(2), 111–122.

Levi, I. (2009). Why indeterminate probability is rational. *Journal of Applied Logic*, *7*(4), 364–376.

Maher, P. (2006). Book review: David Christensen. Putting logic in its place: Formal constraints on rational belief. *Notre Dame Journal of Formal Logic*, *47*(1), 133–149.

Milgram, E. (2009). *Hard truths*. Malden: Wiley.

Moss, S. (2014). Credal dilemmas. Noûs. doi:10.1111/nous.12073.

Paolacci, G., Chandler, J., & Ipeirotis, P. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, *5*(5), 411–419.

Pedersen, P., & Wheeler, G. (2014). Demystifying dilation. *Erkenntnis*, *79*(6), 1305–1342.

Rinard, S. (2013). Against radical credal imprecision. *Thought: A Journal of Philosophy*, *2*(2), 157–165.

Seidenfeld, T., Schervish, M. J., & Kadane, J. B. (2012). Forecasting with imprecise probabilities. *International Journal of Approximate Reasoning*, *53*(8), 1248–1261.

Seidenfeld, T., & Wasserman, L. (1993). Dilation for sets of probabilities. *The Annals of Statistics*, *21*(3), 1139–1154.

Singer, D. J. (2014). Sleeping beauty should be imprecise. *Synthese*, *191*(14), 3159–3172.

Sturgeon, S. (2008). Reason and the grain of belief. *Noûs*, *42*(1), 139–165.

Van Fraassen, B. C. (1990). Figures in a probability landscape, chapter 21. In *Truth or consequences* (pp. 345–356). Dordrecht: Kluwer.

Van Fraassen, B. C. (2006). Vague expectation value loss. *Philosophical Studies*, *127*(3), 483–491.

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

Wheeler, G. (2014). Character matching and the locke pocket of belief. In Context Epistemology (Ed.), *Epistemology, context, and formalism* (pp. 187–195). cham: Springer.

White, R. (2010). Evidential symmetry and mushy credence. In J. Hawthorne (Eds.), *Oxford studies in epistemology*. Oxford: Oxford University Press..

Williamson, T. (1994). *Vagueness*. London: Routledge.

Zadeh, L. A. (1968). Probability measures of fuzzy events. *Journal of mathematical analysis and applications*, *23*(2), 421–427.