TWO APPROACHES TO BELIEF REVISION

TED SHEAR, BRANDEN FITELSON, AND JONATHAN WEISBERG

Abstract. In this paper, we compare and contrast two methods for the qualitative revision of (*viz.*, "full") beliefs. The first ("Bayesian") method is generated by a simplistic diachronic Lockean thesis requiring coherence with the agent's posterior credences after conditionalization. The second ("Logical") method is the orthodox AGM approach to belief revision. Our primary aim will be to characterize the ways in which these two approaches can disagree with each other — especially in the special case where the agent's belief set is deductively cogent.

1. Setup

We will be considering a very simple and highly idealized type of epistemic agent who possesses both credences (*viz.* numerical degrees of confidence) and ordinary (qualitative) beliefs.[1] The objects of these attitudes will be (classical, possible world) propositions in a finite propositional language whose logic is classical. On the credence side, we adopt a naïve Bayesian account of credences where the agent's credences are represented by a classical *probability* function, $b(\cdot)$. When it comes to the qualitative attitudes of our idealized agent, we will attend only to the *beliefs* of the agent (*i.e.*, we will not discuss disbelief or suspension of judgment) and the agent's belief state will be represented by a set, $\mathbf{B}$, comprising the set of propositions the agent believes. At times, it will be helpful to refer specifically to individual beliefs of the agent and we will write $\mathsf{B}(X)$ indicating that $X \in \mathbf{B}$.

The primary purpose of this paper is to investigate the *diachronic* norms governing the revisions made by agents to their beliefs when they receive new information. To do so, we will introduce a novel, broadly Bayesian, method of belief revision and contrast it with the orthodox AGM account. So-called *belief revision operators* are functions that map *prior* belief sets together with a proposition to *posterior* belief sets. So, where $\star$ is an arbitrary belief revision operator, when an agent with the prior belief set $\mathbf{B}$ learns the proposition E, she her posterior belief set is $\mathbf{B}' = \mathbf{B} \star E$. While AGM's belief revision operator ($*$) is defined in wholly qualitative terms, our novel *Lockean* belief revision operator ($\circledast$) is defined by way of a diachronic version of Foley's [14] *Lockean thesis*. That is, Lockean revision requires that when an agent learns E, she adopts the posterior belief set, $\mathbf{B}' = \mathbf{B} \circledast E$, such that $\mathsf{B}(X)$ just in case her posterior credence in X is no less than the *Lockean threshold*, t. So, for an agent to perform Lockean revision, she will require some procedure for updating her *prior* credence function, $b(\cdot)$, to a *posterior* credence function, $b'(\cdot)$. Since it is the simplest and most straightforward option available, we will assume that our agents update their credence function *via* conditionalization.[2] However, this choice is made largely for convenience and most (if not all) of our results will continue to hold for any update generating the posterior b' (on learned proposition E) which satisfies the following two constraints: (i) $b'(E) > b(E)$ and (ii) if $b'(X) \geq t$, then $b(E \supset X) \geq t$.[3]

In the next section, we will introduce Lockean revision and discuss some initial results designed to clarify how the dialectic will unfold. In section three, we will discuss AGM revision and some related issues. In the subsequent four sections, we will provide our novel results, primarily concerning the precise similarities and differences between these two approaches to belief revision. Finally, we will close with some remarks about open questions and future work. In a brief epilogue, we contrast Lockean revision with Leitgeb's [33] recent account of belief revision, which satisfies both the Lockean thesis and AGM's revision postulates.

2. Lockeanism and its Revision Operator

While we are presently concerned with *diachronic* rational norms on belief, both Lockean revision and AGM revision presuppose certain *synchronic* constraints as

Date: 05/02/17. *Draft*: Do not cite or quote without permission.

We would like to thank audiences at Tilburg, Boulder, Maryland, Helsinki, Kent, Taipei, Toronto, Bristol, and Gröningen for stimulating discussions. Individually, we must single out Luc Bovens, Kenny Easwaran, Haim Gaifman, Konstantinos Georgatos, Jeremy Goodman, Justin Fisher, Kevin Kelly, Hannes Leitgeb, Isaac Levi, Hanti Lin, Eric Pacuit, Rohit Parikh, Richard Pettigrew, Eric Raidl, Hans Rott, Jonah Schupbach, Teddy Seidenfeld, and Julia Staffel for providing useful feedback. We also wish to note our sincere gratitude to Hans Rott for his generous and helpful comments on late drafts of this work.

[1]For the purposes of this paper, we remain neutral on ontological questions regarding the relationship between credences and beliefs. Although we are inclined towards a pluralistic approach admitting the existence and independence of both types of cognitive attitudes, none of the content of this paper requires the adoption of any particular view about their existence or relative fundamentality.

[2]Readers who are already familiar with the literature on belief revision will likely recognize that this update procedure on credences is the credal analogue of qualitative update known as *expansions*. Qualitatively, an expansion is performed when an agent simply adds a proposition to her stock of beliefs. As such, expansions capture updates by propositions that are *consistent* with her prior belief set. Revisions, on the other hand, capture the more general case in which there is no guarantee that the new proposition is consistent with the agent's priors. Since conditionalization is undefined when the learned proposition receives a prior probability of zero, it may be seen as the credal analogue of qualitative expansion. This point is relevant to the current application because the novel broadly-Bayesian approach to qualitative revision will be driven by credal expansion. Nonetheless, the new revision operator may be aptly viewed as a qualitative revision operator since it permits revision by a proposition that is *logically* (*viz.* qualitatively) inconsistent with the agent's prior belief set.

[3]In an earlier draft, we had noted that the results hold provided the stronger requirements — which hold only for strict conditionalization — that (i) $b'(E) > b(E)$, (ii) $b'(E) \geq t$ (where t is the agent's Lockean threshold), and (iii) $b(E \supset X) \geq b'(X)$. However, Konstantin Genin [23] has generalized some of our more interesting results to the case of Jeffrey conditionalization by showing that they hold provided the weaker pair of conditions mentioned above.

well. We will begin our discussions of each with brief descriptions of the synchronic presuppositions of each account. For the Lockean, the underlying synchronic constraint on and agent's full beliefs is that they cohere with their credences in a certain way. This constraint of our first approach was first dubbed the *Lockean thesis* by Foley [14].

Intuitively, the thesis requires that an agent believe all and only those propositions to which she assigns sufficiently high credence. More carefully, the Lockean thesis requires that an agent believe X iff her credence in X is at least t, where the agent's *Lockean threshold* $t \in (1/2, 1]$. Formally:

Lockean Thesis. $\mathsf{B}(X)$ iff $b(X) \geq t$, where $t \in (1/2, 1]$.

This normative constraint is intuitively plausible since: (1) it seems irrational for an agent to believe a proposition which she takes to be (sufficiently) improbable, (2) it appears to be a rational shortcoming if an agent fails to believe a proposition that she thinks is (sufficiently) highly likely, and (3) it is never rationally permissible for an agent to concurrently belief both X and $\neg X$ (as would be permissible if $t \leq 1/2$).

Despite its intuitive plausibility (and the fact that it can be given a very elegant justification *via epistemic utility theory* — see §7), the Lockean thesis remains rather controversial. There are a variety of objections one may raise against it. The earliest objections stem from the fact that it permits (indeed, *requires*, in some cases) belief sets that are *logically inconsistent* and which *fail to be closed* under logical consequence. Kyburg's [31] lottery paradox is the most well-known example of this phenomenon. As we will see in the next section, AGM requires both consistency and closure of all qualitative belief sets. So, this is a key difference between our two updating paradigms. We don't have much to add to this traditional and well trodden debate between "Bayesian" and "Logical" approaches to revision [3, 14, 9]. Instead, our analysis will primarily focus on *deductively cogent* agents. Even when our attention is restricted to deductively cogent agents, there remain some crucial differences and interesting connections between the two approaches.

Beyond the failure of closure and consistency, it is commonly complained that there appear to be only two non-arbitrary (*viz.*, non-context-dependent) candidate values for the Lockean threshold: $1/2$ and 1. As alluded to above, given our current formulation of the thesis, were an agent to assign equal credence to a proposition and its negation, a Lockean threshold of $1/2$ would permit belief in both. Nonetheless, even if the thesis were stated with a strict inequality (*viz.* $\mathsf{B}(X)$ iff $b(X) > t$), a $1/2$-threshold would be *too permissive*. Surely, rationality does not (always) *require* belief when an agent is only *slightly* more confident in a proposition than its negation. Similarly, the *extremal* Lockean threshold ($t = 1$) is *not permissive*

enough, since this would make *certainty* a rational requirement for belief.[4] It is typically assumed (for these reasons) that appropriate, particular Lockean thresholds are determined in a *context-dependent* way (although, in §7, we will show how to *derive* rational Lockean thresholds from an agent's epistemic utility function).

One of the novel and important contributions of this paper will be to establish a new non-arbitrary and non-context-dependent Lockean threshold: the inverse of the golden ratio ($\Phi^{-1} \approx 0.618$), which we will refer to as the *golden threshold*. In the second half of this section, we will provide a demonstration that Lockean revision must require $t \in (\Phi^{-1}, 1]$ on pain of violating a well-motivated, purely qualitative constraint on belief revision. But, first, we must define our Lockean (or Bayesian) belief revision operator.

Lockean revision requires that an agent revise her beliefs so as to satisfy the Lockean thesis, relative to her posterior credences (*i.e.*, her prior credence function *conditionalized* on her new evidence).

> **Lockean revision**. Where $\mathbf{B}$ and b are the agent's prior belief set and and credence function respectively, and E is a proposition, the Lockean revision of $\mathbf{B}$ by E (*i.e.* $\mathbf{B} \circledast E$) is defined as:
>
> $$\mathbf{B} \circledast E := \{X \mid b(X \mid E) \geq t\}.$$

This belief revision procedure is entailed by jointly insisting that *at any given time* an agent's beliefs and credences should jointly satisfy the Lockean thesis along with the fact that her credences are updated *via* conditionalization.

Since Lockean revision requires diachronic coherence with the agent's credences, the procedure will be sensitive to certain changes in the agent's non-qualitative information. As we will see shortly, this means that agents may be led to drop some of their beliefs when they receive non-definitive counter-evidence to previously believed propositions. This fact about Lockean revision will ultimately prove to be at the heart of its divergence with AGM revision, which only permits the loss of belief as a result of learning a proposition that is *logically inconsistent* with the agent's prior beliefs.

2.1. **Weak Preservation and the Golden Threshold.** Belief revision operators are often defined by way of qualitative axioms on possible revisions to an agent's prior beliefs. In the next section, we will present one standard axiomatization of AGM's revision operator and, in the subsequent section, assess its standing relative to the Lockean revision operator. But first, as a preview of how this dialectic will unfold, we will discuss two constraints on belief revision operators, which are both entailed

[4] Despite this criticism, we should note that some authors have argued in favor of adopting an extremal Lockean threshold, *e.g.* see [36] or, more recently, [6].

by AGM's axioms. This will serve to expose some interesting features of Lockean revision that will resurface later in the paper.

The first principle that we will consider is known as Weak Preservation, which requires that no beliefs are lost when a belief set is revised by one of its members.[5]

($\star 4^w$) If $E \in \mathbf{B}$, then $\mathbf{B} \subseteq \mathbf{B} \star E$ Weak Preservation

In words: *learning something you already believe should never cause you to stop believing something you already believed.* At first pass, this principle may seem indubitable. After all, when an agent *already* believes E, learning E would not appear to provide her with any new information. Thus, there should be no basis for any change to her beliefs. Although Weak Preservation is *not* generally satisfied by Lockean revision, it *is* generally satisfied by AGM revision.[6]

If a Lockean agent revises by a previously believed proposition, then it is possible for her posterior credence in other previously believed propositions to fall (significantly) below the Lockean threshold. This revision will then result in the loss of a belief. But, there is a bound on the "degree" to which Lockean revision can violate Weak Preservation. That is, there is a precise bound on how low an agent's posterior credence in a previously believed proposition may be after having revised by some other previously believed proposition. This bound can be provided as a function of the Lockean threshold as recorded in the following lemma and visually represented in Figure 1.

Lemma 1 Where $t \in [1/2, 1]$, if $b(X) \geq t$ and $b(E) \geq t$, then $b(X \mid E) \geq \frac{2t-1}{t}$.

Proof. Suppose for reductio that $b(X) \geq t$, $b(E) \geq t$, but $b(X \mid E) < \frac{2t-1}{t}$. First observe that by definition, $b(X \mid E) = \frac{b(X \wedge E)}{b(E)}$, which implies $\frac{b(X \wedge E)}{b(E)} < \frac{2t-1}{t}$. Now assume, with no loss in generality that Assume with no loss in generality that $b(E) = b(X) = t$ and oSince the maximum value of $b(X \wedge E)$ is $b(X) + b(E) - 1 = 2t - 1$ when .

From this it follows that $b(X \mid E) \geq \frac{2t-t}{t}$. □

It follows readily from the proof of Lemma 1 that Lockean revision can violate Weak Preservation.

Proposition 1 Lockean revision *can* violate Weak Preservation. That is, it is possible that:

$$E \in \mathbf{B} \text{ and } \mathbf{B} \not\subseteq B \circledast E.$$

[5]While discussion of Weak Preservation has primarily been provided in the literature surrounding the Ramsey test for conditionals (*e.g.* see [15], [47], or [37]), it is interesting to see that this principle offers a contrastive case between our two approaches to belief revision.

[6]Indeed, this is a weakened version of AGM's characteristic postulate, Preservation, which will be discussed in the next section.

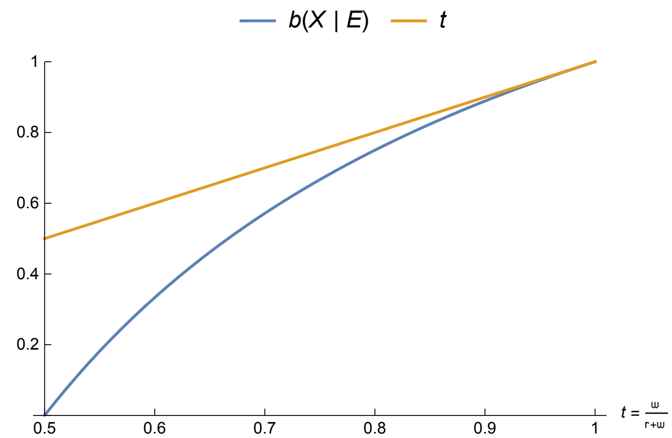


FIGURE 1. Lower bound on $b(X \mid E)$ given $b(X), b(E) \geq t$.

However, notice that as the Lockean threshold *increases*, the "degree" to which Lockean revision can violate Weak Preservation *decreases*. Moreover, in the limit (when $t = 1$), it is *satisfied* by Lockean revision. So, Weak Preservation is "approximately" true, *if* full beliefs have sufficiently high credence (and it is *exactly* true in the extremal case, or when the agents priors are sufficiently high). Additionally, note that the initial motivation for Weak Preservation provided above is consistent with this result. For a Lockean agent, so long as she was not previously certain that E, learning E actually *does* provide her with new information. Although she has not acquired any new *qualitative* information, Lockean revision is sensitive to the finer-grained information represented by her credences.

Next, we report two further results concerning the behavior of Lockean revision when the agent revises by a previously believed proposition. First, Lockean revision can only violate Weak Preservation if the agent's prior beliefs fail to be deductively cogent. Indeed, if we assume that an agent's prior beliefs are deductively closed, then Lockean revision always satisfies Weak Preservation.[7] To wit:

Proposition 2 If **B** is deductively closed, then Lockean revision satisfies Weak Preservation. That is, the following is a theorem:

$$\text{If } \mathbf{B} \text{ is deductively closed and } E \in \mathbf{B}, \text{ then } \mathbf{B} \subseteq \mathbf{B} \circledast E.$$

Proof. Suppose for *reductio* that (a) **B** is deductively closed, (b) $E, X \in \mathbf{B}$, *but* (c) $X \notin \mathbf{B} \circledast E$. From (c), it follows that $b(X \mid E) < t$. Moreover, by the definition of the conditional probability, this implies $b(X \wedge E) < t \cdot b(E) < t$. So, $X \wedge E \notin \mathbf{B}$. But, by (a) and (b), we have $X \wedge E \in \mathbf{B}$. □

[7]We thank Hans Rott for this interesting observation.

The second additional result concerning these sorts of revisions involves a *further* weakening of Weak Preservation, which only requires that an agent not come to *dis*-believe any proposition previously believed upon revising by a previously believed proposition.[8]

$(\star 4^v)$ If $E, X \in \mathbf{B}$, then $\neg X \notin \mathbf{B} \star E$ Very Weak Preservation

While Lockean revision does not, in full generality, satisfy Very Weak Preservation, it does so long as the Lockean threshold is set sufficiently high. Specifically, if we require that the Lockean threshold is greater than the inverse of the Golden ratio ($\Phi^{-1} \approx 0.618$), then Lockean revision will never require the agent to believe the negation of a previously believed proposition as a result of revising by another proposition previously believed. The following proposition and the graphical demonstration provided in Figure 2 confirm that requiring that $t \in (\Phi^{-1}, 1]$ rules out the troublesome cases.[9]

Proposition 3 Lockean revision satisfies Very Weak Preservation if the Lockean threshold $t \in (\Phi^{-1}, 1]$.

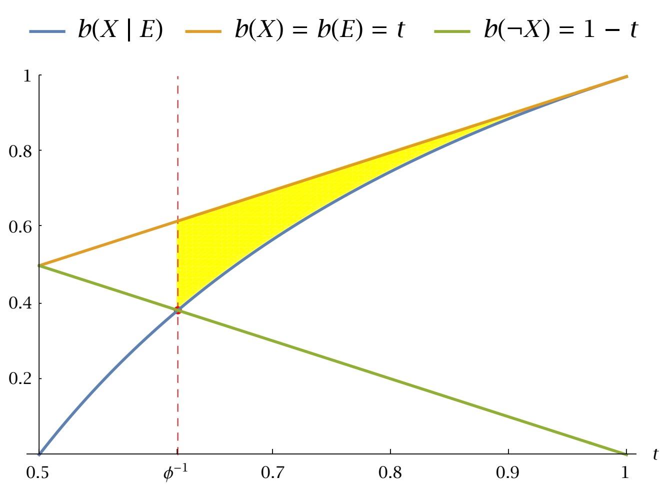


FIGURE 2. $t \in (\Phi^{-1}, 1]$ implies Very Weak Preservation (yellow region).

While the appearance of Φ^{-1} in our constraint may seem surprising, we will see shortly that there is another, independent route to Φ^{-1} that is revealed by a deeper analysis of the relationship between Lockean and AGM revision.

[8] Since neither Lockean revision nor AGM permit belief in both a proposition and its negation after a revision, Very Weak Preservation is seen by both approaches as a strictly weaker constraint than Weak Preservation.

[9] We thank Kenny Easwaran for this helpful observation.

3. AGM and its Revision Operator

The AGM theory of belief revision — first proposed by Carlos Alchourrón, Peter Gärdenfors, and David Makinson in their seminal 1985 paper [1] — has served as the basis for nearly all subsequent research into the dynamics of qualitative belief. In its raw form, the AGM theory offers a characterization of *theory revision* and is often presented in terms of the (so-called) *Gärdenfors postulates*. These postulates are stated in logical and set theoretic terms and are taken as axioms characterizing AGM revision. When understood as postulates for belief revision, they serve to constrain the ways in which an agent may change her *belief set* upon the receipt of new information. Accordingly, belief sets are taken to be *theories* in the mathematical sense, *i.e.* deductively closed sets of sentences.[10] Just as Lockean revision does not uniquely identify a single revision operator (since it is consistent with a range of Lockean thresholds and probability functions), AGM's axioms define a family of operators. Further constraints would be required to generate specific, individual AGM revision operators.

The primary conceptual principle underlying AGM revision is known as *the principle of conservativity* (also called *the principle of informational economy* or *minimal mutilation*).

> **Conservativity**. When an agent with a prior belief set, $\mathbf{B}$, learns E, she should adopt a posterior belief set, $\mathbf{B}'$, such that (i) $\mathbf{B}'$ is *deductively cogent*, (ii) $\mathbf{B}'$ *includes* E, and (iii) $\mathbf{B}'$ is *the closest belief set to* $\mathbf{B}$ that satisfies constraints (i) and (ii).[11]

This motivating principle provides the normative basis for the coherence requirements imposed by the AGM axioms. Condition (i) requires that the agent revises to a belief set that is both *deductively consistent* and *closed under logical consequence*. Condition (ii) simply requires that learned propositions *are adopted* as beliefs. Finally, condition (iii) poses AGM's characteristic restriction on belief sets and specifies that accommodating the learned proposition into the posterior belief set should be accomplished by *making as few changes as possible* to the prior.

We will begin our overview with a discussion of AGM's synchronic coherence requirements by contrasting them with those of Lockeanism. Subsequently, we will complete our overview by briefly explaining each of the Gärdenfors postulates.

[10] The decision to rely on sets of sentences to capture belief sets is largely a matter of historical accident and reflects Gärdenfors's desire to avoid the use of possible worlds, which he viewed as philosophically suspect. Our choice to accord with this convention is made solely out of desire to remain consonant with the most common presentation of the theory.

[11] There are various ways to measure the *distance between belief sets*. The AGM postulates will follow from **Conservativity** for a very wide variety of such distance measures/geodesics [24, 50].

3.1. **The Synchronic Requirements of AGM.** At first pass, one might (mistakenly) be led to think that the AGM theory only imposes diachronic norms on full belief and has no synchronic presuppositions. Not only is this suggested by the statement of **Conservativity** provided above (which only constrains admissible posteriors), but also by the fact that AGM's explicit constraints are put forth in the Gärdenfors postulates that govern its revision operator. Nonetheless, there are compelling reasons to think that the rational requirements of AGM are not exhausted by its diachronic requirements. We will argue — in similar fashion to arguments found in Rott [48, 49] — that AGM must adopt deductive cogency as a standing synchronic requirement. To wit:

> **Cogency**. An agent's belief set should (*always*) be both deductively consistent and closed under logical consequence, *i.e.*, $\mathbf{B} \nvdash \bot$ and $\mathbf{B} = \mathsf{Cn}(\mathbf{B})$, where $\mathsf{Cn}(\mathbf{X})$ is the *deductive closure* of the set $\mathbf{X}$.

This requirement goes beyond **Conservativity**'s insistence that, following revision, agents' posterior belief sets must be deductively cogent. Indeed, it requires that *all* belief sets (both prior and posterior) must be deductively cogent. To appreciate why this strengthening is needed, consider the following attractive principle, which maintains that revising by a *tautology* should not result in any change to an agent's beliefs.

$$(*\top) \qquad \mathbf{B} * \top = \mathbf{B} \qquad \textsf{Idempotence}$$

It is intuitive to think that there would be no rational basis for an agent to change her beliefs if she has simply revised by a proposition that expresses no information at all about the way the world is. Clearly, Idempotence is imminently reasonable and should be satisfied by any adequate belief revision operator.

As expected, both AGM and Lockean revision are guaranteed to satisfy Idempotence.[12] However, if AGM permitted prior belief sets that were either inconsistent or not deductively closed, then Idempotence would expose an inconsistency in the theory.[13] Thus, given AGM's satisfaction of Idempotence and its commitment to **Conservativity**, if AGM is to offer an internally coherent account, it must presuppose **Cogency** as a standing synchronic requirement.[14]

[12] Idempotence follows immediately from AGM's Preservation axiom along largely the same lines as the Weak Preservation principle discussed in the previous section. On the other hand, Lockean revision satisfies Idempotence as a trivial consequence of the fact that $b(\cdot \mid \top) = b(\cdot)$; this straightforwardly implies that $\mathbf{B} \circledast \top = \mathbf{B}$.

[13] The arguments supporting this conclusion are provided in *fn.* 17 and 19.

[14] It is roughly along these lines that Rott [48, 49] has convincingly argued that AGM imposes synchronic coherence requirements in addition to diachronic ones. Rott actually argues for the still stronger conclusion that AGM also imposes *dispositional* requirements governing iterated revisions. However, since our current aim is only to contrast the recommendations of AGM and Lockean revision with respect to single revisions, attending to such considerations are beyond the intended scope of this paper. Thus, we will leave discussion of coherence requirements for iterated revision for future work.

In order to consider the *fundamental* requirements of AGM, it is useful to observe that **Cogency** is just the conjunction of *two* requirements. The first requires an agent's belief set to be *deductively consistent*.

> **Consistency**. At any given time, an agent's belief set $\mathbf{B}$ should be such that there is some possible world w in which every member $\mathbf{B}$ is true, *i.e.*, $\mathbf{B} \nvdash \bot$.

The second requires an agent's belief set to be *closed under logical consequence*.

> **Closure**. At any given time, an agent's belief set, $\mathbf{B}$, should be such that if $\mathbf{B}$ logically implies X, then $\mathsf{B}(X)$, *i.e.*, $\mathbf{B} = \mathsf{Cn}(\mathbf{B})$.

Although both **Consistency** and **Closure** are standing synchronic requirements of AGM, we might wonder whether one ought to be regarded as more *epistemically* fundamental than the other. As convincingly argued by Steinberger [52], if failures of **Consistency** are permitted, then **Closure** will lose much (if not all) of its normative force. For this reason, we suggest that **Consistency** should be seen as the *fundamental* synchronic requirement of AGM.

3.2. **Contrasting Synchronic Requirements of AGM and Lockeanism.** Although our primary interest is in comparing the diachronic requirements of AGM and Lockean revision, we will first briefly discuss how their synchronic requirements relate. We will argue that AGM's synchronic requirements are *more demanding* than those of Lockeanism. For reasons exemplified by Kyburg's lottery paradox [31], the Lockean thesis permits belief sets satisfying neither **Closure** nor **Consistency**. To see why Lockean revision permits violations of **Closure**, observe that it can be probabilistically coherent for an agent to have credences in X and Y such that $b(X) \geq t$, $b(Y) \geq t$, but $b(X \wedge Y) < t$. As such, the Lockean thesis may require belief in two propositions, but not require belief in their conjunction. So, the Lockean thesis permits violations of **Closure**. Moreover, it can be probabilistically coherent for the agent to have credences in $X_1, \ldots, X_n$ such that $b(X_i) \geq t$ for each $i : 1 \leq i \leq n$, but $b(\neg(X_1 \wedge \cdots \wedge X_n)) \geq t$. In this case, the Lockean thesis permits not only a violation of **Closure**, but also of **Consistency**, since it would require the agent to believe each proposition individually, but also believe the negation of their conjunction.

Nonetheless, it is important to recognize that both **Consistency** and **Closure** are *consistent* with the Lockean thesis. After all, they are actually entailed by an extreme form of Lockeanism with a *maximal* threshold ($t = 1$). While this is consistent with Lockeanism, we view **Cogency** as far *too demanding* to be adopted as a universal requirement of rationality. As Foley [14, *p.* 186] explains,

> [...] if the avoidance of recognizable inconsistency were an absolute prerequisite of rational belief, we could not rationally believe each member of a set of propositions and also rationally believe of this set that at least one of its members is false. But this in turn pressures us to be unduly cautious. It pressures us to believe only those propositions that are certain or at least close to certain for us, since otherwise we are likely to have reasons to believe that at least one of these propositions is false. At first glance, the requirement that we avoid recognizable inconsistency seems little enough to ask in the name of rationality. It asks only that we avoid certain error. It turns out, however, that this is far too much to ask.

One might try to resist Foley's suggestion that inconsistency pressures us to believe only near certainties, while maintaining **Consistency** as a rational requirement (*e.g.*, Leitgeb has recently defended just such a view [35]). However, even if one has abandoned the extreme version of Lockeanism, there remains reason to think that **Consistency** is (still) too demanding, epistemically. Pettigrew [45] has recently shown — using the tools of epistemic utility theory — that agents who satisfy **Consistency** are exhibiting epistemically *risk-seeking* behavior. Stated intuitively, he shows that requiring consistency involves disproportionately weighting the epistemic best-cases scenarios over the epistemic worst-cases ones. So, in addition to rejecting the extreme version of Lockeanism which entails **Consistency**, we maintain that **Consistency** is *too strong* to serve as a universal rational requirement. We will return to this point as well as other applications of epistemic utility theory in the final section of this paper. However, first we will rehearse AGM's Gärdenfors postulates before exhaustively comparing them with Lockean revision.

3.3. **The Gärdenfors Postulates.** AGM theory can be presented in a variety of equivalent ways.[15] But, for our purposes, it will be most perspicuous to present the approach in terms of the axioms provided by the Gärdenfors postulates mentioned earlier in this section. Here are the six basic postulates of AGM.[16]

[15]Aside from the axiomatic presentation that we rely on, AGM is well-known to be equivalent to structures provided in terms of revision based on a "selection function" [1], a particular kind of entrenchment ordering [22], a Lewisian system of spheres [26], the rational consequence relation of non-monotonic logic [32], the probability one part of a Popper function [27], or in terms of *minimal change updating* [24, 50].

[16]As it turns out, the basic postulates (*1)–(*6) provide an axiomatization of *partial-meet revision* operators, which can be thought of as emerging from the minimally mutilating revision of some prior belief set $\mathbf{B}$ in accord with an entrenchment ordering on propositions. The addition of the supplementary postulates, (*7) and (*8), yields a characterization of a special class of partial meet revision operators whose entrenchment orderings are transitive. See [17] for an overview of the various ways of characterizing AGM belief revision operators. Finally, one can interpret these axioms more generally, in terms of a generalized entailment relation (which may be non-classical). For simplicity, we will assume a classical entailment relation here. What we say below can be generalized to non-classical (*e.g.*, substructural) entailment relations.

(*1) $\mathbf{B} * E = \mathsf{Cn}(\mathbf{B} * E)$ — Closure

In words, (*1) says that if an agent revises by E, then her posterior belief set $\mathbf{B}' = \mathbf{B} * E$ should satisfy AGM's standing synchronic requirement, **Closure**. As suggested earlier in this section, we take it that (*1) is grounded in AGM's *synchronic* requirements and does not independently impose any genuinely *diachronic* requirements.[17]

(*2) $E \in \mathbf{B} * E$ — Success

In words, (*2) says that if an agent revises by E, then E should be included in her posterior belief set $\mathbf{B}' = \mathbf{B} * E$. (*2) directly encodes **Conservitivity**'s constraint (ii).

(*3) $\mathbf{B} * E \subseteq \mathsf{Cn}(\mathbf{B} \cup \{E\})$ — Inclusion

In words, (*3) says that if an agent revises by E, then her posterior belief set $\mathbf{B}' = \mathbf{B} * E$ should contain *no more* propositions than $\mathsf{Cn}(\mathbf{B} \cup \{E\})$. In effect, this places an upper-bound on the agent's posterior ensuring that she does not adopt belief in propositions that are logically independent of her priors and the new evidence.

(*4) If $\mathbf{B} \nvDash \neg E$, then $\mathbf{B} \subseteq \mathbf{B} * E$ — Preservation[18]

In words, (*4) says that if an agent revises by E and E is consistent with her prior belief set $\mathbf{B}$, then her posterior belief set $\mathbf{B}' = \mathbf{B} * E$ should contain all of her prior beliefs. When the agent is revising by a proposition consistent with her priors, this places a lower-bound on her posterior and guarantees that she does not lose any beliefs as a result.

Note: (*1), (*3) and (*4) jointly imply that if an agent learns (exactly) some E that is consistent with their prior belief set $\mathbf{B}$, then their posterior belief set $\mathbf{B}' = \mathbf{B} * E$ should be *identical to* $\mathsf{Cn}(\mathbf{B} \cup \{E\})$.

(*5) If $E \nvDash \bot$, then $\mathbf{B} * E$ is consistent. — Consistency

[17]Consider the non-closed, but consistent (initial) belief set $\mathbf{B} = \{P, Q\}$, where P and Q are independent, contingent (atomic) claims. Closure implies that $\mathbf{B} * \top$ is closed. Thus, according to AGM, if an agent starts out with the prior belief set $\mathbf{B}$ and then revises by a tautology, they must (as a result of this "revision") come to believe the (contingent) *conjunction* $P \wedge Q$ (since, otherwise, the closure of $\mathbf{B} * \top$ will not be ensured). But, it is counter-intuitive that "learning" a tautology should provide an agent with a conclusive reason to accept a contingent claim. This drives home the point that AGM really needs to presuppose closure as a standing, synchronic constraint on all belief sets. See *fn.* 19 for a similar argument regarding consistency.

[18]Strictly speaking, Gärdenfors's principle (*4) was Vacuity. But, given the other postulates, Vacuity is equivalent to Preservation. And, since we've been working with various forms of Preservation already, it's easiest to stick with this (slightly non-standard) rendition of the Gärdenfors postulates.

In words, (*5) says that if an agent revises by E and E is not itself a contradictory proposition, then her posterior belief set $\mathbf{B}' = \mathbf{B} * E$ should satisfy AGM's synchronic **Consistency** principle. Much like (*1), we think that (*5) is really grounded in AGM's *synchronic* requirements and does not independently impose any diachronic requirements.[19]

(*6) If $\vdash X \equiv Y$, then $\mathbf{B} * X = \mathbf{B} * Y$ Extensionality

In words, (*6) says that if X and Y are *tautologically equivalent*, then updating on X should have *exactly the same effect as* updating on Y.

Each of these postulate places a restriction on which posteriors are admissible under AGM revision and, thereby, constrains the outputs of individual AGM revisions. It should be noted that the theory is often presented as also including the two supplementary postulates, (*7) and (*8), which generalize (*3) and (*4) respectively constraining *iterated* revisions.[20] However, since we are interested in the *diachronic* requirements governing single revisions rather than the *dispositional* ones governing iterated revisions (mentioned in *fn.* 14), we will focus exclusively on (*1)-(*6).

4. Convergences between Lockean Revision and AGM Revision

Now that we have presented the basics of these two approaches to belief revision, we will direct our attention to their relative behavior. In this section, we will begin our exploration by reporting some of the general convergences between Lockean revision and AGM revision.

4.1. ***Extremal* Lockean revision *is* AGM Revision.** In the case of the extremal Lockean threshold, where $t = 1$, our agent believes every proposition to which she assigns *maximal* credence. It is easy to see that this entails that the Lockean agent's prior and posterior belief sets $\mathbf{B}$ and $\mathbf{B}'$ will both satisfy **Cogency**. As a result, extremal Lockean revision must satisfy both **Closure** and **Consistency**. Furthermore, it has been known for some time that extremal Lockean agents must satisfy *all of the other AGM postulates as well.* To wit, we report the following classic theorem.

[19]Consider the closed, but inconsistent (initial) belief set $\mathbf{B} = \{P, \neg P, \top, \bot\}$, where P is a contingent (atomic) claim. Consistency implies that $\mathbf{B} * \top$ is consistent. Thus, according to AGM, if an agent starts out with the prior belief set $\mathbf{B}$ and then revises by a tautology, they must (as a result of this "revision") abandon either their belief in P or their belief in $\neg P$ (since, otherwise, the consistency of $\mathbf{B} * \top$ will not be ensured). But, it is counter-intuitive that "learning" a tautology should provide an agent with a conclusive reason to drop one of their contingent beliefs. This drives home the point that AGM theory really needs to presuppose consistency as a standing, synchronic constraint on all belief sets.

[20]For completeness we include statements of (*7) and (*8) below:

(*7) $\mathbf{B} * (X \wedge Y) \subseteq \mathsf{Cn}((\mathbf{B} * X) \cup \{Y\})$ Superexpansion

(*8) If $\mathbf{B} * E \nvdash \neg Y$, then $\mathbf{B} * (X \wedge Y) \supseteq \mathsf{Cn}((\mathbf{B} * X) \cup \{Y\})$ Subexpansion

Theorem (Gärdenfors [20]) Given a Lockean threshold of $t = 1$, for every E such that $b(E) > 0$, $\mathbf{B} \circledast E$ satisfies (*1)-(*6).[21]

The situation is much more interesting when our agent's Lockean threshold is *non*-extremal. As we will see, the relationship between Lockean revision and AGM revision in these cases is more nuanced.

4.2. **General convergences between Lockean revision and AGM Revision.** In addition to fully converging in the special case described above, the following three of AGM's postulates are satisfied by Lockean revision in full generality.

Proposition 4 Lockean revision *satisfies* Success. That is, the following is a theorem:

$$E \in \mathbf{B} \circledast E.$$

Proof. It is a theorem of probability calculus that $b(E \mid E) = 1$.[22] Therefore, $b(E \mid E) = 1 \geq t$, for *any* Lockean threshold t. So, $E \in \mathbf{B} \circledast E$. □

Proposition 5 Lockean revision *satisfies* Inclusion. That is, the following is a theorem:

$$\mathbf{B} \circledast E \subseteq \mathsf{Cn}(\mathbf{B} \cup \{E\}).$$

Proof. Suppose $X \in \mathbf{B} \circledast E$. Then, $b(X \mid E) \geq t$. And, it is a theorem of probability calculus that $b(E \supset X) \geq b(X \mid E)$. Therefore, $b(E \supset X) \geq t$. So, $E \supset X \in \mathbf{B}$. Hence, by *modus ponens* (for material implication), $X \in \mathsf{Cn}(\mathbf{B} \cup \{E\})$. □

Proposition 6 Lockean revision *satisfies* Extensionality. That is, the following is a theorem:

$$\text{If } \vdash X \equiv Y \text{, then } \mathbf{B} \circledast X = \mathbf{B}. \circledast Y$$

Proof. Suppose X and Y are tautologically equivalent. Then, X and Y are *probabilistically indistinguishable* (under *every* probability function). Therefore, Lockean revisions on X are indistinguishable from Lockean revisions on Y. □

These three positive results exhaust the set of AGM's postulates that Lockean revision is guaranteed (in full generality) to satisfy.

[21]In classical probability theory, conditionalization is undefined when the proposition that the agent conditionalizes on is assigned zero prior probability. For this reason, we must assume that our agents only learn things to which they assign non-zero credence. However, Gärdenfors's Theorem generalizes to accommodate such cases when the agent's credences are represented by Popper functions [27, 40]. Such generalizations allow for Bayesian style modeling of agents who learn propositions with zero credence.

[22]We assume that $b(E) > 0$ for *all* potential pieces of evidence — see *fn.* 21.

Lockean revision's satisfaction of Inclusion is of particular interest, since at first sight it may not have been so obvious that this would obtain. Since Lockean revision is driven by the Bayesian apparatus, we might have been inclined to think that there may be cases in which an agent may acquire a new belief in some proposition X, which is probabilistically (but not logically) dependent on the learned proposition E. However, this is not so. Inspecting the proof of Proposition 5 exposes that Lockeanism's synchronic requirements ensure that any time such a new belief is acquired, it might have equally well been acquired through *modus ponens*.[23] Towards the end of the next section, we will examine whether a convergence between Lockean revision's and AGM's *new* beliefs holds more generally. Ultimately, we will show that *sometimes* AGM will require the agent to form (strictly) *more new beliefs* than Lockean revision.

5. Divergences between Lockean revision and AGM Revision

In the general, non-extremal case, Lockean revision and AGM revision may *diverge significantly*. In this section, we explore this divergence and discuss the ways in which Lockean revision may violate the remaining three postulates: Closure, Consistency, and Preservation. As previously noted, Lockean revision's violation of Consistency and Closure has been widely discussed [3, 14, 9] and so they will only receive a cursory treatment. Instead, we will focus on the more interesting case of the possibility that Lockean revision may violate Preservation and include an instructive counter-example.

First, we know that Lockeanism, in general, does not require **Cogency**. Indeed, since Lockean revision is driven entirely by the Lockean apparatus, we see that it admits counter-examples to both **Closure** and **Consistency**. For brevity, we omit the proofs.

Proposition 7 Lockean revision *violates* **Closure** and **Consistency**. That is, it is possible that:

(1) $\mathbf{B} \circledast E \neq \mathsf{Cn}(\mathbf{B} \circledast E)$, and
(2) $E \nvdash \perp$ and $\mathbf{B} \circledast E$ is not consistent.

It has been known since the early 1960's [31] that non-extremal Lockean representability is compatible with failures of Consistency (*e.g.*, the lottery paradox).

[23] We are grateful to Konstantin Genin and Kevin Kelly for pointing out to us that, on the face of it, this fact suggests that Lockeanism is committed to deductivism about inductive inference. After all, you might think, if any proposition newly learned by a Lockean could have been learned by deduction using the new evidence and old beliefs, then it may seem that the inductive apparatus plays an inessential role in learning. However, we suspect that this inference is a bit too quick. As we will see shortly, acquiring new evidence can undermine an agent's old beliefs and, thus, render them unfit for use in such an inference. Nonetheless, the observation is interesting and warrants further exploration.

And, of course, if Consistency fails, then Closure must also fail (on pain of epistemic triviality). So, the well-known paradoxes of consistency will (inevitably) yield examples of *non*-extremal Lockean revision which violate both Consistency and Closure. For present purposes, we are not so interested in this well-known divergence between the two approaches [3, 14, 9]. Rather, we are more interested in cases where the agent's prior and posterior belief sets *satisfy* **Cogency**, but Lockean revision and AGM revision *still* manage to disagree.

This leads us to the central disagreement between the two approaches provided by Lockean revision's failure to generally satisfy AGM's characteristic postulate, Preservation. The counter-example provided in the proof of the next proposition highlights a deeper (and hitherto not fully understood) possible divergence between these two approaches.

Proposition 8 Lockean revision *can violate* Preservation. That is, it is possible that:

$$\mathbf{B} \nvdash \neg E \text{ and } \mathbf{B} \not\subseteq \mathbf{B} \circledast E$$

Proof. The proof strategy involves constructing a case in which Lockean revision recommends that an agent — whose priors are synchronically coherent by the lights of both Lockeanism and AGM — gives up one of her beliefs after revising by a proposition that is consistent with her prior belief set.

Suppose that the agent has a Lockean threshold $t = 0.85$. Then, the result is established in Table 1, which provides the distribution of the agent's credences over the algebra over the atomic sentences, E and X.

Given the Lockean threshold of 0.85 and the prior credence $b(\neg E \vee X) = 0.9$, it follows that (aside from the tautology) the Lockean agent will *only* have *one* belief: $\mathsf{B}(\neg E \vee X)$. However, learning E would leave her with the posterior credence $b(\neg E \vee X \mid E) = 2/3$. Thus, $\neg E \vee X \notin \mathbf{B} \circledast E$ even though E is consistent with $\mathbf{B}$ and $\mathsf{B}(\neg E \vee X)$. □

The proof above is more illustratively explained using a simple urn case. Suppose that we are tasked with taking a random sample from an urn containing a total of *ten* objects. The objects in the urn — as represented in Figure 3a — include *four* black circles, *three* black squares, *one* red square and *two* red circles.

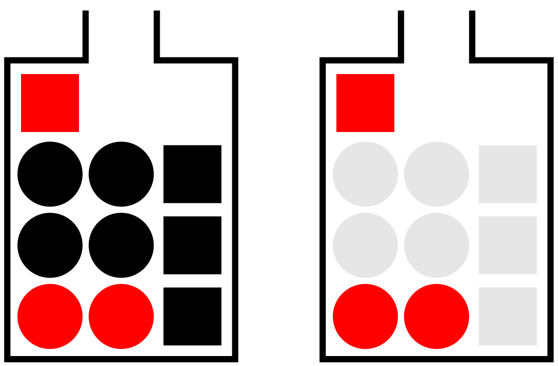
Let the atomic sentences E and X be assigned the following interpretations:

$$E \overset{\text{def}}{=} \text{‘The object sampled from the urn is red’, and}$$

$$X \overset{\text{def}}{=} \text{‘The object sampled from the urn is a circle’.}$$

φ	$b(\varphi)$	$b(\varphi \mid E)$	$\varphi \in \mathbf{B}$?	$\varphi \in \mathbf{B} \circledast E$?	
$E \wedge X$	$2/10$	$2/3$	No	No	
$E \wedge \neg X$	$1/10$	$1/3$	No	No	
$\neg E \wedge X$	$4/10$	0	No	No	
$\neg E \wedge \neg X$	$3/10$	0	No	No	
E	$3/10$	1	No	Yes	
X	$6/10$	$2/3$	No	No	
$E \equiv X$	$5/10$	$2/3$	No	No	
$E \not\equiv X$	$5/10$	$1/3$	No	No	
$\neg E$	$7/10$	0	No	No	
$\neg X$	$4/10$	$1/3$	No	No	
$E \vee X$	$7/10$	1	No	Yes	
$E \vee \neg X$	$6/10$	1	No	Yes	
$\neg E \vee X$	$9/10$	$2/3$	**Yes**	**No**	✓
$\neg E \vee \neg X$	$8/10$	$1/3$	No	No	

TABLE 1. Proof of Proposition 8



(A) Prior distribution (B) Posterior (given E)

FIGURE 3. Visualization of counter-example to Preservation for Lockean revision

Finally, assume that some Lockean agent knows the prior distribution of the objects and, as such, has credences in propositions about the shapes and colors of the objects in the urn that are calibrated to this distribution. In this case, the only proposition (aside from the tautology $\top$) that receives a credence higher than the Lockean threshold of 0.85 is $\neg E \vee X$. Thus, our agent's prior belief set will be the singleton $\mathbf{B} = \{\neg E \vee X\}$.

Now, suppose the agent learns only that the object drawn from the urn was red (*i.e.* she learns E). Upon conditionalizing on her new evidence, the agent's credence in the previously believed $\neg E \vee X$ will drop from 0.9 to $2/3$, as represented in Figure 3b. So, when the Lockean agent revises her beliefs, she will be led to give up her belief that $\neg E \vee X$ and only be left with belief in the newly learned E and its logical consequences, which are all assigned maximal credence. That is, after learning E, the agent's posterior belief set is $\mathbf{B}' = \mathbf{B} \circledast E = \{E, E \vee X, E \vee \neg X\}$.

To see how the described case serves as a counter-example to Preservation, observe that the following four crucial facts obtain:

- both the prior and posterior belief sets, $\mathbf{B}$ and $\mathbf{B} \circledast E$, satisfy **Cogency**;
- E is consistent with $\mathbf{B}$;
- $\neg E \vee X \in \mathbf{B}$; *but*,
- $\neg E \vee X \notin \mathbf{B} \circledast E$.

Each of these facts can be easily verified using Table 1. So, this simple case offers a demonstration that Preservation need not be satisfied by Lockean revision, even given **Cogency**.

In this case, learning E would seem to suffice for her to rationally infer X on the basis of her belief $\neg E \vee X$. In light of this, one may wonder why our Lockean agent is *precluded* from adopting belief in a deductive consequence of her prior beliefs and her new evidence. To see why, it will be illuminating to consider an analogy with the literature on *epistemic closure*. Hawthorne [29, p. 29] defends a closure principle which grants agents knowledge of the conclusions of their logical inferences *only if* they *retain knowledge of the premises throughout said inferences*. While Hawthorne's principle is aimed at the closure of *knowledge*, a similar "premise maintenance" caveat also seems reasonable for rational belief. In our example, the learned proposition, E, actually *serves as counter-evidence* to the (previously believed) second premise, $\neg E \vee X$, required for her to have inferred X. Thus, learning E serves to make a premise crucial for the inference no longer sufficiently likely to warrant belief.

It is important to note that in our counter-example both the agent's prior and posterior belief sets actually *satisfy* **Cogency**. As such, *this* disagreement between Lockean revision and AGM revision is *orthogonal* to the traditional disputes between "Bayesian" and "logical" schools of thought in formal epistemology, which have tended to fixate on their disagreement over **Cogency** [3, 14, 9]. In this sense, the Lockean counter-example to Preservation reveals a more fundamental disagreement between the diachronic requirements of the two approaches. At its core, this disagreement amounts to their differing on the question whether it is ever rational for an agent to give up belief in the face of non-definitive counter-evidence.

The AGM theorist answers in the negative, only allowing for beliefs to be given up when the agent has learned something *logically inconsistent* with her prior beliefs. Whereas the Lockean responds in the affirmative, permitting beliefs to be dropped when the learned proposition causes her prior beliefs to fall below the Lockean threshold.[24]

6. Lockeanism's golden threshold

It may be observed that our counter-example to Preservation provided above relies on the rather high Lockean threshold of 0.85. We have already observed in our discussion of Very Weak Preservation that it is satisfied by Lockean revision when the Lockean threshold is restricted to the range $(\Phi^{-1}, 1]$. We might then wonder whether similar results are available for Preservation. Interestingly, we are able to provide a result that both offers an affirmative answer to this question and also offers another avenue to appreciating the theoretical importance of the golden threshold.

Recall that, when **Cogency** is assumed, Lockean revision *satisfies* Weak Preservation. It immediately follows from this that Lockean revision will satisfy Very Weak Preservation assuming **Cogency** (and not only, as we have shown, when $t \geq \Phi^{-1}$). But, our counter-example from the previous section shows that **Cogency** *alone* is not sufficient for Lockean revision to satisfy Preservation. However, assuming *both* **Cogency** *and* that the Lockean threshold is greater than the golden threshold, we find that Lockean revision will actually satisfy Preservation. Specifically, we have the following theorem.

Theorem 1 Where $\mathbf{B}$ is deductively cogent, if there exists an E such that $\mathbf{B} \nvdash \neg E$, but $\mathbf{B} \not\subseteq \mathbf{B} \circledast E$, then $t \in [\Phi^{-1}, 1)$.

Proof. A proof is available in the Appendix. □

In other words, a *cogent* Lockean agent can violate Preservation only if her Lockean threshold is at least the golden threshold.

[24] While many are sympathetic to Preservation, the literature (primarily on epistemic conditionals) contains a number of arguments against the principle, *e.g.* in [19], [47], [37], and [10]. More recently (and more directly in the context of belief revision), Lin and Kelly [39] have independently argued against Preservation on the basis of their own broadly Bayesian account of revision. However, though motivated by similar considerations, their alternative remains distinct from the Lockean approach and is instead based on odds-ratio thresholds rather than the Lockean's conditional probability thresholds. Specifically, Lin and Kelly's revision procedure (LK revision) differs from Lockean revision in two respects: (a) LK revision is partition-sensitive, and (b) LK permits agents to believe propositions in which they have arbitrarily low credence. Ultimately, the underlying reason that LK revision deviates in these ways from Lockean revision derives from their adoption of **Cogency** as a universal requirement of rational belief. Though we differ over the ultimate standing of **Cogency**, we remain sympathetic to the objections that they provide against AGM. Nonetheless, we take our counter-examples to be more direct and probative in appreciating the fundamental issues.

Not only is this result surprising from a formal point of view, but it offers some philosophically important lessons as well. For one, the theorem provides a straightforward path to consistently endorsing the Lockean thesis in conjunction with AGM. As discussed earlier, it is well known that proponents of AGM may do so by treating their preferred theory as an account of belief revision under *certainty*.[25] In Lockean terms, proponents of AGM have previously viewed their theory as applying to cases in which $\mathsf{B}(X)$ iff $b(X) = 1$. But, by examining some of the consequences of Theorem 1, we can see that AGM theorists may regard their account as relevant to a broader class of situations. To fully appreciate this fact, note the following immediate corollary of our theorem.

Corollary Where $\mathbf{B}$ is deductively cogent, if $t \in (1/2, \Phi^{-1})$, then for every E, $\mathbf{B} \circledast E$ satisfies (*1)–(*6).

In other words, where $\mathbf{B}$ is deductively cogent and $t \in (1/2, \Phi^{-1})$, AGM and Lockean revision will be (qualitatively) equivalent. So, if proponents of AGM additionally accept a Lockean thesis *with a threshold* $t \in (1/2, \Phi^{-1})$, then they may reasonably take their theory to hold more generally without violating any Bayesian/Lockean intuitions. This is because the result establishes that (in the presence of deductive cogency) AGM will never diverge from a Lockean account *for thresholds in that interval.*

A second philosophically interesting consequence of Theorem 1 is that it provides Lockeans a rebuttal to the standard challenge involving the arbitrariness/contex dependence of Lockean thresholds *aside from* 1/2 and 1. This result (and our previous result regarding Very Weak Preservation) establishes an additional non-arbitrary (and context-independent) Lockean threshold at Φ^{-1}.

7. AGM is More Epistemically Risk-Seeking than Lockeanism

The results provided in the previous section do not exhaust the insights available from Lockeanism's failure to satisfy Preservation. Because Lockean revision satisfies Inclusion, it follows that that it can never require an agent to adopt *more* new beliefs than AGM revision would have generated. More precisely, because both approaches satisfy Inclusion, they both *rule out* posterior belief sets $\mathbf{B}' = \mathbf{B} * E$ that are *proper supersets* of $\mathsf{Cn}(\mathbf{B} \cup \{E\})$. In other words, neither approach will ever require an agent to be committed to new beliefs that *go beyond* the logical consequences of their prior belief set together with their new evidence.

Still further, it turns out that when the two approaches diverge, AGM will require the agent to have *strictly more* new beliefs than would be mandated by Lockean

[25] In fact, this motivation is explicitly offered by many AGM theorists including Gärdenfors himself in [21, p. 21].

revision. For example, recall our counter-example to Preservation from above (as provided in Table 1). There, we saw that Lockean revision requires the agent to give up belief in $\neg E \vee X$, while AGM revision does not. When E is learned, AGM revision will result in the agent believing X (along with a variety of other things), while Lockean revision will *preclude* the adoption of these new beliefs. But, this feature is not unique to our case. In fact, *whenever* Lockean revision and AGM revision disagree *in an interesting way* (*i.e.*, not as a result of failures of Closure or Consistency), the former will require the agent to adopt *strictly fewer* new beliefs than the latter. And, the converse holds as well. So, we arrive at our final theorem, which confirms that Lockean revision violates Preservation just in case performing an AGM revision by a proposition consistent with the agent's prior belief set leaves the agent with belief in all the things that she would have with the Lockean revision as well as some *additional* beliefs.

Theorem 2 Where $\mathbf{B}$ is deductively cogent, $\mathbf{B} \circledast E$ violates Preservation iff E is consistent with $\mathbf{B}$ and $\mathbf{B} \circledast E \subset \mathbf{B} * E$.

Proof. ($\Rightarrow$) Suppose $\mathbf{B}$ is deductively cogent and $\mathbf{B} \circledast E$ violates Preservation. Then, (a) E is consistent with $\mathbf{B}$; and, (b) $\mathbf{B} \nsubseteq \mathbf{B} \circledast E$. By (b), there is some $X \in \mathbf{B}$ such that $X \notin \mathbf{B} \circledast E$. It follows from (a), Preservation and Inclusion that $\mathbf{B} * E = \mathsf{Cn}(\mathbf{B} \cup \{E\})$. Therefore, $X \in \mathbf{B} * E$ and $X \notin \mathbf{B} \circledast E$. Finally, recalling from Proposition 5 that Lockean revision satisfies Inclusion, it follows that $\mathbf{B} \circledast E \subseteq \mathsf{Cn}(\mathbf{B} \cup \{E\}) = \mathbf{B} * E$. □

($\Leftarrow$) Suppose E is consistent with $\mathbf{B}$ and $\mathbf{B} \circledast E \subset \mathbf{B} * E$. Then, there exists an X such that $X \in \mathbf{B} * E$ but $X \notin \mathbf{B} \circledast E$. Because E is consistent with $\mathbf{B}$, Closure, Preservation, and Inclusion imply that $\mathbf{B} * E = \mathsf{Cn}(\mathbf{B} \cup \{E\})$. Therefore, $X \in \mathsf{Cn}(\mathbf{B} \cup \{E\})$, but $X \notin \mathbf{B} \circledast E$. Finally, once more appealing to Lockean revision's satisfaction of Inclusion, it follows that $X \in \mathbf{B}$. □

In other words, when Lockean revision and AGM revision (interestingly) diverge, AGM will be *more demanding* on an agent's posterior beliefs, since the Lockean agent's posterior will be a *strict subset* of the AGM agent's posterior. Because AGM will require agents to maintain beliefs in the face of non-definitive counter-evidence, it may be aptly viewed as an epistemically risk-seeking policy for belief revision. So, since Lockean revision will recommend that agents suspend belief in many cases when AGM revision recommends belief, it can be rightly viewed as the more epistemically risk-averse approach. In §3, we mentioned that Pettigrew uses tools from epistemic utility theory to argue that AGM's synchronic requirements are the more epistemically risk-seeking of the two. But, the same tools can be used in conjunction with our result from Theorem 1 to further argue that the

diachronic requirements of AGM imply an epistemically risk-seeking approach to belief revision.

To do so, we rely on Dorst's [7] representation theorem revealing that the Lockean thesis can be *derived* using epistemic utility theory. To see how Dorst's result works, we first equip our Bayesian agent with a (naïve) *epistemic utility function* for belief. Let $u(\mathsf{B}(X), w)$ refer to the *epistemic utility* of believing X in world w, and suppose that u is provided the following simple, piecewise definition:

$$u(\mathsf{B}(X), w) \stackrel{\text{def}}{=} \begin{cases} \mathfrak{r} & \text{if } X \text{ is true at } w \\ -\mathfrak{w} & \text{if } X \text{ is false at } w \end{cases}$$

That is, when an agents believes that X and X is true, then the utility function rewards her accuracy with the "epistemic credit" $\mathfrak{r}$; on the other hand, if her belief is false, then the utility function penalizes her with the "epistemic debit" $-\mathfrak{w}$.[26] The value yielded by this function is wholly determined by the truth of the agent's belief and is insensitive to its content and other considerations. So, the epistemic utility theory approach supposes a *veritistic* and *value monistic* account of the epistemic worth of beliefs. This treatment directly aligns with the those offered by so-called *accuracy-first epistemologists* and has been motivated on the Jamesian [30] grounds that belief has the simultaneous aims of attaining truth, while avoiding error. Accordingly, we take the epistemic utility of individual beliefs to contribute equally to the overall epistemic utility of an agent's total belief state.

For the moment, we impose only the following single constraint on the value ranges of the utility function's parameters $\mathfrak{r}$ and $\mathfrak{w}$:

$$\mathfrak{w} > \mathfrak{r} > 0.$$

The justification for these minimal restrictions is straightforward. If the epistemic benefit of believing a truth were not greater than zero, then there would never be incentive for belief over suspension. Of course, that would be an unwelcome result and so we have justification for the constraint that $\mathfrak{r} > 0$. A similar reason can be given to justify the restriction that the epistemic harm of false belief must be greater than the epistemic benefit of true belief. If they were the same, then it would be no better to suspend judgment on the outcome of a fair coin flip than to believe both that it will come up heads and that it will come up tails, since both would have an expected utility of zero. Finally, if the epistemic harm of false belief

[26]It is important to note that our treatment will assume that belief is the only qualitative attitude for which agents receive any epistemic utility. Accordingly, we treat suspension of belief as nothing more than lacking belief in both the proposition and its negation, and assign suspension of belief neither positive or negative value.

were *less* than the epistemic benefit of true belief, then suspension would actually have a worse expected value than the inconsistent alternative. Thus, $\mathfrak{w} > \mathfrak{r}$.

On the basis of this simple accuracy-centered utility function, it is straightforward to define the *expected epistemic utility* (EEU) for an agent's belief that X (relative to her credence function):

$$EEU(\mathsf{B}(X), b) \overset{\text{def}}{=} \sum_{w \in W} b(w) \cdot u(\mathsf{B}(X), w).$$

Then, we define the overall EEU of an agent's total belief set $\mathbf{B}$ as simply the sum of the EEUs of all of her individual beliefs.

$$EEU(\mathbf{B}, b) \overset{\text{def}}{=} \sum_{X \in \mathbf{B}} EEU(\mathsf{B}(X), b).$$

This basic apparatus is all that is needed to generate Dorst's theorem, which establishes that a belief set *maximizes* EEU relative to a credence function b just in case it satisfies the following precise (normative) Lockean thesis.[27]

Theorem (Dorst [7]) Where b is an agent's credence function, her belief set $\mathbf{B}$ *maximizes EEU* just in case

$$\begin{array}{ll} \text{if } \mathsf{B}(X), & \text{then } b(X) \geq \dfrac{\mathfrak{w}}{\mathfrak{r} + \mathfrak{w}}, \text{ and} \\ \text{if } b(X) > \frac{\mathfrak{w}}{\mathfrak{r} + \mathfrak{w}}, & \text{then } \mathsf{B}(X). \end{array}$$

It is straightforward to see that this implies that if an agent's beliefs and credences jointly satisfy the Lockean thesis, then she will maximize EEU. Thus, Lockean revision, as we have explored it, is entailed by the more general norm requiring that agents have belief sets that maximize EEU *at any given time*. Assuming conditionalization[28] as the rational procedure for credal update, the norm entails that Lockean revision is the unique procedure that will guarantee that agents maximize overall EEU.

The definition of the Lockean revision operator can now be equivalently restated with the aid of this new apparatus using the free-parameter values $\mathfrak{r}$ and $\mathfrak{w}$:

$$\mathbf{B} \circledast E = \left\{ X \;\middle|\; b(X \mid E) \geq \frac{\mathfrak{w}}{\mathfrak{r} + \mathfrak{w}} \right\}.$$

Now, notice that the greater $\mathfrak{w}$ is relative to $\mathfrak{r}$, the greater the resulting Lockean threshold will be. That is, the larger the debit incurred by the agent for believing a falsehood relative to the credit for her believing a truth, the larger the Lockean

[27] It is worth noting that a similar result is also proved (independently) in Easwaran [8], although Easwaran's applications of his result are much different than Dorst's. Historically, this method of deriving Lockean constraints traces back to the work of Hempel [28].

[28] Conditionalization can *itself* be given a justification using epistemic utility theory — *e.g.* see [25].

threshold will be. In the limiting case, we see that a maximal Lockean threshold is established by letting there be no benefit at all for believing truths (*i.e.* letting $\mathfrak{r} = 0$).

In the corollary to Theorem 1, we saw that AGM coheres with (cogent) Lockeanism when $t \in (1/2, \Phi^{-1}]$. Using these new tools from epistemic utility theory, we can see that this restriction on t is equivalent to requiring that $\mathfrak{w} \leq \Phi \cdot \mathfrak{r}$. Thus, the adoption of AGM can be seen as placing more weight on the epistemically best-case scenarios, which corresponds to a kind of epistemically risk-seeking behavior.[29] So, not only are the *synchronic* requirements of AGM more epistemically risk-seeking than those of Lockeanism — as argued by Pettigrew [45] — but its *diachronic* requirements are as well.

8. Conclusion and Future Work

We have pinpointed the precise ways in which a (broadly Bayesian) Lockean approach to belief revision agrees (and disagrees) with the more traditional AGM theory. Setting aside issues surrounding **Cogency**, Lockean revision and AGM revision exhibit a surprising degree of convergence. Our analysis reveals that, holding **Cogency** fixed, the two approaches to belief revision disagree *only* regarding the universal validity of Preservation. Intuitively, this simply results from the fact that Lockean revision is sensitive to non-definitive counter-evidence, while AGM revision is not.

In this paper, we have chosen to focus on the *diachronic* coherence requirements of the two theories. However, there are further issues relating to their *dispositional* coherence requirements, which govern iterated revision. Critics of AGM have often complained that it does not easily generalize to offering an account of how iterated revision should proceed. On the other hand, Lockean revision has no special problem with iterated revision.[30] In future work, we plan to compare Lockean revision to other systems of belief revision beyond AGM. In doing so, it will be of particular interest to consider systems whose specific aim is the accommodation of iterated revision. More specifically, we plan to investigate the dispositional norms of Lockean revision as contrasted with the Darwiche and Pearl postulates for iterated revision [4].

Another interesting next step in the exploration of Bayesian qualitative revision is to investigate how Lockean revision changes when the agent's credence function is a non-classical probability function (thus permitting conditionalization

[29] Lockeanism's risk-aversion (in this sense) should not be wholly surprising, since it is driven by the expected utility calculus *via* a *concave* utility function u. Nonetheless, it is interesting to see that, from multiple perspectives, non-extremal Lockean revision is more risk-averse than AGM.

[30] That said, insofar as we have relied on conditionalization to define $\circledast$, there is a problem with conditionalizing on any proposition assigned a prior probability of 0 (*fn.* 21).

by propositions with zero unconditional credence) or when combined with other, more general, credal update procedures in place of conditionalization.[31] One especially interesting application along these lines would be to the problem of explicating a Bayesian notion of *contraction*. We have some preliminary ideas about "Bayesian contraction," which we plan to explore in a sequel to this paper.[32]

Finally, we would (ideally) like to have a *purely qualitative axiomatization* of the Lockean revision operator. Some progress toward such an axiomatization has recently been made — *e.g.* see [40]. Of particular interest, van Eijck and Renne [12] recently provided an axiomatization for the modal logic of belief with a Lockean threshold of $1/2$. We think that our results involving the non-arbitrary Lockean threshold at Φ^{-1} suggest that a fruitful next step may be to investigate the logic of belief satisfying this threshold. However, there remains significant theoretical work to be done in order to determine precisely which axioms would be needed to characterize Lockean revision.

9. Epilogue: A (Third) Approach to Belief Revision

Those well versed in the recent literature may wonder how our results relate to Leitgeb's [33] results concerning his *stability theory of belief*. The fundamental synchronic requirement of Leitgeb's stability theory is the *Humean thesis*, which he proves to be equivalent to jointly assuming the Lockean thesis in conjunction with probabilism and **Cogency** for belief. In addition to his synchronic requirement, Leitgeb develops an account of belief revision[33] for his Humean agents that ultimately satisfies *all* of the Gärdenfors postulates, (*1)-(*6). But, how does this square with our results, which seem to drive a wedge between Lockean revision for cogent agents and AGM revision? In this brief epilogue, we will directly attend to this tension and establish a result for Humean revision, which may appear problematic.

Like Lockeanism, the stability theory provides an account of the joint coherence requirements governing beliefs and credences. The credences of Humean agents

[31]As we mentioned in the introduction, all of the results we reported here will continue to hold for any mechanical/minimal change Bayesian credal update procedure that satisfies the following three constraints: (i) $b'(E) > b(E)$, (ii) $b'(E) \geq t$ (where t is the agent's Lockean threshold), and (iii) $b(E \supset X) \geq b'(X)$. It would be nice to explore these (and other) non-standard Bayesian updating procedures in more depth (especially, in conjunction with Lockeanism).

[32]The basic idea behind our approach to "contracting a Bayesian belief set **B** on proposition E" would involve (a) defining b' as *the closest probability function to* b *such that* $b'(E) \leq t$, and then (b) checking which propositions X are such that $b'(X) > t$. The set $\mathbf{B} \div E \stackrel{\text{def}}{=} \{X \mid b'(X) > t\}$ would be our (initial) explication of what it means to "contract a Bayesian agent's belief set **B** on proposition E."

[33]Although Leitgeb actually discusses these requirements as requirements on *conditional belief* — *viz.* belief *given* some proposition — it is unproblematic to translate his account of conditional belief into an account of belief revision. To remain consistent in our notation, we will explain his account in the latter terms; however, nothing rests on this. For a helpful over of conditional belief, see [11].

satisfy all the same norms as those of Lockean agents. So, they are taken to taken to have a probabilistically coherent prior, $b(\cdot)$, which is updated *via* conditionalization yielding the posterior $b'(\cdot) = b(\cdot \mid E)$. Also like Lockean agents, Humean agents are equipped with a set of qualitative beliefs, **B**, which satisfies the Lockean thesis relative to her credences. These agents update their belief set, **B**, when learning some proposition, E, *via* his novel *Humean revision* operator ($\circ$) to generate the posterior $\mathbf{B}' = \mathbf{B} \circ E$. However, unlike Lockeanism in general, the stability theory imposes *Cogency* as a synchronic requirement on belief.

Leitgeb's most general (and notable) result is a representation theorem showing that these synchronic requirements — **Cogency**, probabilism, and the Lockean thesis — are equivalent to the single requirement imposed by what he calls *Humean thesis*. This requirement says that an agent should believe all and only those propositions to which she assigns "resiliently high" credence. More precisely, the Humean thesis requires agents to believe all and only those propositions that they would continue regard as sufficiently likely were they to learn any proposition logically consistent with their prior belief set. A simple and more careful formulation of the Humean thesis is provided below:

(HT^r) $\mathsf{B}(X)$ iff $b(X \mid Y) > r$ for all Y such that $\neg\mathsf{B}(\neg Y)$ and $b(Y) > 0$.

As with the Lockean thesis, the Humean thesis relies on the probability threshold, r, to capture the notion of sufficient likelihood provided in the intuitive statement above. With the aid of this definition, we can now provide a precise statement of Leitgeb's central representation theorem:

Theorem (Leitgeb [33]) Where **B** is an agent's belief set and b is her credence function, the agent's beliefs and credences satisfy (HT^r) for some r iff

(1) **B** is deductively cogent,
(2) b is a probability function, and
(3) **B** and b satisfy the Lockean thesis for some t.

Thus, (HT^r) is equivalent to requiring **Cogency**, probabilism for credences, and the Lockean thesis for some t.[34] So, the stability theory offers a univocal and well-defined account that accommodates both the Lockean thesis and **Cogency**.

In addition these synchronic coherence requirements, Leitgeb [33, Ch. 4] offers diachronic requirements based on similar motivations. These requirements serve to characterize Humean revision, which ultimately satisfies AGM's (*1)-(*6). The first two principles required are the following bridge principles:

[34]The reader is invited to note at this point that the Lockean threshold, t, is distinct from the *Humean threshold*, r, found in (HT^r). Although the t and r may converge, some pairs of credences and beliefs will satisfy Lockean thresholds significantly greater than the greatest Humean threshold that they satisfied, as we will see shortly.

(°BP1^r) If $\mathbf{B} \nvdash \neg E$ and $b(E) > 0$, then $X \in \mathbf{B} \circ E$ only if $b(X \mid E) > r$

(°BP2) $\mathbf{B} \circ E$ is inconsistent iff $b(E) = 0$

The first principle, (°BP1^r), requires that the left-to-right direction of the Lockean thesis is satisfied relative to propositions consistent with the agent's prior beliefs. The second, (°BP2), requires that Humean revision treats every proposition that the agent assigns zero credence the same way as it treats *logically* impossible propositions.

In addition to (°BP1) and (°BP2), Humean revision requires that the following AGM-like axioms are satisfied.

(°1)	$X \in \mathbf{B} \circ X$	Reflexivity
(°2)	If $Y \in \mathbf{B} \circ X$ and $Y \vdash Z$, then $Z \in \mathbf{B} \circ X$	Single Premise Closure
(°3)	If $Y \in \mathbf{B} \circ X$ and $Z \in \mathbf{B} \circ X$, then $Y \wedge Z \in \mathbf{B} \circ X$	Finite Conjunction
(°4)	For any $\mathcal{Y} = \{Y \mid Y \in \mathbf{B} \circ X\}$, $\bigwedge \mathcal{Y} \in \mathbf{B} \circ X$	General Conjunction
(°5)	$\mathbf{B} \circ \top \nvdash \bot$	Consistency
(°6)	If $\mathbf{B} \circ X \nvdash \neg Y$, then $\mathbf{B} \circ (X \wedge Y) = \mathsf{Cn}(\mathbf{B} \circ X \cup \{Y\})$	General Revision

In conjunction with (°BP1^r) and (°BP2), these requirements suffice to guarantee that Humean revision satisfies *all* of AGM's postulates — including (∗7) and (∗8). The inclusion of General Revision,[35] reveals that Humean revision is constructed so as to eliminate the possibility that ∘ will violate Preservation.

Not only does Humean revision satisfy the axioms of AGM, but it is guaranteed to yield posterior states satisfying the (HT^r). In order to explain this result, we first define *P-stability^r*, which is similar to (HT^r), but applies to individual propositions rather than belief sets.

Definition (P-stability^r) A proposition, X, is *P-stable^r* iff $b(X \mid Y) > r$, for any Y such that $X \nvdash \neg Y$ and $b(Y) > 0$

Clearly, if a belief set $\mathbf{B}$ satisfies HT^r, then the strongest proposition in $\mathbf{B}$ will be P-stable^r. With this new notion in hand, we may now state Leitgeb's representation theorem (modified only to cohere with the current notational conventions).

Theorem (Leitgeb [33, p. 206]) Provided a deductively cogent belief set $\mathbf{B}$ and a probabilistic credence function b, the revision operator ∘ satisfies °BP1^r, °BP2, and (∘1) – (∘6) relative to $\mathbf{B}$ and b iff *there exists a class $\mathcal{X}$* of non-empty P-stable^r propositions such that:

- $\mathcal{X}$ contains the least set of probability 1 in the algebra,

[35]It is easily seen that General Revision (in conjunction of the other principles) implies Preservation. First, note that ∘ satisfies Idempotence since ∘ satisfies the AGM axioms. Then, observe that General Revision directly implies Preservation when we let X be ⊤ and weaken its consequent.

- all other members of $\mathcal{X}$ have probability less than 1,
- for any Y, such that $b(Y) > 0$, if X is the strongest proposition in $\mathcal{X}$ such that $Y \cap X \neq \emptyset$, then for all Z:

$$Z \in \mathbf{B} \circ Y \text{ iff } Z \supseteq Y \cap X,$$

  and
- for all Y, if $b(Y) = 0$, then $\mathbf{B} \circ Y$ is inconsistent.

Intuitively, the result establishes an equivalence between the satisfaction of his principles and the *existence* of some P-stable^r set. Leitgeb suggests that the right-to-left direction offers the benefit of providing a recipe for building models for his revision postulates by finding some P-stable^r set $\mathcal{X}$ and imposing the restrictions listed.

At first pass, this might seem to suggest that an agent whose prior belief set $\mathbf{B}$ satisfies (HT^r) and revises by E in such a way that satisfies the right side of the theorem relative to a set $\mathcal{X}$ whose members are P-stable^r, her revision must be representable by $\mathbf{B} \circ E = \mathbf{B} * E$. Nonetheless, we will demonstrate that this need not be so. A reconsideration of our counter-example from the proof of Proposition 8 will demonstrate that the agent's Lockean revision also satisfies these conditions as well despite violating Preservation. Recall the probability distribution across the strongest worlds:

w	$b(w)$	$b(w \mid E)$
$E \wedge X$	$2/10$	$2/3$
$E \wedge \neg X$	$1/10$	$1/3$
$\neg E \wedge X$	$4/10$	0
$\neg E \wedge \neg X$	$3/10$	0

TABLE 2. Counter-example to Preservation for Leitgeb's stability theory

As we saw earlier, the agent's belief set, $\mathbf{B}$, is deductively closed and satisfies the Lockean thesis for $0.8 < t < 0.9$. However, a close examination of the distribution confirms that the prior belief set $\mathbf{B} = \{\neg E \vee X\}$ also (uniquely) satisfies (HT$^{.65}$).

Now, we compare the two diverging recommendations the stability theory and Lockean revision and show that both are revisions that lead to posteriors including only P-stable$^{.65}$ propositions. First, consider Leitgeb's revision:

$$\mathbf{B} \circ E = \{E \wedge X, E, H, E \vee H, E \vee \neg H, \neg E \vee H\}.$$

In this case, the class of P-stable$^{.65}$ propositions used to generate the posterior was $\mathcal{X} = \{\neg E \vee X, \top\}$. Now recall the Lockean revision from earlier, which generated the following posterior.

$$\mathbf{B} \circledast E = \{E, E \vee H, E \vee \neg H\}$$

Notice that this is the revision that would follow from the right side of the theorem if we choose $\mathcal{X}' = \{\top\}$ as the class of P-stable[.65] propositions. Not only does this satisfy the required conditions, but $\mathbf{B} \circledast E$ also satisfies °BP1, °BP2, and $(\circ 1) - (\circ 5)$. The Lockean revision only violates $(\circ 6)$ (which is, of course, just the conjunction of the generalizations of Inclusion and Preservation).

Naturally, this observation does not show that Leitgeb's theorem was mistaken. After all, his theorem merely required that there is *some* P-stable^r set that can be used to construct an AGM revision that satisfies his bridge principles. Indeed, there is *some* such set (as demonstrated above). But, it does show that further information is required to determine *which* class of P-stable^r sets is the appropriate one.

Appendix: Proof of Theorem 1

Theorem 1 Where **B** is deductively cogent, if there exists an E such that $\mathbf{B} \nvdash \neg E$, but $\mathbf{B} \not\subseteq \mathbf{B} \circledast E$, then $t \in [\Phi^{-1}, 1)$.

Proof. It is well-known that if a Lockean agent's threshold is *extremal* (*i.e.*, *equal to 1*), then such an agent will (a) have deductively cogent beliefs, and (b) satisfy Preservation. This explains why the interval $[\Phi^{-1}, 1)$ is open on the right. The interesting (and new) case involves the left-hand (closed) side of the interval. From Proposition 1, we know that assuming **Cogency**, Lockean revision satisfies *Weak Preservation.* Therefore, cogent Lockean agents will *never* lose beliefs by revising by a previously believed proposition. The remaining cases are handled by Lemmas 2 and 3 below. Together, these results suffice to demonstrate that cogent Lockean Preservation violators (like our agent above) must have Lockean thresholds of *at least* Φ^{-1}. □

Lemma 2 Let $\langle W, b\rangle$ be a (finite) probability space, over a set of possible worlds W, and suppose that the three propositions $\{X, Y, Z\}$ form a *partition* of W.[36] Then, the following four conditions:

$$b(X \vee Y) \geq t \tag{1}$$

$$b(X \vee Z) < t \tag{2}$$

$$b(Y \vee Z) < t \tag{3}$$

$$b(X \mid X \vee Z) < t \tag{4}$$

[36]That is, the disjunction $X \vee Y \vee Z$ is a tautology (*i.e.*, a proposition that's true in all members of W), and each pair from $\{X, Y, Z\}$ is a contradiction (*i.e.*, a proposition that's false in all members of W).

jointly entail

$$t \geq \frac{\sqrt{5}-1}{2} = \Phi^{-1} \tag{5}$$

Proof. Suppose conditions (1)–(4) hold. We adopt the following abbreviations.

$$x \overset{\text{def}}{=} b(X)$$

$$y \overset{\text{def}}{=} b(Y)$$

$$z \overset{\text{def}}{=} b(Z)$$

Conditions (1)–(4) then may be re-written as follows

$$x + y \geq t \tag{1}$$

$$x + z < t \tag{2}$$

$$y + z < t \tag{3}$$

$$\frac{x}{x+z} < t \tag{4}$$

Because $\{X, Y, Z\}$ form a partition of W, it follows from probability calculus that $x, y, z \in [0, 1]$ and $z = 1 - x - y$. Substitution for z in (1)–(4) then yields

$$x + y \geq t$$

$$1 - y < t$$

$$1 - x < t$$

$$\frac{x}{1-y} < t$$

It is then simply a matter of elementary algebraic reasoning to derive that

$$t^2 + t - 1 \geq 0$$

from which it follows (since $t \geq 0$) that

$$t \geq \frac{\sqrt{5}-1}{2} = \Phi^{-1} \tag{5}$$

which completes the proof of Lemma 2. □

Lemma 3 Let $\langle W, b\rangle$ be a (finite) probability space, over a set of possible worlds W, and let $t \geq 0$ be a Lockean agent's Lockean threshold. Suppose that the agent's prior belief set **B** is deductively cogent, and let E be some proposition consistent with **B**, and which is not already contained in **B**. If such an agent's Lockean-revised belief set $\mathbf{B} \circledast E$ is deductively cogent, but yields a violation of Preservation, then there exists a partition $\{X, Y, Z\}$ of W such that the following four conditions are met:

(1) $$b(X \vee Y) \geq t$$

(2) $$b(X \vee Z) < t$$

(3) $$b(Y \vee Z) < t$$

(4) $$b(X \mid X \vee Z) < t$$

Proof. Because **B** is cogent, there is a proposition C (the conjunction of all the members of **B**) which is a member of **B** and which entails every member of **B**. Let

$$X \overset{\text{def}}{=} E \wedge C$$

$$Y \overset{\text{def}}{=} \neg E \wedge C$$

$$Z \overset{\text{def}}{=} \neg C$$

Clearly, $\{X, Y, Z\}$ constitute a partition of W. So, we just need to show that $\{X, Y, Z\}$ satisfy conditions (1)–(4) above.

(1) $b(X \vee Y) \geq t$. To see this, note that $X \vee Y$ is logically equivalent to C, and $C \in \mathbf{B}$ by hypothesis.

(2) $b(X \vee Z) < t$. To see this, note that $\vdash X \vee Z \equiv C \supset E$. Suppose for *reductio* that $b(X \vee Z) \geq t$. Then, $C \supset E \in \mathbf{B}$, since $b(C \supset E) \geq t$. But, since $C \in \mathbf{B}$ and **B** is cogent, this would imply (*via modus ponens*) that $E \in \mathbf{B}$, which contradicts one of the preconditions of the lemma (that $E \notin \mathbf{B}$).

(3) $b(Y \vee Z) < t$. To see this, note that $Y \vee Z$ is logically equivalent to $C \supset \neg E$. Suppose for *reductio* that $b(Y \vee Z) \geq t$. Then, $C \supset \neg E \in \mathbf{B}$, since $b(C \supset \neg E) \geq t$. But, since $C \in \mathbf{B}$ and **B** is cogent, this would imply (*via modus ponens*) that $\neg E \in \mathbf{B}$, which contradicts one of the preconditions of the lemma (that $\mathbf{B} \nvDash \neg E$).

(4) $b(X \mid X \vee Z) < t$. To see this, note that

$$\begin{aligned} b(X \mid X \vee Z) &= \frac{b(X \wedge (X \vee Z))}{b(X \vee Z)} \\ &= \frac{b(X)}{b(X \vee Z)} \\ &= \frac{b(E \wedge C)}{b(\neg C \vee E)} \\ &\leq \frac{b(C \wedge E)}{b(E)} \end{aligned}$$

Thus, $b(X \mid X \vee Z) \leq b(C \mid E)$. Moreover, $b(C \mid E) < t$. To see this, suppose (for *reductio*) that $b(C \mid E) \geq t$. Then, $C \in \mathbf{B} \divideontimes E$. But, this implies that *both C and E* are members of $\mathbf{B} \divideontimes E$. And, by the cogency of $\mathbf{B} \divideontimes E$, this implies $\mathbf{B} \divideontimes E \supseteq \mathsf{Cn}(\mathbf{B} \cup \{E\})$, which contradicts one of the preconditions of the lemma (that this is a counter-example to Preservation). Therefore, $b(X \mid X \vee Z) \leq b(C \mid E)$ and $b(C \mid E) < t$, which implies $b(X \mid X \vee Z) < t$.

This completes the proof of Lemma 3. □

References

[1] Carlos Alchourron, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *The Journal of Symbolic Logic*, 50(2): 510–530, June 1985.

[2] Horacio Arló Costa. Conditionals and Monotonic Belief Revisions: the Success Postulate. *Studia Logica*, 49: 557–566, 1990.

[3] David Christensen. Putting Logic in its Place. Oxford University Press, 2004.

[4] Adnan Darwiche and Judea Pearl. On the Logic of Iterated Belief Revision *Artificial Intelligence*, 89: 1–29, 1996.

[5] Persi Diaconis and Sandy Zabell. Updating subjective probability *Journal of the American Statistical Association* 77: 822-830, 1982.

[6] Dylan Dodd. Belief and Certainty. *Synthese*, forthcoming.

[7] Kevin Dorst. An Epistemic Utility Argument for the Threshold View of Outright Belief. Manuscript, 2014.

[8] Kenny Easwaran. Dr. Truthlove, or how I learned to stop worrying and love Bayesian probabilifies. Noûs, Forthcoming.

[9] Kenny Easwaran and Branden Fitelson. Accuracy, Coherence, and Evidence. *Oxford Studies in Epistemology*, Vol. 5, 61–96. Oxford University Press, 2015.

[10] Horacio Arló-Costa. Belief Revision Conditionals: Basic Iterated Systems. *Annals of Pure and Applied Logic*. 96, 3–28, 1999.

[11] Dorothy Edgington. On Conditionals. *Mind*, 104(414): 235–329, 1995.

[12] Jan van Eijck and Bryan Renne. Belief as Willingness to Bet. Manuscript, 2014.

[13] Ronald Fagin, Jeffrey Ullman, and Moshe Vardi. On the semantics of updates in databases. *Proceedings of the Second ACM SIGACT-SIGMOD*, 352–365, 1983.

[14] Richard Foley. Working Without a Net. Oxford University Press, 1992.

[15] Peter Gärdenfors. Belief Revisions and the Ramsey Test for Conditionals. *The Philosophical Review*, 95(1): 81–93, 1986.

[16] Peter Gärdenfors, Sten Lindström, Michael Morreau, and Wlodzimierz Rabinowicz. The negative Ramsey test: Another triviality result. In André Fuhrmann and Michael Morreau (eds.), *The Logic of Theory Change*, pages 127–134. Springer, 1991.

[17] Peter Gärdenfors and Hans Rott. Belief revision. *Handbook of logic in artificial intelligence and logic programming*, Vol. 4, 35–132. Oxford University Press, 1995.

[18] Peter Gärdenfors. Conditionals and Changes of Belief. *Acta Philosophica Fennica*, 30: 381–404, 1978.

[19] Peter Gärdenfors. Belief Revisions and the Ramsey Test for Conditionals. *The Philosophical Review*, 95(1): 81–93, 1986.

[20] Peter Gärdenfors. The dynamics of belief: Contractions and revisions of probability functions. *Topoi*, 5: 29–37, 1986.

[21] Peter Gärdenfors. *Knowledge in Flux*. MIT Press, 1988.

[22] Peter Gärdenfors and David Makinson. Revision of Knowledge Systems Using Epistemic Entrenchment. *TARK '88 Proceedings of the Second Conference of Theoretical Aspects of Reasoning about Knowledge*, 83–95, 1988.

[23] Konstantin Genin. Jeffrey Conditioning and the Inclusion Principle. Manuscript, 2017

[24] Konstantinos Georgatos. Geodesic revision. *Journal of Logic and Computation* 19(3): 447–459, 2009.

[25] Hilary Greaves and David Wallace. Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility. *Mind*, 115(459): 607–632, 2006.

[26] Adam Grove. Two Modellings for Theory Change. *Journal of Philosophical Logic*, 17(2): 157–170, 1988.

[27] William Harper. Rational belief change, popper functions and counterfactuals. *Synthese*, 30(1-2): 221–262, 1975.

[28] Carl Hempel. Deductive-nomological vs. statistical explanation. *Minnesota studies in the philosophy of science* 3: 98–169, 1962.

[29] John Hawthorne. The Case for Closure In M. Steup and E. Sosa (eds.), *Contemporary Debates in Epistemology*, 26–42. Blackwell, 2005.

[30] William James. The will to believe. *The New World*, 5: 327–347, 1896.

[31] Henry Kyburg. Probability and The Logic of Rational Belief. Wesleyan University Press, 1961.

[32] Daniel Lehmann and Menachem Magidor. What does a conditional knowledge base entail?. *Journal of Artificial Intelligence*, 44(1): 167–207.

[33] Hannes Leitgeb. *Stability Theory of Belief*. Oxford University Press, 2016.

[34] Hannes Leitgeb. The Stability Theory of Belief. *Philosophical Review*, 123(2): 131–171, 2014.

[35] Hannes Leitgeb. The review paradox: On the diachronic costs of not closing rational belief under conjunction. *Noûs*, 78(4): 781–793, 2013.

[36] Isaac Levi. *Gambling with Truth: An Essay on Induction and the Aims of Science*. MIT Press, 1973.

[37] Isaac Levi. *For the Sake of Argument*. Cambridge University Press, 1996.

[38] Isaac Levi. *Is a miss as good as a mile?* In S. Pihlström, P. Raatikainen, and M. Sintonen (eds.), *Approaching truth: Essays in honour of Ilkka Niiniluoto*, 209–223. College Publications, 2007.

[39] Hanti Lin and Kevin Kelly. Propositional Reasoning that Tracks Probabilistic Reasoning. *Journal of Philosophical Logic*, 41(6): 957-981, 2012.

[40] David Makinson and James Hawthorne. Lossy Inference Rules and their Bounds: A Brief Review. *The Road to Universal Logic: Festschrift for 50th Birthday of Jean-Yves Béziau*, Vol. 1, 385–407, Springer, 2015.

[41] Ilkka Niiniluoto. *Truthlikeness*. D. Reidel, 1987.

[42] Ilkka Niiniluoto. Belief Revision and Truthlikeness. In Sören Halldén, Bengt Hansson, Wlodek Rabinowicz, Nils-Eric Sahlin (eds.) *Spinning Ideas: Electronic Essays Dedicated to Peter Gärdenfors on His Fiftieth Birthday*, 1999.

[43] Graham Oddie. *Likeness to the Truth*. D. Reidel, 1986.

[44] Richard Pettigrew. Epistemic Utility Arguments for Probabilism. *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition), E. Zalta (ed.), URL = <`http://plato.stanford.edu/archives/spr2016/entries/epistemic-utility/`>.

[45] Richard Pettigrew. Jamesian Epistemology Formalised: An Explication of 'The Will to Believe'. *Episteme*, 13(3): 253–268, 2016.

[46] Alexander Pruss. The badness of being certain of a falsehood is at least $\frac{1}{\log(4)-1}$ times greater than the value of being certain of a truth. *Logos & Episteme*, 3(2): 229–238, 2012.

[47] Wlodek Rabinowicz. Stable Revision, or Is Preservation Worth Preserving?. *Logic, Action and Information: Essays on Logic in Philosophy and Artificial Intelligence*, A. Fuhrmann and H. Rott (*eds.*), 101–128, 1995.

[48] Hans Rott. Coherence and Conservatism in the Dynamics of Belief Part I: Finding the Right Framework. *Erkenntnis*, 50: 387–412, 1999.

[49] Hans Rott. *Change, choice, and inference*. Oxford University Press, 2001.

[50] Odinaldo Rodrigues, Dov Gabbay and Alessandra Russo. Belief Revision. *Handbook of Philosophical Logic v.16*, D. Gabbay and F. Guenthner (*eds.*), 1–114, 2011.

[51] Robert Stalnaker. Iterated belief revision. *Erkenntnis*, 70(2): 189–209, 2009.

[52] Florian Steinberger. Explosion and the Normativity of Logic. *Mind*, 125(498): 385–419, 2016.