

Beliefs, buses and lotteries: Why rational belief can't be stably high credence

Julia Staffel¹

Published online: 25 September 2015
© Springer Science+Business Media Dordrecht 2015

Abstract Until recently, it seemed like no theory about the relationship between rational credence and rational outright belief could reconcile three independently plausible assumptions: that our beliefs should be logically consistent, that our degrees of belief should be probabilistic, and that a rational agent believes something just in case she is sufficiently confident in it. Recently a new formal framework has been proposed that can accommodate these three assumptions, which is known as “the stability theory of belief” or “high probability cores.” In this paper, I examine whether the stability theory of belief can meet two further constraints that have been proposed in the literature: that it is irrational to outright believe lottery propositions, and that it is irrational to hold outright beliefs based on purely statistical evidence. I argue that these two further constraints create a dilemma for a proponent of the stability theory: she must either deny that her theory is meant to give an account of the common epistemic notion of outright belief, or supplement the theory with further constraints on rational belief that render the stability theory explanatorily idle. This result sheds light on the general prospects for a purely formal theory of the relationship between rational credence and belief, i.e. a theory that does not take into account belief content. I argue that it is doubtful that any such theory could properly account for these two constraints, and hence play an important role in characterizing our common epistemic notion of outright belief.

Keywords Credence · Belief · Rationality · Stability theory · Formal epistemology

✉ Julia Staffel
staffel.julia@gmail.com

¹ Department of Philosophy, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130-4899, USA

1 Introduction

Current debates in epistemology employ two different notions of belief: the notion of outright belief and the notion of credence, or degree of belief. Both notions are taken to describe real psychological attitudes, and each notion comes with its own conditions of rationality. If we have both kinds of beliefs, it seems natural to think that there are constraints on which combinations of degrees of belief and outright beliefs an agent can rationally hold. Plausibly, an agent cannot rationally believe p unless she is very confident that p is true. Unfortunately, it has turned out to be very difficult to turn this suggestion into a robust theory of the relation between degrees of belief and outright belief in rational agents. Until recently, it seemed like no theory about the relationship between rational credence and rational outright belief could reconcile three independently plausible assumptions: that our beliefs should be logically consistent, that our degrees of belief should be probabilistic, and that a rational agent believes something just in case she is sufficiently confident in it.

Recently a new formal framework has been proposed that can accommodate these three assumptions. One version of the view is presented as “the stability theory of belief” by Hannes Leitgeb, and a slightly different version is proposed under the name “high probability cores” by Horacio Arló-Costa and Arthur Paul Pedersen.¹ I will focus on Leitgeb’s most recent formulation of the view, and note when other versions of the view are different in ways relevant to the discussion.

In what follows, I examine whether the stability theory of belief has the resources to meet two further constraints on rational outright belief that have been proposed in the literature: that it is irrational to outright believe lottery propositions, and that it is irrational to hold outright beliefs based on purely statistical evidence. I argue that these two further constraints create a dilemma for a proponent of the stability theory: she must either deny that her theory is meant to give an account of the common epistemic notion of outright belief, or supplement the theory with further constraints on rational belief that render the stability theory explanatorily idle. This result sheds light on the general prospects for a purely formal theory of the relationship between rational credence and belief, i.e. a theory that does not take into account belief content. I argue that it is doubtful that any such theory could properly account for these two constraints, and hence play an important role in characterizing our common epistemic notion of outright belief.

¹ Leitgeb’s presentation of the view differs from Arló-Costa and Pedersen’s version in a number of ways, some of which will become important later in the paper. Leitgeb starts from the standard Kolmogorov axiomatization of probability, whereas Arló-Costa and Pedersen prefer a version of the view that is based on primitive conditional probabilities. Also, Leitgeb (2014) highlights the importance of the fineness with which the space of possibilities is partitioned, which isn’t emphasized in his earlier paper, or by Arló-Costa and Pedersen. The authors also differ on the way the threshold for rational belief is set in a context. Arló-Costa and Pedersen, following Leitgeb (2013), claim that the threshold is set by the strongest available proposition (2012, Definition 3.11), whereas Leitgeb (2014, 2015) leaves this open.

2 The stability theory of belief

Leitgeb's (2013, 2014, 2015) stability theory of belief tells us which combinations of credences and outright beliefs agents can rationally hold. Leitgeb is motivated by the challenge to reconcile assumptions (1)–(3):

1. A rational agent's outright beliefs are logically consistent and closed under conjunction.
2. Rational credences must obey the probability axioms.
3. The Lockean Thesis: a rational agent believes a claim X just in case the agent's degree of belief in X is equal to or greater than some threshold r , where r is above $1/2$ and at most 1.

The problem with accepting (1)–(3) simultaneously is this: if we choose a fixed threshold for belief that is less than a credence of 1, then there are sets of inconsistent propositions such that an agent has high enough credence to outright believe each proposition in the set. However, this is incompatible with assumption 1, that an agent's beliefs should be logically consistent. But if we choose credence 1 as the threshold, then we must be certain of everything we believe, which seems too restrictive.

Leitgeb develops a view that reconciles (1) and (2) with a modified version of (3). He shows that we can have thresholds below 1 without violating belief closure if we adopt the following view: for a belief in A to be rational, the agent's credence in A must both be above some threshold that is greater than $1/2$, and remain above the threshold when the agent gains evidence that is consistent with A. Precising the Lockean thesis by adding this second component is the key to modifying (3). To state this more formally, we will first define what it means for a proposition A (construed as a set of possible worlds) to be P-stable²:

P-stability Let P be a probability measure on the sample space of worlds W, and $A \subseteq W$: For all A, A is P-stable iff for all $B \subseteq W$, such that B is consistent with A, and $P(B) > 0$: $P(A|B) > 1/2$.

In other words, a proposition is P-stable iff the agent's credence in it is greater than $1/2$ given any proposition consistent with it. Leitgeb goes on to prove that statements (I) and (II) are equivalent:

Let W be a finite, non-empty set of worlds, let Bel (the agent's belief function) be a set of subsets of W, and let P assign to each subset of W a number in the interval $[0,1]$. Let B_w be the conjunction of all of the agent's beliefs (i.e. the strongest proposition she believes, and the least subset of W she believes).

- (I) Bel satisfies (1), P satisfies (2), P and Bel satisfy: for all B, $\text{Bel}(B)$ iff $P(B) \geq P(B_w) > 1/2$.
- (II) P satisfies (2), and there is a (uniquely determined) $A \subseteq W$, such that

² I am omitting some of the precise formal details of Leitgeb's theory in order to convey the essence of the view to the reader. For a fully precise statement of the view, readers are invited to consult Leitgeb's work directly. The same holds for Arló-Costa and Pedersen's view.

- A is a non-empty P-stable proposition,
- if $P(A) = 1$, then A is the least subset of W with probability 1; and
- for all $B \subseteq W$: $\text{Bel}(B)$ iff $A \subseteq B$, and hence $B_w = A$.

This means that in order for our three assumptions about belief and credence to be satisfied, it has to be the case that B_w , the strongest proposition the agent believes, is P-stable. This sets constraints for the threshold for belief: given that there is a strongest believed proposition B_w , the threshold for rational belief cannot be higher than $P(B_w)$. It also can't be (much) lower than $P(B_w)$, because this might allow for beliefs that aren't supersets of B_w , which conflicts with (1), the closure condition.

In many cases, there is more than one P-stable proposition, each of which is a candidate for being the strongest proposition the agent believes. In other words, just knowing which propositions are P-stable doesn't yet tell us which propositions are believed. To see how this works, let's consider an example. Start with an algebra consisting of six possible worlds, w_1 – w_6 . The agent assigns them the following credences:

$$P(\{w_1\}) = 0.5, P(\{w_2\}) = 0.3, P(\{w_3\}) = 0.11, P(\{w_4\}) = 0.05, \\ P(\{w_5\}) = 0.03, P(\{w_6\}) = 0.01$$

In which propositions does the agent have P-stable credence?³ Here, the logically strongest proposition that meets the stability condition is $\{w_1, w_2\}$, which would fix the threshold r for rational belief at $P(\{w_1, w_2\}) = 0.8$. If we choose this threshold, it means the agent rationally believes $\{w_1, w_2\}$ and anything entailed by it, which is going to be any proposition with a probability above 0.8. But there are also weaker propositions that are P-stable, namely $\{w_1, w_2, w_3\}$, $\{w_1, w_2, w_3, w_4\}$, $\{w_1, w_2, w_3, w_4, w_5\}$, and $\{w_1, w_2, w_3, w_4, w_5, w_6\}$. Each of these propositions is a candidate for the strongest claim believed by the agent. For example, if $\{w_1, w_2, w_3\}$ were the strongest claim believed by the agent, this would set the threshold for rational belief at 0.91. In other words, if there are multiple propositions in which the agent has stably high credence, each of them is a candidate for being the strongest thing believed by the agent, and each of them would fix a threshold for rational belief.

In order to determine which of the different candidate thresholds for belief is acceptable in a given context, further aspects of the context must be consulted. Leitgeb (2014) proposes that we model this kind of context-sensitivity in the same way subject-sensitive invariantism (SSI) accounts for the context-sensitivity of knowledge. This means that whether a high or low threshold for belief is selected in

³ Footnote 5 in Leitgeb (2014) gives a sketch of the algorithm by which one can determine which propositions have stably high credence in a given context: "Assume that $W = \{w_1, \dots, w_n\}$, and $P(\{w_1\}) \geq P(\{w_2\}) \geq \dots \geq P(\{w_n\})$. If $P(\{w_1\}) > P(\{w_2\}) + \dots + P(\{w_n\})$, then $\{w_1\}$ is the first, and least, non-empty P-stable set, and one moves on to the list $P(\{w_2\}), \dots, P(\{w_n\})$; e.g. if $P(\{w_2\}) > P(\{w_3\}) + \dots + P(\{w_n\})$, then $\{w_1, w_2\}$ would be the next P-stable set. On the other hand, if $P(\{w_1\}) \leq P(\{w_2\}) + \dots + P(\{w_n\})$ then consider $P(\{w_2\})$: if it is greater than $P(\{w_3\}) + \dots + P(\{w_n\})$ then $\{w_1, w_2\}$ is the first P-stable set, and one moves on to the list $P(\{w_3\}), \dots, P(\{w_n\})$; but if $P(\{w_2\})$ is less than or equal to $P(\{w_3\}) + \dots + P(\{w_n\})$ then consider $P(\{w_1\}), P(\{w_2\}), P(\{w_3\})$: and so forth. The procedure is terminated when the least subset of W of probability 1 is reached."

a given context depends on the stakes, interests, and focus of attention of the believer.

With respect to how thresholds are determined, Leitgeb's current view differs from Leitgeb (2013), and Arló-Costa and Pedersen, who state that the threshold for rational belief in a given context is always given by the strongest P-stable proposition. Hence, on their view, the agent is rationally required to believe the proposition $\{w_1, w_2\}$ in this scenario, and everything that is entailed by it, whereas it is open to Leitgeb (2014, 2015) to set the threshold for rational belief higher, for example at 0.91, which is given by $P(\{w_1, w_2, w_3\})$, or even at 1.

There is one more ingredient to Leitgeb's current version of the stability theory, namely the way the agent's perspective partitions the space of possibility. We have so far assumed that the agent is taking a maximally fine-grained perspective on the example, by considering each of the six possible worlds separately, as its own cell in a partition. But we often don't distinguish between all the possibilities we could in principle distinguish, because we don't always need to pay attention to every detail. Hence, agents can lump separate worlds together into partition cells, and thereby adopt a more coarse-grained view of a scenario. Sometimes this can make a difference with respect to which propositions are P-stable, and hence which beliefs are rational, which will become important in the following sections.

I will now proceed to explain the motivations for two further constraints on rational belief, and examine whether the stability theory can accommodate them.

3 Lottery propositions and statistical evidence

The following two constraints on what constitutes rational outright beliefs have been proposed in the literature. They are meant to capture our intuitions about the ordinary notion of belief.

4. It is irrational to have outright beliefs in lottery propositions.
5. It is irrational to hold outright beliefs based on purely statistical evidence.

To see the reasoning behind (4), we will first consider an instance of the lottery paradox. The lottery paradox is a well-known problem for theories of rational outright belief that endorse simple versions of the Lockean thesis (Kyburg 1961). Suppose it is rational for an agent to believe any proposition p as long as her credence in p is 0.99 or higher. If there is a fair lottery with 100 tickets, the agent will have a credence of 0.99 that any given ticket will lose, which makes it rational for her to believe of each ticket that it is a loser. If belief is closed under conjunction, she will also believe the conjunction of all these claims. However, she also rationally believes the negation of this conjunction, because she has credence 1 that one of the tickets will win. But rational agents shouldn't have contradictory beliefs. We can generate the same paradox for other thresholds by varying the size of the lottery. Hence, we must reject one of the assumptions generating the paradox. One obvious assumption to reject is that it is rational to believe any 'lottery propositions', such as "ticket # n is a loser." This response seems quite intuitive: if I can rationally believe that my ticket will lose, why buy it in the first place? Why not

throw it away? Moreover, I can rationally think that I might win, which is not consistent with believing that I will lose.

We just considered a fair, one-guaranteed-winner lottery. But, as has been pointed out by numerous authors, intuitions supporting the irrationality of believing lottery propositions generalize to all sorts of lotteries. They don't depend on whether there is a guaranteed winner or not, how high the prize is, or whether all the tickets have the same chance of winning. Moreover, they seem to generalize to situations that merely resemble lotteries.⁴ These considerations speak in favor of accepting (4) as a constraint on rational outright belief.

Constraint (5) has recently been argued for by Lara Buchak (2014), on the basis of intuitive as well as legal considerations. She argues that whether a belief is rational can depend on the type of evidence it is based on.⁵ Here is a version of the well-known blue bus case that illustrates the point:

Your car was hit by a bus in the middle of the night. The bus could either belong to the blue bus or the red bus company.

Version A You know that the blue bus company operates 90 % of the buses in the area, and the red bus company only 10 %. Hence, you have a 0.9 credence that a blue bus is to blame.

Version B The red and the blue bus company each operate half of the buses in the area, and a 90 % reliable eyewitness says that she saw a blue bus hit your car. Hence, you have a 0.9 credence that a blue bus is to blame.

Buchak argues that it is rational to have an outright belief that a blue bus is to blame in version B, but not in version A. This is because you have different types of evidence in each case: your evidence is purely statistical in version A, but you have causal evidence that has a specific relation to the infraction in version B. Her argument relies partly on intuitions about cases, but she also shows that these verdicts are reflected in the law. There are numerous legal cases that show that for a court to reach a conviction, purely statistical evidence is insufficient, even if it

⁴ For example, consider a fair lottery with ten tickets, in which there is only a 10 % chance that the lottery will have any winner. No inconsistency ensues from believing that every ticket (each of which has a 1 % chance of winning) is a loser, because the agent might simply believe that there will be no winner (as long as a credence of 0.9 is sufficiently high to rationalize belief. But we can of course change the numbers to accommodate other thresholds). It seems intuitively plausible to say that even in this case, belief in the lottery propositions is irrational, even though the agent's beliefs aren't inconsistent.

For the claim that it shouldn't matter whether a lottery has an even probability distribution or not, see for example: McKinnon (2013), Christensen (2004, p. 58), Hawthorne (2004, pp. 8–9), Nelkin (2000), Williamson (2000, p. 248), DeRose (1996) and Vogel (1990).

Some authors discuss lottery propositions in the context of whether they can be known, rather than believed. Yet, some supporters of knowledge norms, such as Williamson, endorse the further claim that one ought not to believe something unless one knows it. It follows that one ought not to believe lottery propositions. (Williamson 2000, p. 255)

Some of the propositions that Hawthorne classifies as lottery propositions could also be subsumed under (5), because they are propositions for which we have purely statistical evidence, for example "I will not die in an accident in the next year." Or "Not all of these golfers will get a perfect score."

⁵ Buchak's paper is not the first to make the distinction between statistical and nonstatistical evidence, and to discuss its relation to rational belief. For further references, I invite the reader to consult the extensive bibliography in her paper.

justifies high confidence. Hence, Buchak endorses (5). Intuitions against the rationality believing on the basis of purely statistical evidence also help explain why (4) seems compelling for all sorts of different lottery propositions: usually, the only evidence we have for lottery propositions is statistical, so that might be the reason why we think they should not be believed.

Buchak further concludes that we cannot identify rational belief with a credence above a threshold within a context, because the type of evidence also matters. Hence, any view that relies only on a threshold and a context to explicate the relationship between rational belief and credence must fail. Yet, Leitgeb's view is richer than the types of views Buchak considers, so it might be able to accommodate (5).

4 Can Leitgeb's view accommodate constraints (4) and (5)?

4.1 The easy cases

The stability theory of belief can handle the fair, one-guaranteed-winner version of the lottery paradox without rejecting the threshold assumption or conjunctive closure, but the solution depends on the viewpoint of the agent in a given context. Suppose first that she is interested in the question of which ticket will be drawn. She considers each ticket as a separate possibility, thereby adopting a viewpoint that assumes a maximally fine-grained partition, with each ticket being a separate cell. In this case, no proposition of the form "ticket # n will lose" is P-stable, and hence no such proposition can be rationally believed. The reason why, for example, the proposition "ticket #1 will lose" is not P-stable is the following: if the agent learns that all tickets have been ruled out except #1 and one other ticket, then she must update her credences so that $P(\text{ticket \#1 will lose}) = 1/2$. But this is incompatible with the definition of P-stability. The same holds for every other ticket. The only P-stable proposition is "Some ticket will win", in which the agent has credence 1. Hence, the threshold for rational belief in this context is 1, the paradox is avoided, and (4) is accommodated (Leitgeb 2014).

Yet, the agent could also adopt a different viewpoint, which partitions the space of possibilities in a more coarse-grained fashion. If, for example, she wonders about whether ticket # n will be drawn versus whether any other ticket will be drawn, she considers a partition with only two cells, where the former has a probability of 0.01 and the latter has a probability of 0.99. In this case, the proposition "some ticket other than # n will be drawn" is P-stable, and the threshold for rational belief is 0.99. (The reader may confirm the results in this section via the algorithm in footnote 3). While the lottery paradox doesn't arise in this case, it looks like the view permits belief in a lottery proposition. However, Leitgeb can respond as follows: while a situation can be partitioned in different ways, some ways of partitioning the space of possibilities are far more natural than others, and hence should be preferred when deciding whether a proposition can rationally be believed. In the case of a lottery, it is arguably much more natural to consider each ticket as a separate cell in the partition than to lump tickets together into more coarse-grained partitions. As long as we grant Leitgeb this response, his treatment of the fair, one-guaranteed-winner lottery accommodates (4).

The appeal to natural ways of partitioning the space of possibilities also works for the blue bus case. Version A of the bus case is naturally viewed as a kind of lottery: suppose there are 100 buses in the city, and each of them has an equal epistemic probability of 0.01 of being involved in the accident. Each bus thus constitutes a single cell in the partitioning of the case, and so the claim that a blue bus is responsible amounts to the disjunction “blue bus 1 caused the accident or blue bus 2 caused the accident ... or blue bus 90 caused the accident.” The agent’s credence of 0.9 in this disjunction is not stably high according to Leitgeb’s definition, and hence his view correctly forbids rational belief in the disjunction. (This can be easily confirmed with the algorithm in footnote 3.)

By contrast, version B of the case lends itself more naturally to coarse partitioning: the agent only considers whether a blue bus or a red bus caused the accident. The reliability of the eyewitness is taken into account in assigning a rational credence, but it doesn’t lead to a perspective on the case that is lottery-like. If the agent adopts this coarse-grained view of the case, then her 0.9 credence that a blue bus caused the accident meets the stability condition, and the proposition is a candidate for rational outright belief. Hence, in the blue bus case, the natural ways of partitioning up the space of possibilities nicely coincide with the distinction between statistical and causal evidence, and belief on the basis of statistical evidence is forbidden, while belief on the basis of causal evidence is allowed.

4.2 The hard cases

For Leitgeb’s view to fully accommodate constraints (4) and (5), it is important that beliefs in lottery propositions, or on the basis of merely statistical evidence, are systematically declared irrational. But unfortunately, the neat results we saw in the last section don’t generalize.

Let’s return to the example from Sect. 2. We assumed that the agent has the following credences over an algebra of six possible worlds:

$$P(\{w_1\}) = 0.5, P(\{w_2\}) = 0.3, P(\{w_3\}) = 0.11, P(\{w_4\}) = 0.05, \\ P(\{w_5\}) = 0.03, P(\{w_6\}) = 0.01$$

We saw that there are various propositions in the algebra in which the agent has stably high credence, and which are therefore candidates for rational outright belief. The strongest of these propositions is $\{w_1, w_2\}$. I originally didn’t fill the example with content, but we can simply interpret it as a six-ticket lottery, in which the tickets have different chances of winning: in world 1, ticket 1 wins, and so on. As we can easily see, even a fine-grained partitioning of the case, which views each world—or each ticket—as a separate cell in the partition, doesn’t forbid outright belief. If we accept the lowest possible threshold of 0.8, set by the proposition $\{w_1, w_2\}$, then the owners of tickets 3–6 can rationally believe that their tickets will lose. Hence, belief in a lottery proposition such as “ticket #3 will lose” is not automatically forbidden by Leitgeb’s view. This result by itself might strike the reader as surprising, given that the chance that each of the tickets 3–6 is a loser is much

lower than in a lot of even lottery cases in which belief is forbidden by the stability theory.

The problem for uneven lotteries arises in a similar way for statistical evidence examples. The stability theory can handle version A of the blue bus case, because it can naturally be viewed as a fair lottery. But there are other statistical evidence cases that are most naturally construed as analogs to uneven lotteries. We can simply use our six-world toy model and credence function P again, and fill it with different content. Imagine a scenario in which you are facing a group of six people, who have all attended a soccer game. You know that one of the six people jumped the fence to watch the game, the other five paid for their tickets. The only evidence about who might be the fence jumper is this: you know that each person sat in a different section, and you know the proportion of fence jumpers and paying visitors in each section. In section one, 50 % of the visitors were fence-jumpers, in section two, 30 % were fence-jumpers, and so on. The numbers are just the same as in our uneven lottery. According to (5), it would be irrational for you to form any outright beliefs about which person is or isn't the fence-jumper, since the only available evidence is statistical evidence about the percentage of fence-jumpers in the sections in which each person sat. But the verdict of the stability theory is just the same as before: if the lowest available threshold is chosen, you can believe that either the person who sat in section one or the person who sat in section two jumped the fence, and rule out everyone else.

Leitgeb (2013), and Arló-Costa and Pedersen are committed to the view that the logically strongest P -stable proposition determines the threshold for rational outright belief, hence they are committed to the problematic claims about which beliefs are rational in the examples I just discussed, and hence their views are definitely incompatible with (4) and (5). Leitgeb's current view is more flexible, because he allows that the threshold can be chosen from the range of permissible thresholds. What are the possible responses he could adopt?

One possibility that is open to all proponents of the stability theory is to deny that the theory is intended to capture our ordinary notion of belief, and hence deny that (4) and (5) are adequate constraints that the view needs to accommodate. Yet, this response leaves unexplained why the alternative notion of belief is worth considering, and it is also at odds with Leitgeb's own aims (Leitgeb 2014, 2015). If, on the other hand, the view is meant to give an account of the ordinary notion of belief, Leitgeb has another response available to him.

We need to ensure that in lottery cases and statistical evidence cases, the threshold for belief is always 1, regardless of whether there are lower thresholds available. Leitgeb can help himself to his idea here that threshold selection for the stability theory is intended to mirror SSI about knowledge. But instead of determining whether or not a subject knows that p , the agent's attention, stakes, and interests are supposed to determine the threshold for rational belief (Leitgeb 2014). Thus, according to the credence-belief version of SSI, whether or not a subject S can rationally believe that p depends on whether S 's credence in p meets the relevant conditions in the context under consideration. The conditions for rational belief can be more or less demanding, depending on what the subject's interests are, what error

possibilities the subject attends to, and what the stakes of the situations are, i.e. whether having a true belief about p is of great practical importance or not.

The knowledge version of SSI explains why we generally don't take ourselves to know lottery propositions when we explicitly consider them. Lottery-like situations are generally accompanied with the realization that our evidence doesn't distinguish between the case in which we're right about p and the case in which we're wrong about p . For a rational agent to consider her evidence sufficient for knowledge of some proposition p , she needs to think not only that her evidence makes it likely that p is true, but also that, were p false, she'd probably not have the exact same evidence she currently has. Someone who holds a lottery ticket knows that her ticket will most likely lose. But she also knows that, if her ticket were the winning one, her evidence would be exactly the same. Hence, she doesn't take her evidence to be sufficient for knowing that her ticket is a loser. (This type of reasoning is called 'parity reasoning' by Hawthorne 2004). By contrast, suppose the ticket holder reads in the newspaper that some ticket other than hers won the lottery. She has good reason to think that, if her ticket did in fact win, her evidence would most likely be different—the newspaper would say that she won. Hence, she takes the evidence from the newspaper to be sufficient for knowing that she lost.

We can give an analogous explanation for why we usually don't take ourselves to know propositions based on purely statistical evidence. In the example introduced earlier, my evidence that person six is the fence jumper is exactly the same, whether or not she actually jumped the fence. By contrast, if I instead found out that person six had a ticket in her name, then my evidence would be different if she were a fence-jumper, because then she wouldn't have a ticket. Hence, I could come to know that person six wasn't a fence-jumper based on seeing her ticket, but not based on the fact that she was sitting in a section with very few fence-jumpers.

Using a version of SSI to determine thresholds for belief will arguably cover all instances of (4) and (5): given that rational agents are generally disinclined to believe lottery propositions, or propositions that are purely based on statistical evidence, their point of view will always determine a threshold of one in the situations in question.

This seems like a satisfying solution at first glance, because it covers all instances of (4) and (5). Yet, considerations of systematicity raise the worry that this result isn't as good as it appears. This is because it is not obvious that the details of the stability theory now play a systematic role in determining when belief is appropriate and when it isn't. Leitgeb presents it as a good-making feature of his view that it helps avoid the lottery paradox, as I explained in the previous section. However, if an agent's attention, stakes, and interests, prevent agents from believing lottery propositions and propositions based on statistical evidence in general, then (4) and (5) can be accommodated by appealing to subject sensitive invariantism alone, regardless of whether the stability theory forbids belief or not in a given case. And this leads to a more general worry: while SSI is not a formal theory of the relationship between rational credence and rational belief, it might in the end do all the heavy lifting: if we could argue that a rational agent's attention, stakes and interests prevent them from forming inconsistent beliefs, alongside a range of other beliefs that aren't rational to have, then SSI looks like a promising theory of the

relationship between rational credence and rational belief all by itself. And if this is right, then it is not at all obvious what role the stability theory plays in characterizing the relationship between rational belief and rational credence.

A proponent of the stability theory might defend the usefulness of her view as follows: all we get from the stability theory of belief is a set of necessary conditions for rational belief. For some proposition p to be a candidate for rational belief relative to an agent's credence function in a given context, it is necessary for the agent to have a stably high credence in p . Then, in order to determine whether any of the propositions in which the agent has stably high credence are in fact rationally believable by the agent, we must consult the agent's interests, stakes, and attention, which will give us sufficient conditions for rational belief. The idea here is that SSI alone isn't always precise enough to determine necessary conditions for belief, but it can help us select which propositions can be believed from the set of candidates singled out by the stability theory.

Now, if we combine SSI and the stability theory in this way, there is something SSI can't do: it can't lift restrictions on belief, i.e. it can't override the stability theory's verdicts on when belief is impermissible. Thus, it would be problematic if there were cases in which our intuitive judgment and SSI agree that belief should be permitted, yet, belief is forbidden by the stability theory. If we can find a case like this, then we cannot invoke the stability theory to provide necessary conditions for outright belief. It turns out that there are examples in which this situation obtains.

We saw earlier that a proponent of the stability theory can exploit the partition-sensitivity of the view to explain why belief is forbidden in at least some instances of (4) and (5). In this regard, partition-sensitivity looked like a good-making feature of the view. Unfortunately, partition-sensitivity leads to very problematic results in some contexts. Consider an agent who has the following credences regarding p and q .⁶

$$\begin{aligned} P(p \& q) &= 0.44 && \text{(call the } p \& q \text{ world } w_1) \\ P(p \& \sim q) &= 0.32 && (w_2) \\ P(\sim p \& q) &= 0.18 && (w_3) \\ P(\sim p \& \sim q) &= 0.06 && (w_4) \end{aligned}$$

The strongest p -stable proposition here is $\{w_1, w_2\}$, which means that the agent can rationally believe that she is in a world in which p is true (assuming we're in a scenario where SSI doesn't forbid belief). Now, we can imagine that the agent goes on to consider a slightly more fine-grained partition of worlds. For each world, she

⁶ This problem with the partition-sensitivity of Leitgeb's view is pointed out independently by Fitelson (2015) and Schurz (2015). Fitelson shows how coin-toss expansions of an agent's doxastic space render initially permissible beliefs impermissible, and he uses these results to demonstrate that the stability theory's verdicts look bad in light of confirmation theory.

Schurz proves a general theorem about the kinds of epistemic expansions that always render initially permissible beliefs impermissible: take any analytically independent expansion of an agent's doxastic space, where each possibility of the old partition is split up into the same number of new possibilities. Let n be the number of doxastic possibilities in the set that is used to expand the agent's original doxastic space, and let t be the threshold for rational belief before expanding the agent's doxastic space. According to Schurz' Theorem 3, the set of initially permissible beliefs will not remain permissible after an analytically independent expansion where $n \geq t/(1 - t)$.

will consider the possibility that a coin flip landed either heads (h) or tails (t). This coin flip is completely independent of p and q. Hence, considering the outcome of the coin flip in addition to p and q should not make any difference to the agent's rational belief in p. Here's the new, more fine-grained set of worlds under consideration:

$$\begin{aligned}
 P(p \& q \& h) &= 0.22 & (w_1) \\
 P(p \& q \& t) &= 0.22 & (w_2) \\
 P(p \& \sim q \& h) &= 0.16 & (w_3) \\
 P(p \& \sim q \& t) &= 0.16 & (w_4) \\
 P(\sim p \& q \& h) &= 0.09 & (w_5) \\
 P(\sim p \& q \& t) &= 0.09 & (w_6) \\
 P(\sim p \& \sim q \& h) &= 0.03 & (w_7) \\
 P(\sim p \& \sim q \& t) &= 0.03 & (w_8)
 \end{aligned}$$

We said that intuitively, the agent should still be able to rationally believe p when she considers the outcome of the coin flip, since the coin flip has nothing to do with p. But that is not the verdict that the stability theory delivers. The strongest p-stable proposition is now $\{w_1, w_2, w_3, w_4, w_5, w_6\}$. Since w_5 and w_6 are $\sim p$ -worlds, the agent can no longer believe that p is true. Yet, SSI, if considered on its own, would side with our intuition here that belief in p is still permitted, even after the coin flip is taken into consideration. If the subject's interests, attention, and stakes permitted belief in p before the coin flip was taken into account, and those factors remain unchanged, then considering the coin flip should not make a difference. Hence, we have a scenario in which belief is intuitively permitted, and it is permitted by SSI, but it's forbidden by stability theory.

This is not a good result for the stability theory. When we explored how the view can accommodate (4) and (5), we found that it only sometimes forbids belief in lottery propositions and propositions that are based on statistical evidence. We then considered the suggestion that we can appeal to a credence-belief version of SSI to explain why belief is forbidden in these cases. Yet, appealing to SSI raised the question of what contribution the stability theory makes to determining what an agent can rationally believe, once it is supplemented with SSI. A promising proposal seemed to be that the stability theory provides necessary conditions for rational belief. From the candidate propositions provided by the stability theory, SSI selects the ones that the agent can rationally believe in her circumstances. Yet, we just demonstrated that this proposal is undermined by the observation that the stability theory, due to its partition sensitivity, is too restrictive in certain cases. Hence, it is hard to see how it can even be said to provide necessary conditions for outright belief.

A proponent of the stability theory might insist that in the coin flip example, our intuitions are misguided, and it is in fact irrational for the agent to believe that p once she considers the coin flip. However, this leads us back to the first horn of the dilemma: if the stability theory must reject the intuition that considering irrelevant propositions should not change our rational beliefs, then it is no longer clear whether the stability theory gives an account of the ordinary notion of belief.

Alternatively, if the proponent of the stability theory allows that judgments about rational belief that are supported by SSI can sometimes override the stability theory's prohibitions on belief, then she must accept that the view provides neither necessary nor sufficient conditions for rational outright belief. This response makes it difficult to see how the stability theory makes any important, systematic contribution to determining which beliefs are rational for an agent.

5 Conclusion and prospects

The stability theory of belief is not the only formal theory that has been proposed as an explication of the relationship between rational credence and rational outright belief. However, it is more sophisticated than some earlier views, such as simple threshold views (e.g. Foley 2009), because it can accommodate the idea that there are consistency and closure constraints on our beliefs. We discussed reasons to question the adequacy of the stability theory in particular, but some of the lessons we learned along the way will be instructive for evaluating formal theories of the relationship between rational credence and rational outright belief in general.

We saw that (4) and (5) impose restrictions on our rational outright beliefs that are independent of the shape, or formal features, of a credence function. We can take any credence function and fill the propositional variables with content so as to generate a case in which the relevant propositions are lottery propositions, or where they are supported by statistical evidence. As a result, we cannot have any formal theory of the relationship between rational belief and rational credence that can get by without supplementary constraints to forbid belief.⁷ These supplementary constraints should not be ad hoc—they should provide a natural way of extending the formal theory. However, once the supplementary constraints are in place, it is important to ensure that the formal part of the view still carries some of the explanatory burden, which can be difficult, given that we need to add supplementary constraints that are quite powerful. One way to ensure that the formal view still plays a significant role is to argue that it functions as a first filter, and the candidates for rational belief are then further narrowed down by the supplementary constraints. However, this strategy only works if the formal theory is sufficiently permissive, and doesn't rule out candidates for outright belief that we judge to be clearly permissible. Of course, the formal view must not be too permissive either, since then it becomes questionable how much work it does in explaining which of our beliefs are rational. Meeting these constraints is not trivial, and it is an open question whether there is an alternative to the stability theory that fares better. The criteria I have developed in my discussion of the stability will hopefully help guide our judgment about these matters.

Acknowledgments I would like to thank Fay Edwards, Branden Fitelson, Alan Hájek, Hannes Leitgeb, Hanti Lin, Richard Pettigrew, Ryan Platte, Brian Talbot, and an anonymous referee for helpful comments and discussion.

⁷ Of course, we could have a theory that only allows for rational belief if the agent has a rational credence of 1. But this view is not a serious contender in the first place.

References

- Arló-Costa, H., & Pedersen, A. P. (2012). Belief and probability: A general theory of probability cores. *International Journal of Approximate Reasoning*, 53(3), 293–315.
- Buchak, L. (2014). Belief, credence, and norms. *Philosophical Studies*, 169, 285–311.
- Christensen, D. (2004). *Putting logic in its place*. Oxford: Oxford University Press.
- DeRose, Keith. (1996). Knowledge, assertion, and lotteries. *Australasian Journal of Philosophy*, 74, 568–580.
- Fitelson, B. (2015). *Belief and credence: The view from Naïve epistemic utility theory*. Manuscript.
- Foley, R. (2009). Beliefs, degrees of belief, and the Lockean thesis. In: F. Huber, & C. Schmidt-Petri (Eds.), *Degrees of belief*, Synthese Library 342. Springer.
- Hawthorne, J. (2004). *Knowledge and lotteries*. Oxford: Oxford University Press.
- Kyburg, H. E. (1961). *Probability and the logic of rational belief*. Middletown: Wesleyan University Press.
- Leitgeb, H. (2013). Reducing belief *Simpliciter* to degrees of belief. *Annals of Pure and Applied Logic*, 164, 1338–1389.
- Leitgeb, H. (2014). The stability theory of belief. *Philosophical Review*, 123(2), 131–171.
- Leitgeb, H. (2015). I—The humean thesis on belief. *Aristotelian Society Supplementary*, 89(1), 143–185.
- McKinnon, Rachel. (2013). Lotteries, knowledge and irrelevant alternatives. *Dialogue*, 52, 523–549.
- Nelkin, D. K. (2000). The lottery paradox, knowledge, and rationality. *The Philosophical Review*, 109(3), 373–409.
- Schurz, G. (2015). *Impossibility results for rational belief*. Manuscript.
- Vogel, J. (1990). Are there counterexamples to the closure principle? In M. Roth & G. Ross (Eds.), *Doubting: Contemporary perspectives in Skepticism*. Boston: Kluwer.
- Williamson, T. (2000). *Knowledge and its limits*. Oxford: Oxford University Press.