Daniel Rothschild February 13, 2020

LOCKEAN BELIEFS, DUTCH BOOKS, AND SCORING SYSTEMS (DRAFT)

Abstract

On the Lockean thesis one ought to believe a proposition if and only if one assigns it a credence at or above a threshold (Foley 1992). The Lockean thesis, thus, provides a way of characterizing sets of all-or-nothing beliefs. Here we give two independent characterizations of the sets of beliefs satisfying the Lockean thesis. One is in terms of betting dispositions associated with full beliefs and one is in terms of an accuracy scoring system for full beliefs. These characterizations are parallel to, but not merely derivative from, the more familiar Dutch Book (de Finetti 1974) and accuracy (Joyce 1998) arguments for probabilism.

1 INTRODUCTION

The Lockean thesis is a way of connecting all-or-nothing beliefs to graded beliefs (Foley 1992). On the Lockean thesis to have an all-or-nothing belief in a proposition is just to assign that proposition a credence at or above a certain threshold.[1] One of the interesting and controversial features of Lockeanism is that it does not require that one's beliefs either be closed under logical implication or even be logically consistent. For a failure of closure note that at any threshold less than one it is possible, on the Lockean thesis, to believe two atomic propositions, P, Q, while not believing their conjunction, $P\&Q$.[2] For a failure of consistency note that at a threshold of .6 it is possible to believe two propositions P and Q, while also believing the negation of their conjunction, $\neg(P\&Q)$.[3]

So, the Lockean thesis (at thresholds less than one) does not require that one's beliefs be either logically closed or logically consistent. We might want to ask what constraints the Lockean thesis *does* put on one's beliefs. Some constraints are well-known: for example, if the threshold for belief is greater than .5, then,

I am grateful to Gary Chamberlain, Nilanjan Das, Kevin Dorst, Kenny Easwaran, Branden Fitelson, Wes Holliday, Matt Mandelkern, Richard Pettigrew and Linda Rothschild for comments and discussion. Special thanks to two anonymous referees at *Review of Symbolic Logic*.

[1] Sometimes the Lockean thesis is framed as one about what one ought to believe, rather than about what one does believe. The difference is not essential here.

[2] See Leitgeb (2014) for an exploration of constraints, other than having a threshold of 1, on credences and thresholds that ensure closure (and consistency).

[3] Assume you assign a credence of .6 to P and .6 to Q, then you can assign at most .8 to $\neg(P\&Q)$.

on the Lockean thesis, one cannot hold pairwise inconsistent beliefs.[4] Such an observation, however, does not yield necessary and sufficient conditions for a set of beliefs to be that of an agent whose beliefs satisfy the Lockean thesis with threshold .5. That is, the following two statements are not equivalent.[5]

(a) A set of propositions, B, is pairwise consistent.

(b) There is a probability function that assigns each member of B a probability greater than .5.

In this paper, I present two types of characterizations of sets of beliefs for Lockean agents. I draw on a pair of related traditions for characterizing graded beliefs. The first tradition is that of the *Dutch Book argument* for probabilism. In this tradition, graded beliefs can be characterized by how they rationalize bets. Very roughly speaking, the Dutch Book argument establishes that a set of numerically graded beliefs is probabilistically coherent if and only if there is no collection of bets it rationalizes that lead to a sure loss.[6] The second tradition is that of the *accuracy argument* for probabilism. In this tradition, numerical credences are given scores corresponding to their accuracy.[7] The standard genre of result in this area is to establish that the numerical credences that are probabilistically coherent are equivalent to those that are not dominated by another set of credences in all worlds.

So, in the case of graded beliefs, both Dutch Book arguments and accuracy arguments are used to characterize probabilistically coherent beliefs (as opposed to numerical beliefs that don't satisfy the axioms of probability theory). What I show here is that these arguments can *also* be used to characterize all-or-nothing beliefs satisfying the Lockean hypothesis. There is a certain sense in which this trivial: since we can characterize probabilistic coherent graded beliefs using the Dutch Book and accuracy arguments, and there is a mapping from graded beliefs to sets of Lockean beliefs. So there are already *indirect* characterization of Lockeanism via the two characterizations of probabilism.

What I give here are *direct* characterizations of Lockeanism via betting dispositions and accuracy scoring rules. To do this I propose mappings from sets of all-or-nothing beliefs to bets and scores and show that, relative to these mappings, Lockean belief sets can be characterized by properties of betting dispositions and accuracy scores.

In the accuracy tradition there is already a small literature that addresses how the Lockean thesis fares vis-a-vis accuracy arguments.[8] This paper advances

---

[4]See, e.g., Hawthorne & Bovens (1999) for this and related observations.

[5]The relevant notions are made more precise in the next section.

[6]This argument goes back to Ramsey (1926), de Finetti (1974). See Pettigrew (2019) for a recent book-length review and discussion.

[7]See, e.g., Joyce (1998) and Pettigrew (2016*a*) . This work is, in turn, dependent on the tradition of scoring graded predictions for their accuracy going back at least to Brier (1950); see Gneiting & Raftery (2007) for a recent discussion.

[8]For example, Easwaran & Fitelson (2015), Pettigrew (2016*b*), Easwaran (2016), Dorst (2019).

that literature by providing a more general result than previously available and by proving, as a corollary of this result, the main conjecture of Easwaran (2016). There is, by contrast, no literature that I am aware of relating Dutch Book arguments to the Lockean account of belief. This paper fills that lacuna by giving two Dutch Book results for the Lockean account of belief.

Here is the plan: In the next section, §2, I outline the basic formal framework for full beliefs. In section §3 I define three different ways in which sets of beliefs can be said to satisfy the Lockean thesis. In §4, I discuss ways of deriving betting dispositions from full beliefs and what it takes for those dispositions to be subject to a Dutch Book. In §5, I give an accuracy scoring system for beliefs and state some standard decision-theoretic properties of beliefchoice relative to the system. In the main section, §6, I give results linking these different ways of thinking about beliefs. I discuss the significance of these results in §7 and outline a few directions for future work in §8.

2 FRAMEWORK

Let W be a finite set of m worlds, which we will enumerate $w_1 \dots w_m$, let P denote the set of proposition 2^W,[9] enumerated $p_1 \dots p_n$, with $n = 2^n$. Let $B \subseteq P$ be a set of o beliefs, $o \leq n$, enumerated $b_1 \dots b_o$. Let t be a real number, $0 \leq t \leq 1$, which we'll call the threshold. We think about B as representing the set of propositions an agent believes (and hence $P \backslash B$ is the set of propositions the agent doesn't believe). In the following we characterize various properties that B can have, always in terms of the threshold t. These fall into three categories: Lockean properties, betting properties, and accuracy properties.

3 CHARACTERIZING LOCKEAN BELIEF SETS

On the Lockean thesis one should only believe a proposition p if you assign it a credence greater than or equal to some threshold t. In this section we'll explore some ways of expressing this and related claims as constraints on the belief set B relative to the threshold t.

Let us begin by defining credences. A CREDENCE FUNCTION is a function, c, from P to $[0, 1]$ such that $c(W) = 1$, $c(\emptyset) = 0$ and for any $X \subseteq W$, $c(X) = \sum_{w \in W} c(\{w\})$.[10]

Since we use linear algebra to prove most of our main results it will be useful to rephrase some of the notions in terms of vectors. We correspond to each proposition p an m-dimensional vector, $\mathbf{p}$, such that

$$p_i = \begin{cases} 1 \text{ if } w_i \in p \\ 0 \text{ otherwise} \end{cases} .$$

[9]*Notation*: 2^X is the powerset of X.

[10]This is equivalent to saying $\langle W, P, c \rangle$ is a probability space. Note that, unlike in standard Dutch Book arguments, we are here stipulating that credences satisfy the axioms of probability theory.

We correspond to the credence function c the vector $\mathbf{c}$ such that $c_i = c(\{w_i\})$. In this case it is easy to see that $c(p) = \mathbf{c} \cdot \mathbf{p}$, where $\mathbf{c} \cdot \mathbf{p}$ is the inner product of $\mathbf{c}$ and $\mathbf{p}$.[11]

We begin with the weakest sense in which B can satisfy the Lockean thesis relative to the threshold t. The set of beliefs B is LOCKEAN COMPATIBLE relative to the threshold t iff there is a credence function c such that $c(b) \geq t$ (or $\mathbf{c} \cdot \mathbf{b} \geq t$) for all $b \in B$.

To say that B is LOCKEAN COMPATIBLE is not equivalent to saying that B can be a Lockean's total set of beliefs. For example, suppose P includes the propositions 'it's raining', 'it's snowing' and 'it's raining or it's snowing'. The set of beliefs that just contains 'it's raining' is Lockean compatible. But that belief cannot be one's *only* belief, since if you believe it's raining you must also believe 'it's raining or it's snowing' on the Lockean thesis (since the probability you assign the latter must be greater than or equal to what you assign the former).

For this reason, we will define two stronger notion as well. The set of beliefs B is LOCKEAN COMPLETE at threshold t iff there is a credence function c such that $c(p) \geq t$ iff $p \in B$ for all $p \in P$. We will also use a weakening of this last notion to allow cases where belief is optional at the threshold t. We say B is ALMOST LOCKEAN COMPLETE relative to a threshold t iff there is a credence function c such that for all $p \in P$ if $c(p) > t$ then $p \in B$ and if $c(p) < t$ then $p \notin B$. As we shall see, this last notion connects more naturally with the decision theoretic notions we define in relation to accuracy scoring.

4 BETTING ON BELIEFS

Dutch Book arguments traditionally assume that an agent's rational betting behavior is determined by their credences. Dutch Book arguments are used to argue for the thesis that their credences are probabilistic (something we assumed above in the definition of credences). The idea behind the Dutch Book argument is that credences determine one's betting preferences (or ought to). The arguments typically assume some divisible good in which an agent's utility is linear (conventionally assumed to be dollars). Expected utility theory provides a standard way of linking rational betting dispositions to credences: a rational agent acts in order to maximize her expected utility. Thus if offered a bet, an agent ought to be willing to take it only if she expects to gain by it or at least not lose by it. Indeed the very idea of credences or subjective probability is often considered to be determined by or to determine betting behavior.[12] On this framework it is typically assumed that if an agent has a credence y in a proposition p then she ought to be willing to buy or sell a bet that pays out \$1 if p is true and \$0 otherwise (we'll call this bet *a bet for p*) for the price \$$y$.

[11] Note that, in general, if I use a variable 'x' for something I will use boldface '$\mathbf{x}$' as a variable for the vector associated with x according to some defined mapping.

[12] Beginning with Ramsey (1926).

This is because the agent's expected return from the bet is $\$y$ (In the decision theoretic framework, the agent also is rationally obliged to buy the bet for any lower value and sell it for any higher value.) A Dutch Book is a way of exploiting these dispositions, by presenting an agent with a collection of bets she is willing to take which will result in a sure loss for the agent, and, hence, a sure gain for the bookie (i.e. the person offering the bets). Traditional Dutch Book arguments show that, roughly, an agent will not be subject to a Dutch Book if and and only if her credences are probabilistically coherent.[13]

It is less straightforward to link all-or-nothing beliefs to betting behavior. The Lockean thesis, however, by providing a link between belief and credence suggests some ways of linking beliefs to betting behavior which I spell out here. On the Lockean thesis an agent ought to believe p if and only if she assigns p a credence greater than or equal to t. The agent, then, for each belief b ought to be willing to buy a bet for b for $\$t$ (or a lower price). (Since her expected return of the bet is greater than or equal to zero.) So we start with the assumption that there are some bets that a Lockean agent ought to be willing to buy (in any quantities). If there is a collection of such bets for an agent that guarantee her a loss at each world then the agent's beliefs B is subject to a ONE-WAY DUTCH BOOK with respect to t.

To state the notion of a one-way Dutch Book formally we will again use vector notation again. For an agent's beliefs, $b_1 \ldots b_o$, we can use a non-negative o-dimensional vector $\mathbf{x}$, the stake vector, to represent the number of bets the agent buys in each proposition $b_1 \ldots b_o$ (i.e. the agent buys x_i bets in b_i).[14] We assume here that all bets are priced at $\$t$ since that is the highest price that the Lockean view guarantees the agent will be willing to pay. At a world w these bets lead to the payoff,

$$\sum_{1 \leq i \leq o, w \in b_i} x_i(1-t) + \sum_{1 \leq i \leq o, w \notin b_i} -x_i t.$$

By associating each world w with an o-dimensional vector $\mathbf{w}$, such that

$$w_i = \begin{cases} (1-t) \text{ if } w \in b_i \\ -t \text{ otherwise} \end{cases}, \tag{1}$$

the payoff at w can be concisely stated as $\mathbf{w} \cdot \mathbf{x}$. A collection of beliefs $b_1 \ldots b_o$ is subject to a one-way Dutch Book just in case there is some vector of positive stakes $\mathbf{x}$ such that for all worlds w, $\mathbf{x} \cdot \mathbf{w} < 0$.

A one-way Dutch Book only takes advantage of the beliefs, $B = b_1 \ldots b_o$, the agent has, it does not exploit those propositions she *does not* belief, i.e. $P \backslash B$. On the Lockean view, the agent must have less than t credence in the propositions

[13]See Pettigrew (2019) for a recent review of these arguments.

[14]We are using an o-dimensional vector to represent a function from B to positive stakes.

not in B. She, thus, ought to be willing to *sell* \$1 bets on each proposition in $P \backslash B$ for \$$t$ each. A TWO-WAY DUTCH BOOK is a collection for the agent of bets in B she is willing to buy at \$$t$ *and* a collection of bets in $P \backslash B$ she is willing to sell at \$$t$ that guarantee her a loss.

We will formally state the notion of two-way Dutch Book in vector notation. Let a non-negative n-dimensional vector $\mathbf{x}$ represent the stakes an agent takes in bets costing \$$t$ on the propositions $p_1 \ldots p_n$ as follows: if p_i is one of the agent's beliefs then she *buys* x_i bets, if p_i is not one of the agents beliefs then she *sells* x_i bets.[15] That $\mathbf{x}$ is required to be nonnegative captures the fact that the choice of buying or selling a bet in a proposition is determined by whether or not the proposition is believe by the agent.

The payout an agent gets for stakes x_i with the set of beliefs B at a world w is this (unwieldy) sum:

$$\sum_{p_i \in B, p_i \in w} x_i(1-t) + \sum_{p_i \in B, p_i \notin w} -x_i t + \sum_{p_i \notin B, p_i \in w} -x_i(1-t) + \sum_{p_i \notin B, p_i \notin w} x_i t.$$

We can represent this as $\mathbf{x} \cdot \mathbf{w}$, where the n-dimensional vector $\mathbf{w}$ is defined as follows:

$$w_i = \begin{cases} 1-t \text{ if } p_i \in B, p_i \in w \\ -t \text{ if } p_i \in B, p_i \notin w \\ -(1-t) \text{ if } p_i \notin B, p_i \in w \\ t \text{ if } p_i \notin B, p_i \notin w \end{cases}. \tag{2}$$

An agent is subject to a two-way Dutch Book, if there is an non-negative n-dimensional stakes vector $\mathbf{x}$ such that for each world w, $\mathbf{x} \cdot \mathbf{w} < 0$. In other words, there is a set of bets the agents beliefs rationalize that lead to a loss at each world.

5 SCORING BELIEFS FOR ACCURACY

Another perspective associates sets of beliefs with scores, depending on the accuracy of the beliefs in the set. The picture here is that an agent chooses a set of beliefs B and is scored according to her choice. Scoring all-or-nothing beliefs is simpler conceptually than scoring graded beliefs: while saying whether a real-valued confidence in a proposition is right or wrong at a world is a tricky business, saying whether a belief is right or wrong is straightforward.[16] For this reason, we'll jump right into how to score sets of all-or-nothing beliefs without comparison to the scoring of graded beliefs.

We'll consider a scoring system for beliefs following Easwaran & Fitelson

[15] Note that we reuse $\mathbf{x}$ and $\mathbf{w}$ for both o-dimensional and n-dimensional vectors. This abuse of nottion is justified because of the similar construction and role these vectors play across different arguments and definitions. It will always be clear, I hope, which one we are discussing.

[16] See Gneiting & Raftery (2007) for a review of scoring rules for graded beliefs.

(2015), Easwaran (2016) and Dorst (2019). The crucial assumption made in this literature is that we assign scores to an agent for their total beliefs at a world by assigning numerical scores to individual propositions believed and summing those scores.[17] We will further assume that at a world w if a belief in a proposition p_i turns out to be correct (i.e. $w \in p_i$) then the agent gets a certain positive score $r_i \geq 0$, and if it is incorrect (i.e. $w \notin p_i$) the agent gets a negative score, $s_i \leq 0$. Note that if an agent fails to believe a proposition there is no score associated with that proposition. We will further assume that the ratio between the reward for true belief and penalty for false belief is constant across propositions so that for any two propositions p_i and p_j, $\frac{r_i}{s_i} = \frac{r_j}{s_j}$.[18] So, on this system of scoring beliefs, at a world w a set of beliefs B gets the score

$$S(w, b) = \sum_{w \in p_i} r_i + \sum_{w \notin p_i} s_i.$$

We now rephrase this scoring system in a way that makes more transparent its connection to Lockeanism. Using the threshold t, for reasons that will become apparent, let $t = \frac{-s_i}{r_i - s_i}$. Note that t does not depend on the choice of i and $0 \leq t \leq 1$. Let $x_i = r_i - s_i$. Note that $-tx_i = s_i$ and $(1-t)x_i = r_i$. We will refer to x_i as the WEIGHT on the score for proposition p_i, as the higher an x_i is the higher scores, both positive and negative, the proposition p_i has, relative to a fixed t. Note that x_i is always non-negative. Let $\mathbf{x}$ be the m-dimensional weight vector whose ith entry is x_i. Given this mapping, it is clear that any scoring system, as defined in the previous paragraph, can be fully specified by giving a $t : 0 \leq t \leq 1$ and a non-negative n-dimensional vector $\mathbf{x}$. In what follows, ee will specify scoring systems in terms of t and $\mathbf{x}$ without loss of generality.

Let $\mathbf{w}$, again, be the o-dimensional vector defined in (1). Let $\mathbf{x}$, the weight vector, be a non-negative o-dimensional vector giving the weights on a scoring system for the propositions $b_1 \dots b_n$.[19] The agents total score at w, $S(w, B)$ can be concisely stated in vector notation as $\mathbf{w} \cdot \mathbf{x}$.

Since we have a scoring system for beliefs, we can apply some standard notions from decision theory to the choice of belief sets. A choice of a set beliefs B is a SURE LOSS if for every world $S(w, B) < 0$. A belief set B is RATIONAL for an agent given a credence function c if and only if there is no other set B' such the expected score of B on c, $E(S(w, B'))$, is greater than the expected score of B' on c, $E(S(w, B'))$.[20] The choice of one set of beliefs B STRICTLY DOMINATES another

[17] The additivity assumption, that the total scores for one's beliefs depends on the sum of the individual scores, is defended by Dorst (2019).

[18] We can remove this assumption, but then we need to relate this scoring system to a slightly different definition of Lockeanism in which there is a different threshold for belief in each proposition, see Dorst (2019) for details.

[19] These are the only weights necessary for determining the score for the set of beliefs B, which is why we do not use our original n-dimensional weight vector also called $\mathbf{x}$. Again the reuse of the $\mathbf{x}$ and $\mathbf{w}$ for both o-dimensional and n-dimensional vectors is to emphasize their parallel roles.

[20] Where $E(S(w, B)) = \sum_{w \in W} c(\{w\}) S(w, B)$

choice of beliefs, B' if for all worlds w, $S(w,B) > S(w,B')$. A set of beliefs B WEAKLY DOMINATES another set of beliefs, B', if for all worlds w, $S(w,B) \geq S(w,B')$ and there is some world w' such that $S(w,B) > S(w,B')$. A set of beliefs B is not weakly/strictly dominated, if there is no set B' that weakly/strictly dominates it. Note that all of these notions are only defined relative to the scoring system, and hence the weight vector $\mathbf{x}$ (for all propositions) as well as the threshold t.

6 CONNECTING THE THREE PERSPECTIVES

First we will draw some connections between properties of Dutch Books and properties of scoring systems on our framework.

Observation 1. *For any set of worlds $W = w_1 \ldots w_m$ and set of beliefs $B = b_1 \ldots b_o$, and a positive real number t the following two statements are equivalent.*

(a) The agent holding B is subject to a one-way Dutch Book at threshold t.

(b) At some weight vector $\mathbf{x}$, the agent holding B will realize a sure-loss at threshold t.

Proof. Corresponding to each world w there is an o-dimensional vector $\mathbf{w}$ as defined in (1) on page 1. Note that $\mathbf{x} \cdot \mathbf{w}$ is both the payoff from holding bets on B with stake vector $\mathbf{x}$ at w and the score one gets for having beliefs B with weight vector $\mathbf{w}$ and threshold t. Both (a) and (b) are equivalent to stating that there is some non-negative o-dimensional vector $\mathbf{x}$ such that for every world w, $\mathbf{x} \cdot \mathbf{w} < 0$. □

Another, less obvious, connection between scoring and betting can be made as follows:

Observation 2. *For any set of worlds $W = w_1 \ldots w_m$ and set of beliefs $B = b_1 \ldots b_o$, and a positive real number t the following two statements are equivalent.*

(a) The agent holding B is subject to a two-way Dutch Book at threshold t.

(b) The choice of beliefs B is strictly dominated by another set of beliefs B', on the scoring system with threshold t, and some weight vector $\mathbf{x}$.

Proof. We will start by showing (a) implies (b): Let $\mathbf{w}$ be a non-negative n-dimensional vector as defined in (2) on page 2. Since B is subject to a two-way Dutch Book we know that there is some weight vector $\mathbf{x}$ such that for every world w, $\mathbf{w} \cdot \mathbf{x} < 0$. In scoring terms $\mathbf{w} \cdot \mathbf{x}$ represents the score at world w for holding B minus the score for holding $P \backslash B$. So $P \backslash B$ strictly dominates B on weights $\mathbf{x}$.

Now we will show that (b) implies (a). If B is strictly dominated we know that there is some other set of bets B' and some set of non-negative weights $\mathbf{x}$

such that for every world w the score for holding B with t and $\mathbf{x}$ is strictly less then the score for holding B'. Now define $\mathbf{x}'$ such that,

$$x_i' = \begin{cases} x_i \text{ if } b_i \in B \backslash B' \cup B' \backslash B \\ 0 \text{ otherwise} \end{cases}$$

It is straightforward to see that on $\mathbf{x}'$, B' strictly dominates B, and additionally $P \backslash B'$ strictly dominates B. Since $\mathbf{x}' \cdot \mathbf{w}$ is the score for holding B minus the score for holding $P \backslash B$, it must be negative for every w. Therefore $\mathbf{x}'$ are the stakes for a two-way Dutch Book.

□

We have just seen the close relationships between betting and accuracy characterizations of B. Now we need to see the connections to the scoring notions.

We begin by reviewing a connection between the scoring system and Lockeanism that have been established in the literature. The first is the following.[21]

Observation 3. *Given a scoring system with threshold t, a set of beliefs is rational relative to some credences c if it is almost Lockean complete with respect to t.*

Proof. We give a sketch, referring the reader to Dorst (2019) and Easwaran (2016) for more detail. Given the additivity of our scoring system and of expectations an agent maximizes expected score for the total set of beliefs, by making the 'best' choice for each proposition p whether to belief it or not. If the scoring system assign a positive weight to p this requires simply believing p only if $c(p) \geq t$ and whenever $c(p) > t$, hence being almost Lockean complete. Since weights are not required to be positive the converse fails. □

Relating rationality, which is expressed in terms of credences, to Lockeanism is relatively straightforward. The notion of sure-loss and the dominance relations are not stated in terms of credences and so their relationship to Lockeanism is not as obvious. We begin with a negative observation, due to Easwaran (2016). Suppose a belief set B is not weakly dominated on the scoring system with weights $\mathbf{x}$ and threshold t. This does *not* entail that B is (almost) Lockean compatible. A simple example suffices, where there are two worlds w_1, w_2 and threshold .6 and all propositions have weight 1 except $\{w_1\}$ which has weight 100. In this case the belief set $\{\{w_1\}, \{w_2\}, \{w_1, w_2,\}\}$ is not weakly dominated. It yields score 40 at w_1 and -59.2 at w_2. Dropping $\{w_1\}$ would reduce the score at w_1 while dropping $\{w_2\}$ would reduce the score at w_2, while dropping both would reduce the score at w_1. However, given the threshold of .6 an almost Lockean

[21] This is observed by Easwaran (2016) and Dorst (2019). The latter provides a generalization to cases where t is propositions dependent.

compatible set of beliefs cannot include two pairwise incompatible propositions such as a $\{w_1\}$ and $\{w_2\}$.

We now turn to our main results: characterizations of Lockean Compatibility and Almost Lockean Completeness in betting and scoring terms.

Here is our first result:

Theorem 1. *For any set of worlds $W = w_1 \ldots w_m$ and set of beliefs $B = b_1 \ldots b_o$, and a positive real number t the following three statements are equivalent.*

(a) The agent holding B is not subject to a one-way Dutch Book at threshold t.

(b) There is no weight vector $\mathbf{x}$ such that the agent holding B will realize sure-loss at threshold t.

(c) B is Lockean compatible with threshold t.

Proof. As we have already showed that (a) and (b) are equivalent in Observation 1, we will only need to show that (c) is equivalent to those two. To do so we will need to represent the entire situation in terms of matrix-vector multiplication.

We will take the matrix $m \times o$ matrix $\mathbf{A}$, constructed so that each row i consists of the row vector $\mathbf{w}_i$ corresponding to the world w_i as defined in (1) on page 5.

$$\mathbf{A} = \begin{bmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_m \end{bmatrix}$$

Or equivalently we can directly define $\mathbf{A}$ as follows:

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1o} \\ \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix},$$

where:

$$a_{ij} = \begin{cases} 1 - t \text{ if } w_i \text{ is in } b_j \\ -t \text{ otherwise} \end{cases}.$$

Note that if we multiply $\mathbf{A}$ by an o-dimensional column vector $\mathbf{x}$ we get the following:

$$\mathbf{Ax} = \begin{bmatrix} \mathbf{w}_1 \cdot \mathbf{x} \\ \vdots \\ \mathbf{w_m} \cdot \mathbf{x} \end{bmatrix}.$$

We can see, then, that if there is no non-negative $\mathbf{x}$ such that $\mathbf{Ax} < 0$ then (a) and (b) hold.[22] So, to put it more compactly, (a) and (b) are equivalent to the

[22]Note that, as is standard. $\mathbf{v} < 0$ means each entry of $\mathbf{v}$ is less than zero.

following holding:

$$\nexists \mathbf{x} \geq 0 : \mathbf{A}\mathbf{x} < 0 \tag{3}$$

Consider the m-dimensional credence vector $\mathbf{c}$ defined in section 3. Now consider multiplying such a vector in row format to $\mathbf{A}$ as follows:

$$\begin{aligned}
\mathbf{c}^T\mathbf{A} &= \begin{bmatrix} c_1, \ldots, c_m \end{bmatrix} \begin{bmatrix} a_{11} & \ldots & a_{1o} \\ \ldots & \ldots & \ldots \\ a_{m1} & \ldots & \ldots \end{bmatrix} \\
&= \left[\sum_{i:w_i \in b_1} c_i(1-t) - \sum_{i:w_i \notin b_1} c_i(t), \ldots, \sum_{i:w_i \in b_n} c_i(1-t) - \sum_{i:w_i \notin b_n} c_i(t) \right] \\
&= \left[c(b_1)(1-t) - (1-c(b_1))(t), \ldots, c(b_n)(1-t) - (1-c(b_n))(t) \right]
\end{aligned}$$

We can see that the ith coordinate of $\mathbf{c}^T\mathbf{A}$ is greater than or equal to 0 iff $c(b_i) \geq t$. Another way of seeing this is to note that the ith column of $\mathbf{A}$ is the vector $\mathbf{b}_i$ minus the vector $\mathbf{t}$ which has value t everywhere.[23] Thus, the ith entry of $\mathbf{c}^T\mathbf{A}$ is $\mathbf{c} \cdot (\mathbf{p} - \mathbf{t}) = \mathbf{c} \cdot \mathbf{p} - t$. If this is greater than or equal to 0, then $c(p) \geq t$.

We can now compactly state what it is for a set of beliefs to satisfy (c) as follows:

$$\exists \mathbf{y} \geq \mathbf{0} : \mathbf{y} \neq \mathbf{0}, \text{ and } \mathbf{y}\mathbf{A} \geq 0. \tag{4}$$

Note that we do not require $\mathbf{y}$ to sum to 1. However, it is easy to see that 4 is equivalent to

$$\exists \mathbf{y} \geq \mathbf{0}, \mathbf{y} \neq \mathbf{0}, \mathbf{y}\mathbf{A} \geq 0, \text{ and } \sum_{i=1}^{m} y_i = 1,$$

since multiplying $\mathbf{y}$ by a positive scalar does not affect any of the inequalities.

So all we need to show is that (3) and (4) are equivalent statements about the matrix $\mathbf{A}$, which is the content of this theorem from linear algebra, a variant of Farkas's Lemma, whose proof is in the appendix.

Theorem 2. *Let* $\mathbf{A}$ *be any* $m \times n$ *matrix, then:*

$$(\nexists \mathbf{x} \geq \mathbf{0} : \mathbf{A}\mathbf{x} < \mathbf{0}) \longleftrightarrow (\exists \mathbf{y} \geq \mathbf{0} : \mathbf{y} \neq 0 \textit{ and } \mathbf{y}\mathbf{A} \geq \mathbf{0})$$

□

The next theorem relates the notion of strict dominance to the scoring and betting frameworks.

Theorem 3. *The following three statements are equivalent:*

(a) There is no two-way Dutch Book on B.

(b) There is no weight vector $\mathbf{x}$ *on which B is strictly dominated.*

[23] Vectors for propositions were defined in section 3 and simply are m-dimensional vectors that are 1 at index i when proposition is true at w_i and 0 otherwise.

(c) *B is almost Lockean complete.*

Proof. We have already established that (a) and (b) are equivalent in Observation 2.

Consider the following $m \times n$ matrix $\mathbf{D}$.

$$\mathbf{D} = \begin{bmatrix} d_{11} & \dots & d_{1n} \\ \dots & \dots & \dots \\ d_{m1} & \dots & d_{mn} \end{bmatrix},$$

where,

$$d_{ij} = \begin{cases} 1-t \text{ if } w_i \in p_j \text{ and } p_j \in B \\ -(1-t) \text{ if } w_i \in p_j \text{ and } p_j \notin B \\ -t \text{ if } w_i \notin p_j \text{ and } p_j \in B \\ t \text{ if } w_i \notin p_j \text{ and } p_j \notin B \end{cases}.$$

Note that the ith row of $\mathbf{D}$ corresponds to the vector $\mathbf{w}_i$ associated with the world w_i as defined in (2) on page 6. Thus, given an n-dimensional non-negative weighting vector $\mathbf{x}$, $\mathbf{x} \cdot \mathbf{w}_i < 0$ iff taking stakes $\mathbf{x}$ on the bets associated with the belief set B leads to a sure loss at w_i. Since

$$\mathbf{Dx} = \begin{bmatrix} \mathbf{w}_1 \cdot \mathbf{x} \\ \vdots \\ \mathbf{w_n} \cdot \mathbf{x} \end{bmatrix},$$

the statement,

$$\nexists \mathbf{x} : \mathbf{x} \geq 0 \text{ and } \mathbf{Dx} < 0, \tag{5}$$

is equivalent to saying that there is no two-way Dutch Book on B (at the threshold t), i.e. that (a) and (b) are true.

Considered column-wise the ith column of $\mathbf{D}$ is $\mathbf{p_i} - \mathbf{t}$ if $p_i \in B$ and $-(\mathbf{p_i} - \mathbf{t})$ if $p_i \notin B$, where $\mathbf{p_i}$ is the m-dimensional proposition vector as defined in section 3 and $\mathbf{t}$ is the m-dimensional vector with value t at all coordinates. There being a credence function that is almost Lockean complete, (c) above, is thus equivalent to there being credence vector $\mathbf{c}$ such that $\mathbf{c}^T A \geq 0$. This is equivalent to this simplified condition:

$$\exists \mathbf{y} : y \geq \mathbf{0}, \mathbf{y} \neq \mathbf{0}, \text{ and } \mathbf{yD} \geq 0. \tag{6}$$

To complete our proof we note that Theorem 2 establishes that (5) and (6) are equivalent statements about $\mathbf{D}$. It follows that (a), (b), and (c) are equivalent. □

The following conjecture from Easwaran (2016) follows directly from Theorem 3.

Lemma 1 (Easwaran's Conjunecture)**.** *If for all positive weight vectors* $\mathbf{x}$*, B is not weakly dominated by some B', then B is almost Lockean complete.*

Proof. Given Theorem 3 what we need to show is that if B is not weakly dominated on any positive weighting function, then there is no non-negative weighting function on which B is strictly dominated. We will prove the contrapositive. Suppose B is strictly dominated by B' on some non-negative weight vector $\mathbf{x}$. Let $\mathbf{x}'$ be a strictly positive weight vector such that

$$x_i' = \begin{cases} x_i \text{ if } x_i > 0 \\ 1 \text{ otherwise} \end{cases} .$$

Let B'' be the following belief set:

$$B'' = \{p_i : p_i \in B'\} \cup \{p_i : p_i \in B \text{ and } x_i = 0\}.$$

One can see that since B' strictly dominates B on $\mathbf{x}$ then B'' will strictly dominate B on $\mathbf{x}'$. So B' is weakly dominated on the positive weight vector $\mathbf{x}'$. □

Note that the converse of Lemma 1 fails. B can be almost Lockean compatible but still weakly dominated on some strictly positive weighting function. This is because if the credence function c assigns 0 to the proposition $\{w_1\}$ and t to $\{w_2\}$ then B might include $\{w_1\}$ and not include $\{w_1, w_2\}$ and still be almost Lockean compatible. It will be weakly dominated (on an even weighting vector) by a variation B' that is like B except it includes $\{w_1, w_2\}$ but not $\{w_1\}$.

7 DISCUSSION

Theorems 1 and 3 provide characterizations of Lockean belief sets (both Lockean compatible and almost Lockean complete) in terms of the betting and accuracy frameworks.

These results can be used to motivate the Lockean thesis itself. If beliefs are taken either to guide betting behavior or are subject to the scoring system above for their accuracy, then the results above give insight into why beliefs should conform to the Lockean thesis. It is worth noting though that, with respect to the accuracy argument the requirement for belief choice to be Lockean is rather strong: it is not just that you chose a strategy that is not strictly dominated but rather one that is not strictly dominated at *any* weighting vector. It is not clear why on the accuracy framework belief-choice should be subject to such a strong constraint.

8 FURTHER DIRECTIONS

8.1 CONNECTIONS TO WALD'S COMPLETE CLASS THEOREM

I would conjecture that an alternative characterization of almost Lockean completeness can likely be found using Wald's Complete Class Theorem.[24] Wald's Complete Class Theorem establishes that in a certain class of decision problems any strategy that is not weakly dominated is a rational strategy (i.e. one maximizing expected reward subject to her credences). Wald's Complete Class theorem does not *directly* apply to the choice problem here: for, as I noted in section 6, there are some belief sets that are not weakly dominated at some weighting function but that are *also* not almost Lockean complete. The reason Wald's Complete Class theorem does not apply is that it requires the class of strategies to be convex: i.e. if two strategies are available to an agent then any probabilistic mix of them also is. If we had allowed mixed strategies in our decision problem, then the Complete Class Theorem would apply and any strategy that was not weakly dominated by any mixed strategy would be almost Lockean complete. However, the main point of modeling outright belief is to have a notion of belief that contrasts with graded belief. If we allowed mixed strategies involving outright belief, it is not so clear that we have not introduced an analogue of graded beliefs. Nonetheless, I conjecture that by applying Wald's Complete Class Theorem we can show that outright beliefs that are almost Lockean complete will be just those that are not weakly dominated by any *mixture* of other outright beliefs. Future work might also investigate how we ought to think about mixed strategies in a system of outright belief and how such mixed strategies relate to graded beliefs.

8.2 EXTENSIONS TO OTHER SCORING SYSTEMS

Note that we assumed here that at the threshold t, the betting odds, and the scoring system ratios of cost to rewards, were the same for every proposition. A simple extension of these results could handle cases in which the odds are allowed to vary for each proposition.[25] In addition to this last extension, future work might also explore other scoring systems and see whether they also give rise to Lockean patterns of beliefs. Most interesting, I think, would be to see what can be found when the additivity assumption is weakened.

8.3 CONNECTION TO ANOTHER CHARACTERIZATION

In this paper I characterized the Lockean belief sets by way of their relationship to bets and accuracy scoring systems. Fernando (1998, Theorem 4, p. 230) gives

[24] I am grateful to Gary Chamberlain here for pointing out the relevance of Wald's Complete Class Theorem. See, e.g., Ferguson (1967) for details.

[25] See Dorst (2019) for the basic formal layout of such a scoring system. The basic idea is that a n-dimensional vector $\mathbf{t}$ would specify the threshold for each proposition. Introducing this complexity does not alter any of the results above given the linearity of the system.

a different characterization of Lockeanism that is based on qualitative properties of sets of beliefs.[26] It would be interesting to explore the connections between this result and those presented here.

APPENDIX: PROOF OF THEOREM 2

Theorem 2 states that this equivalence for any matrix **A**:

$$(\nexists \mathbf{x} \geq \mathbf{0} : \mathbf{A}\mathbf{x} < \mathbf{0}) \longleftrightarrow (\exists \mathbf{y} \geq \mathbf{0} : \mathbf{y} \neq 0 \text{ and } \mathbf{y}\mathbf{A} \geq \mathbf{0})$$

We start with this widely variant of Farkas's Lemma result:[27]

$$(\text{there is no } \mathbf{x} \geq \mathbf{0} \text{ such that } \mathbf{A}\mathbf{x} \geq \mathbf{b}) \longleftrightarrow (\text{there is } \mathbf{y} \text{ with } \mathbf{y}\mathbf{A} \geq \mathbf{0} \text{ and } \mathbf{y}\mathbf{b} < 0)$$

We can get the following by taking a universal instantiation of the both sides of the biconditional. (i.e. going from ' p iff q' → '(for all x p) iff (for all x q)').

$$(\forall \mathbf{b} > 0, \nexists \mathbf{x} \geq \mathbf{0} : \mathbf{A}\mathbf{x} \geq \mathbf{b}) \longleftrightarrow (\forall \mathbf{b} > \mathbf{0}, \exists \mathbf{y} : \mathbf{y}\mathbf{A} \geq 0 \text{ and } \mathbf{y}\mathbf{b} < 0)$$

This can be simplified to:

$$(\nexists \mathbf{x} \geq \mathbf{0} : \mathbf{A}\mathbf{x} > \mathbf{0}) \longleftrightarrow (\exists \mathbf{y} \leq \mathbf{0} : \mathbf{y} \neq \mathbf{0} \text{ and } \mathbf{y}\mathbf{A} \geq \mathbf{0})$$

We can then switch signs on the right-hand side to get:

$$(\nexists \mathbf{x} \geq \mathbf{0} : \mathbf{A}\mathbf{x} > \mathbf{0}) \longleftrightarrow (\exists \mathbf{y} \geq \mathbf{0} : \mathbf{y} \neq 0 \text{ and } \mathbf{y}\mathbf{A} \leq \mathbf{0})$$

Theorem 1 follows immediately (by substitution of **A** with −**A**):

$$(\nexists \mathbf{x} \geq \mathbf{0} : \mathbf{A}\mathbf{x} < \mathbf{0}) \longleftrightarrow (\exists \mathbf{y} \geq \mathbf{0} : \mathbf{y} \neq 0 \text{ and } \mathbf{y}\mathbf{A} \geq \mathbf{0})$$

□

REFERENCES

Adams, E. W. (1965), 'Elements of a theory of inexact measurement', *Philosophy of Science* **32**(3/4), 205–228.
URL: *https://doi.org/10.1086/288044*

Brier, G. W. (1950), 'Verification of forecasts expressed in terms of probability', *Monthly Weather Review* **78**(1), 1–3.
URL: *https://doi.org/10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2*

de Finetti, B. (1974), *Theory of probability*, Vol. 1, John Wiley & Sons.

[26] Wes Holliday pointed out to me that Fernando's result can also be found as a direct consequence of results in Adams (1965), see also Fishburn (1969).

[27] For statements and proof see Matoušek & Gärtner (2007, pp. 89–92), Strang (2006, p. 441–42), and or the wikipedia entry on Farkas's lemma (Wikipedia contributors 2019).

Dorst, K. (2019), 'Lockeans maximize expected accuracy', *Mind* **128**, 175–211.

Easwaran, K. (2016), 'Dr. Truthlove or: How I learned to stop worrying and love Bayesian probabilities', *Noûs* **50**(4), 816–853.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/nous.12099*

Easwaran, K. & Fitelson, B. (2015), 'Accuracy, coherence, and evidence', *Oxford Studies in Epistemology* **5**, 61–96.

Ferguson, T. S. (1967), *Mathematical Statistics, A decision theoretic approach*, Academic Press.

Fernando, T. (1998), 'In conjunction with qualitative probability', *Annals of Pure and Applied Logic* **92**(3), 217–234.

Fishburn, P. C. (1969), 'Weak qualitative probability on finite sets', *The Annals of Mathematical Statistics* **40**(6), 2118–2126.
URL: *http://www.jstor.org/stable/2239524*

Foley, R. (1992), 'The epistemology of belief and the epistemology of degrees of belief', *American Philosophical Quarterly* **29**(2), 111–124.
URL: *http://www.jstor.org/stable/20014406*

Gneiting, T. & Raftery, A. E. (2007), 'Strictly proper scoring rules, prediction, and estimation', *Journal of the American Statistical Association* **102**(477), 359–378.
URL: *https://doi.org/10.1198/016214506000001437*

Hawthorne, J. & Bovens, L. (1999), 'The preface, the lottery, and the logic of belief', *Mind* **108**(430), 241–264.
URL: *http://www.jstor.org/stable/2660219*

Joyce, J. M. (1998), 'A nonpragmatic vindication of probabilism', *Philosophy of Science* **65**(4), 575–603.

Leitgeb, H. (2014), 'The stability theory of belief', *Philosophical Review* **123**(2), 131–171.

Matoušek, J. & Gärtner, B. (2007), *Understanding and Using Linear Programming*, Springer.

Pettigrew, R. (2016*a*), *Accuracy and the Laws of Credence*, Oxford University Press.

Pettigrew, R. (2016*b*), 'Jamesian epistemology formalised: An explication of 'The Will to Believe'', *Episteme* **13**(3), 253–268.

Pettigrew, R. (2019), *The Dutch Book Arguments*, Cambridge University Press.

Ramsey, F. P. (1926), *Truth and Probability*, Routledge, pp. 52–94. (1978).

Strang, G. (2006), *Linear Algebra and Its Applications*, 4th edn, Thomson Brooks/Cole.

Wikipedia contributors (2019), 'Farkas' lemma — Wikipedia, the free encyclopedia', `https://en.wikipedia.org/w/index.php?title=Farkas%27_lemma&oldid=879978648`. [Online; accessed 6-February-2019].