AVOIDING DUTCH BOOKS DESPITE INCONSISTENT CREDENCES

ALEXANDER R. PRUSS

Abstract. It is often loosely said that Ramsey (1931) and de Finetti (1937) proved that if your credences are inconsistent, then you will be willing to accept a Dutch Book, a wager portfolio that is sure to result in a loss. Of course, their theorems are true, but the claim about acceptance of Dutch Books assumes a particular method of calculating expected utilities given the inconsistent credences. I will argue that there are better ways of calculating expected utilities given a potentially inconsistent credence assignment, and that for a large class of credences—a class that includes many inconsistent examples—these ways are immune to Dutch Books and single-shot domination failures. The crucial move is to replace Finite Additivity with Monotonicity (if $A \subseteq B$, then $P(A) \leq P(B)$) and then calculate expected utilities for positive U via the formula $\int_0^\infty P(U > y)\,dy$. This shows that Dutch Book arguments for probabilism, the thesis that one's credences should be consistent, do not establish their conclusion. Finally, I will consider a modified argument based on multi-step domination failure that does better, but nonetheless is not as compelling as the Dutch Book arguments appeared to be.

1. Introduction

The following is accepted wisdom: If your credence assignments do not satisfy the axioms of finitely-additive probability, then you will be subject to a Dutch Book, i.e., a bookie can offer you a deal such that (a) you are rationally permitted (or even required, on some versions) to accept the deal, but (b) no matter what happens, you are guaranteed to lose overall (Ramsey 1931; de Finetti 1937). This fact is supposed to, either pragmatically (see the literature in Hájek 2008, Section 2.1) or by vividly demonstrating a non-pragmatic inconsistency (e.g., Howson and Urbach 1993), show that your credences *should* be consistent in the sense of satisfying the axioms.

But whether this line of thought works depends on *how* you decide using your inconsistent credence assignments. For instance, here is one way to decide with inconsistent credence assignments: maximize the mathematical expectation of the utility function with respect to the consistent probability

Date: July 7, 2020. This is a post-peer-review, pre-copyedit version of an article to be published in *Synthese*. The final authenticated version is not yet available online.

Key words and phrases. credence; Dutch Book; probability; expected value; inconsistency; monotonicity.

function closest (in some metric, and with some tie-breaking procedure) to your credence assignment. Then, since consistent credences are not subject to Dutch Books (Kemeny 1955; Lehman 1955; Hájek 2008), you won't be subject to them either. So, even if your credences are inconsistent, you need not be subject to Dutch Books.

There are two problems with this response to the accepted wisdom. First, one might have a pragmatic approach to credences on which credences are *defined* by propensities for decision-making. On such a view, in the above example, you are deciding just like you would if you had that closest consistent assignment instead of your inconsistent one, and hence that consistent assignment *is* your credence assignment. I think such pragmatic approaches are mistaken in general, but my disagreement with them will go beyond the scope of this paper.

But there is a second more serious problem with the above example of how to decide with inconsistent credences. For a typical metric (e.g., Euclidean distance or maximum deviation) there will be infinitely many inconsistent credences with the same closest consistent assignment, and hence it is implausible to say that the procedure reflects a reasonable take on "deciding with inconsistent credence P". What we want is a procedure that is more sensitive to the details of the credence assignments. But at the same time, the procedure has to be in some intuitive sense natural.[1]

Promisingly, Hedden (2013) has recently shown that there is a decision procedure for binary decisions that in some cases of inconsistent credences provably escapes Dutch Books. However, Pettigrew (2019) has also shown that the natural extension of Hedden's decision procedure to more complex choices will require one to choose a dominated wager over a dominating one, unless one's credences are a scaled version of consistent credences.

I will show that there are two fairly natural procedures for deciding with inconsistent credences that, for certain classes of inconsistent credences, are not subject to Dutch Books. Moreover, these procedures will be quite sensitive to the details of the inconsistent credences. This will show the failure of Dutch Book arguments for probabilism (the thesis that our credences should satisfy the axioms of probability) on the basis of Dutch Books. I will also show that the procedures never lead to one's choosing a dominated wager over the dominating wager, and discuss continuity properties.

Next, I will move on to consider a different kind of argument, based on avoiding dominated outcomes in *sequences* of wagers, an argument close to one from Pettigrew's (2019) concept of "finite exploitability". We will see

[1]On any finite sample space Ω, the set A of all possible credence functions, i.e., assignments of probabilities to all the subsets of Ω, and the set C of all consistent credence functions have the cardinality of the continuum. Thus, there is a bijection ϕ between A and C. We could fix such a bijection ϕ, and then make decisions with a credence function P by maximizing mathematical expectation with respect to $\phi(P)$. Here we have maximal sensitivity to the details of P: no two credence functions will result in the same decision dispositions. But this is a very unnatural decision-making procedure.

that the two decision procedures do suffer a failure here, but I will argue that it is not clear that this is as intuitively as serious a problem as being subject to a Dutch Book.

Finally, I will argue that Pettigrew's (2019) extension of Hedden's (2013) prevision fails an intuitive continuity property for inconsistent credences and there is thus independent reason to reject it.

2. Stage-setting

2.1. **Credence assignments.** *Events* can be thought of as members of an algebra $\mathcal{F}$ of subsets of a sample space Ω (i.e., $\mathcal{F}$ is a non-empty collection of subsets of Ω closed under complements and finite unions). If $\mathcal{F}$ is a σ-algebra (i.e., it is also closed under countable unions), then a *probability function* is an assignment P of numbers to events that satisfies the Kolmogorov axioms:

(1) *Non-negativity:* $P(A) \geq 0$
(2) *Normalization:* $P(\Omega) = 1$
(3) *Countable Additivity:* $P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \cdots$ if the A_i are pairwise disjoint ($A_i \cap A_j = \varnothing$ whenever $i \neq j$).

In the simpler case where $\mathcal{F}$ is finite, Countable Additivity is equivalent to:

(4) *Finite Additivity:* $P(A \cup B) = P(A) + P(B)$ whenever $A \cap B = \varnothing$.

I will say that an assignment P of numbers to members of some algebra of subsets is *consistent* or *a finitely additive probability* provided that it satisfies Non-negativity, Normalization and Finite Additivity.

Some other axioms follow from Non-negativity and Finite Additivity, but will be of independent interest:

(5) *Zero:* $P(\varnothing) = 0$
(6) *Monotonicity:* $P(A) \leq P(B)$ whenever $A \subseteq B$
(7) *Binary Non-Disappearance:* If $P(A) = 0$ and $P(B) = 0$ and A and B are disjoint, then $A \cup B \neq \Omega$.

For the sake of mathematical simplicity, we are considering probabilities to be defined on subsets of a sample space rather than for sentences or propositions. This procedure by itself enforces a certain degree of consistency: logically equivalent propositions must receive the same probability. Since it is not the purpose of this paper to provide a reasonable decision procedure for all inconsistent credence assignments, but only to show that one exists for some, this will not be a problem.

2.2. **Wagers, utility functions and previsions.** A *credence space* will be a pair $\langle \Omega, \mathcal{F} \rangle$ of a set Ω and an algebra $\mathcal{F}$ of subsets or events. For convenience, I will often refer to the space as just Ω. Let a *wager* W be a finite sequence of pairs $\langle \langle F_1, u_1 \rangle, \dots, \langle F_n, u_n \rangle \rangle$ where F_i is an event on a credence space Ω and u_i is the utility of that event eventuating. Thus, e.g., if Ω represents the toss of a coin, the wager

$$W = \langle \langle \Omega, -\$26 \rangle, \langle \{\text{heads}\}, \$100 \rangle \rangle$$

represents a case where you pay \$26 no matter what, and receive \$100 on heads.

For any wager $W = \langle\langle F_1, u_1\rangle, \ldots, \langle F_n, u_n\rangle\rangle$, there is an associated utility function U_W defined on Ω by:

$$U_W = \sum_{i=1}^{n} u_i \cdot 1_{F_i},$$

where 1_A is the indicator function of A, i.e., the function whose value is 1 on A and 0 everywhere else.

A *simple function* on a credence space $\langle\Omega, \mathcal{F}\rangle$ is a function U from Ω to the reals that takes on only finitely many values and that has the property that for each real y, the set $\{\omega \in \Omega : U(\omega) = y\}$ of all the points where U has the value y is an event (i.e., is in $\mathcal{F}$).

The utility function associated with a wager is a simple function, and I will assume that all utility functions are simple. Conversely, every simple function f is associated with a canonical wager $W_f = \langle\langle F_1, y_1\rangle, \ldots, \langle F_n, y_n\rangle\rangle$ where $y_1 < \cdots < y_n$ are the values of U and F_i is the set of all points of Ω where U has the value y_i. It is easy to check that $U_{W_f} = f$ for any simple function f, but the converse is false, since multiple wagers can share the same utility function (e.g., $\langle\langle A, 1\rangle, \langle B, 1\rangle\rangle$ has the same utility function as $\langle\langle A \cup B, 1\rangle\rangle$ if A and B are disjoint).

A wager W_2 *dominates* a wager W_1 provided that $U_{W_2} > U_{W_1}$ everywhere on Ω.

A *prevision* on the set of all wagers or on the set of all simple functions on a credence space Ω is a function E from wagers or simple functions to real numbers, intended to represent the expected utility of a wager.

We can get a prevision E^* on the space of all wagers derivatively from a prevision E on the space of simple functions with the rule $E^*W = EU_W$, and, conversely, given a prevision E on the space of wagers, one can use $E^\dagger f = EW_f$ to define a prevision on the space of simple functions. When no ambiguity results, in the case of previsions on wagers derived from previsions on functions, I will drop the asterisk.

I will say that two wagers, W and W', are *equivalent* provided that they define the same utility function, i.e., $U_W = U_{W'}$. A prevision on the space of wagers is *integral-like* if and only if it assigns the same values to equivalent wagers. Any prevision derived as above from a prevision on the space of simple functions is integral-like and, conversely, if a prevision E on the space of wagers is integral-like, then it is easily checked that $E = (E^\dagger)^*$ and hence E is derived from the prevision $E^\dagger$ on the space of simple-functions.

Classically, the prevision mathematicians and decision theorists have been most interested in was the mathematical expectation with respect to a consistent credence assignment. This is a prevision on the simple functions and from which we can derive an integral-like prevision on the space of wagers. But our main interest in this paper is in other previsions, ones that can be naturally defined in terms of an inconsistent credence assignment.

I will say that a prevision E on the space of wagers satisfies the Zero Condition provided that $EW = 0$ when W is the empty wager $\langle\rangle$, and a prevision E on the space of simple functions satisfies the Zero Condition provided that $EU = 0$ if U is a function that is equal to zero everywhere on Ω. All the previsions we will be interested in will satisfy the Zero Condition, at least as long as they are defined using credences P that satisfy the Zero Axiom for credences, i.e., $P(\varnothing) = 0$.

2.3. **Single-shot decisions using previsions.** Suppose I have a prevision E on the space of wagers satisfying the Zero Condition and a wager W is offered to me. There is a natural decision theory associated with E: I am required to accept W if and only if $EW > 0$ and permitted if and only if $EW \geq 0$.

This nicely generalizes to cases of a choice between multiple wagers. If I am choosing between finitely many wagers, then on the natural decision theory associated with E, I am required to choose some wager with maximal prevision, and permitted to choose any such wager. The binary case where I just choose whether to accept or reject a wager is a special case of this if we take, as we shall, a choice to accept or reject a wager W to be equivalent to a choice between W and the empty wager $\langle\rangle$, assuming the Zero Condition which ensures that $E\langle\rangle = 0$.

A prevision E on the space of wagers is subject to a weak (respectively, strong) single-shot Dutch Book provided that there is a wager W such that the associated utility function U_W is strictly negative everywhere on Ω and yet it is E-permissible (E-required) that I accept W, i.e., $EW \geq 0$ ($EW > 0$).

Our treatment of betting on wagers involves a slightly different setting from the traditional one in the literature beginning with de Finetti (1937). In the literature, the question isn't just whether to accept or reject a wager simply, but whether to accept or reject a wager at a specific price. But we can model the question of whether to accept a wager W at a price r in our setting simply by asking whether the agent should simply accept the modified wager

$$W_r = \langle\langle F_1, u_1\rangle, \ldots, \langle F_n, u_n\rangle, \langle\Omega, -r\rangle\rangle$$

(where $W = \langle\langle\langle F_1, u_1\rangle, \ldots, \langle F_n, u_n\rangle\rangle\rangle$). There are reasons to prefer the setting of this paper. The paying of the price r is a part of what the agent is agreeing to, and hence it should simply be treated as yet another part of the wager, albeit one that is certain (i.e., occurs on the certain event Ω). Moreover, let us imagine that a coin is flipped, and you're an agent who assigns credence $1/4$ to heads, credence $1/4$ to tails, and credence $1/2$ for the whole probability space. Suppose that for \$10 you are offered a wager that pays \$24 on heads. The traditional approach will have you computing the value of the \$24-on-heads as $(1/4)\$24 = \$6$ (and all the previsions we will consider will agree in this computation), and hence rejecting buying the wager for \$10. That makes sense if you were to consider the probability of the \$10 payment to be one, but you don't: you think the probability of that payment is 1/2, since that's the probability of the whole space. So, rather than comparing the \$6 to \$10, you would be more reasonable to compare it to $(1/2) \cdot \$10 = \$5$, and hence accept the wager. And that acceptance is, indeed, the more reasonable thing to do.

In any case, for the previsions we will consider in this paper, our approach is equivalent to the traditional one in the special case where P satisfies Normalization, i.e., $P(\Omega) = 1$. The assumption of Normalization is largely a conventional one, since as long as $P(\Omega) \neq 0$ we can just normalize P by dividing it by $P(\Omega)$, and so nothing of much philosophical interest is lost by making the assumption as needed, except for the special case where the agent is so inconsistent as to assign credence 0 to Ω, a special case we won't need to consider since our purpose is to provide a decision procedure that works for some, but not all, inconsistent agents.

Next, let us say that the prevision E is subject to a weak (respectively, strong) single-shot failure of domination provided there are wagers W_1 and W_2 such that W_2 dominates W_1 and yet E permits (requires) one to choose W_1 over W_2.

Note that if E satisfies the Zero Condition, then subjection to weak or strong single-shot Dutch Books entails the corresponding failure of single-shot domination. For if W is a Dutch Book (either weak or strong), then W is dominated by $\langle\rangle$.

A prevision is *weakly (strongly) monotonic* provided that it satisfies Zero and does not suffer from single-shot strong (weak) failure of domination, i.e., $E\langle\rangle = 0$ and if W_2 dominates W_1 then $EW_2 \geq EW_1$ ($EW_2 > EW_1$). Classical mathematical expectation with respect to a consistent probability assignment is strongly monotonic.

2.4. **Multi-shot decisions using previsions.** Suppose that instead of being offered a single wager, I am offered multiple wagers in a fixed finite sequence, with me being able to accept or reject each one as it comes, and with the wagers all becoming resolved at the end of the sequence. There is then a natural decision theory given a prevision E: I am required to accept all wagers with positive E-value and permitted to accept all with a non-negative E-value.

However, this natural decision theory for multi-shot decisions is not the only one. For it might be that whether I accept a given wager is affected by what wagers I have already accepted. After all, if I have already accepted a wager W_1, which has yet to be resolved, and now am being asked whether to accept a wager W_2, then I am not really choosing between W_2 and nothing (i.e., $\langle\rangle$). Rather, I am choosing between $W_1 + W_2$ and W_1, where "+" is concatenation of wagers:

$$\langle\langle F_1, u_1\rangle, ..., \langle F_n, u_n\rangle\rangle + \langle\langle F'_1, u'_1\rangle, ..., \langle F'_{n'}, u'_{n'}\rangle\rangle$$
$$= \langle\langle F_1, u_1\rangle, ..., \langle F_n, u_n\rangle, \langle F'_1, u'_1\rangle, ..., \langle F'_{n'}, u'_{n'}\rangle\rangle.$$

This suggests that intuitively a better decision theory would be one on which I take into account the already-accepted wagers, so that if I've already accepted $W_{i_1}, ..., W_{i_j}$, then I am permitted to accept a new wager W_k when

$$E(W_{i_1} + \cdots + W_{i_j}) \leq E(W_{i_1} + \cdots + W_{i_j} + W_k)$$

and required to do so when this inequality is strict.

I will call the simpler decision theory, where I ignore previously accepted wagers, the *independent* decision theory, and the theory that compares concatenations of wagers the *cumulative* decision theory.

In the case of the classical mathematical expectation prevision with respect to consistent probabilities, the independent and cumulative theories are equivalent, since classical expectation not only satisfies Zero but is *additive* in the sense that $E(W_1 + W_2) = EW_1 + EW_2$ for any wagers W_1 and W_2, so that

$$E(W_{i_1} + \cdots + W_{i_j}) \leq E(W_{i_1} + \cdots + W_{i_j} + W_k)$$

if and only if $0 \leq EW_k$, and similarly for strict inequalities.

We can extend the two theories to multi-shot choices between multiple wagers. Specifically, if at a given step one is offered a choice of a wager from some finite set, then the independent decision theory for a prevision E says one is permitted to accept any wager with maximal E-value, and one is required to only accept such a wager. The cumulative decision procedure says that one is permitted to accept any wager W such that $E(W_{i_1} + \cdots + W_{i_j} + W)$ is maximal, and required to only accept such a wager, where $W_{i_1}, ..., W_{i_j}$ are the already accepted wagers.

Now, we say that a prevision E with a multi-shot decision theory T is subject to a weak (strong) multi-shot Dutch Book provided there is a fixed sequence of wagers $W_1, ..., W_n$ that one is permitted (required) by T to take all of in sequence and yet that is such that $W_1 + \cdots + W_n$ is dominated by $\langle\rangle$.

The following result will be quite useful:

Proposition 1. *If a prevision E is subject to a multi-shot weak or strong Dutch Book with respect to the cumulative decision theory, then it is respectively subject to a single-shot weak or strong Dutch Book.*

Proof. Suppose E is subject to the multi-shot weak (strong) Dutch Book $W_1, ..., W_n$ on cumulative decision theory. Then the utility function associated with $W = W_1 + \cdots + W_n$ is strictly negative. But EW is nonnegative (positive) or else the agent wouldn't have been permitted (required) to choose W_n after having chosen $W_1 + \cdots + W_{n-1}$ on cumulative decision theory, and hence W is a single-shot weak (strong) Dutch Book. □

Given Zero, being subject to weak (strong) single-shot Dutch Book W implies weak (strong) single-shot domination failure (with respect to the choice between W and $\langle\rangle$).

Corollary 1. *A strongly monotonic prevision E that satisfies Zero is not subject to weak single- or multi-shot Dutch Books given cumulative decision theory.*

Thus, to escape Dutch Books and single-shot domination failures, we simply need to have a strongly monotonic prevision. After a digression concerning the De Finetti prevision, I will give two previsions that are strongly monotonic for credence functions that satisfy Zero, Non-negativity, Monotonicity, and for one of them, Binary Non-Disappearance.

3. The De Finetti Prevision

The classical Dutch Book results that are used to support probabilism assume the following *De Finetti prevision* on the space of wagers:

$$\text{DeFinetti}_P\langle\langle F_1, u_1\rangle, ..., \langle F_n, u_n\rangle\rangle = \sum_{i=1}^{n} u_i P(F_i).$$

If P fails Non-negativity or Finite Additivity, decisions based on maximizing DeFinetti_P will be subject to one-shot weak Dutch Books.

To see this, observe first that if $P(\Omega) \leq 0$, then $\langle\langle\Omega, -1\rangle\rangle$ has a De Finetti prevision of $-P(\Omega) \geq 0$, and yet is a sure loss. Next, if $P(\Omega) > 0$, then we can assume without loss of generality that P satisfies Normalization, simply by replacing P with $P'(A) = P(A)/P(\Omega)$, as that will simply rescale the De Finetti previsions by a positive factor and hence not change whether they are zero, positive or negative, and thus will not change what wagers are permissibly accepted. But if we have Normalization, then as noted in Section 2.3 our setting does not differ from the traditional setting of de Finetti (1937), and we can use de Finetti's result to see that if we lack Non-negativity or Finite Additivity, then there is a weak Dutch Book.

Given a probability function P, the prevision DeFinetti_P on the space of wagers is equivalent to mathematical expectation on the space of simple functions. Thus, in that case DeFinetti_P is integral-like.[2] But if P is not finitely additive, then DeFinetti_P is not integral-like. To see this, note that if $P(A) + P(B) \neq P(A \cup B)$ but A and B are disjoint, then the wagers

$$W_1 = \langle\langle 1, A\rangle, \langle 1, B\rangle\rangle$$

and

$$W_2 = \langle\langle 1, A \cup B\rangle\rangle$$

[2]This is also true if P is consistent but not countably additive, but requires a slightly more complicated argument. Suppose $U_{W_1} = U_{W_2}$. Let $\mathcal{F}'$ be the finite algebra generated by the events entering into W_1 and W_2. Then since P is finitely additive and $\mathcal{F}'$ is finite, P will be countably additive on $\mathcal{F}'$. Hence, DeFinetti_P will be integral-like with respect to wagers defined with respect to $\mathcal{F}'$, and in particular $\text{DeFinetti}_P\, W_1 = \text{DeFinetti}_P\, W_2$.

have the same utility function, namely $1_{A\cup B}$, but different De Finetti previsions, namely $P(A)+P(B)$ and $P(A\cup B)$, respectively.

Thus, in addition to having Dutch Book failures in all cases without Finite Additivity, in those cases the De Finetti prevision also has the unfortunate property of assigning different previsions to equivalent wagers. It would be obvious even to many agents employing inconsistent probabilities that in the above example W_1 and W_2 are the same wager described differently, and so we should not base our arguments for consistency on imputing a strategy to these agents based on a prevision that evaluates W_1 and W_2 differently. In the next section, I will thus consider two different previsions, which will be integral-like because unlike the De Finnetti prevision they will be derived from a prevision on the space of utility functions.

4. Level Set Integrals

4.1. **Definitions.** Let U be any non-negative measurable function on a space Ω with a countably additive probability measure P. The following is a standard consequence of the Fubini-Tonelli Theorem and is much use in probability theory and analysis:

$$\int_\Omega U\,dP = \int_0^\infty P(U > y)\,dy, \tag{1}$$

where the left-hand-side is the Lebesgue integral of U—which defines the mathematical expectation of U with respect to P—while the integrals on the right hand side are improper Lebesgue or Riemann integrals (it does not matter which), and where $P(U > y)$ is short for $P(\{\omega \in \Omega : U(\omega) > y\})$, the probability of the event of U being bigger than y.[3]

Indeed, sometimes (1) is even used to define the Lebesgue integral with respect to P (in this case, the right-hand-side is interpreted as a Riemann integral; see Lieb and Loss 2001, p. 14), which supports the mathematical naturalness of (1).

The intuition behind (1) is that we can think of the Lebesgue integral $\int_\Omega U\,dP$ as the area under the graph U (where the x-axis represents Ω equipped with the measure P). We can compute this area by slicing the graph into thin horizontal slices and adding up the lengths of these slices multiplied by their thickness. Roughly speaking, the slice at height y above the x-axis has length $P(U > y)$. The sum of the areas of such slices is the integral $\int_0^\infty P(U > y)\,dy$ if U is a measurable function.

Note that while the Lebesgue integral $\int_\Omega U\,dP$ is only defined for a countably additive measure P, the improper integral $\int_0^\infty P(U > y)\,dy$ can in principle make sense for an inconsistent credence assignment P. In particular,

[3]Sketch of proof: Define the function $\phi(\omega, y)$ on $\Omega \times [0,\infty)$ as equal to 1 if $U(\omega) > y$ and 0 otherwise. One can prove that ϕ is measurable (Lieb and Loss 2001, p. 14). Thus, by Fubini-Tonelli $\int_\Omega \int_0^\infty \phi(\omega, y)\,dy\,dP(\omega) = \int_0^\infty \int_\Omega \phi(\omega, y)\,dP(\omega)\,dy$. But $\int_0^\infty \phi(\omega, y)\,dy = U(\omega)$ (since $\phi(\omega, y) = 1$ if $y \le U(\omega)$ and is zero otherwise) and $\int_\Omega \phi(\omega, y)\,dP(\omega) = P(U > y)$.

if P satisfies Zero, then the improper integral makes sense for any simple function U (since the function $y \mapsto P(U > y)$ will have at most finitely many discontinuities, corresponding to the finitely many values of U, and will be zero if y is large enough).

I will call $\int_0^\infty P(U > y)\,dy$ the *Level Set Integral* $\mathrm{LSI}_P\,U$ of the non-negative function U with respect to P when it is defined. If all the utility functions we were interested in were non-negative, we could take the Level Set Integral to be our prevision of choice. However, that would not be very useful, since for Dutch Books to be an issue we need to deal with utility functions that take on some negative values.

There are two mathematical ways of extending our definition of a Level Set Integral to all simple functions U. We can split the utility function U into positive and negative parts U^+ and U^-, where $U^\pm(\omega) = |U(\omega)|$ when $\pm U(\omega) \geq 0$ and $U^\pm(\omega) = 0$ otherwise, and then define the *Split Level Set Integral*:

$$\mathrm{LSI}_P^\pm\,U = \mathrm{LSI}_P\,U^+ - \mathrm{LSI}_P\,U^- = \int_0^\infty P(U > y)\,dy - \int_0^\infty P(U < -y)\,dy.$$

Alternately, we can shift U upwards to make it non-negative, compute the Level Set Integral, and then readjust it back down. More precisely, we choose a real number M such that $M + U$ is non-negative and then define the *Shifted Level Set Integral*:

$$\mathrm{LSI}_P^\uparrow\,U = (\mathrm{LSI}_P(M + U)) - M.$$

As long as P satisfies Normalization, the definition of the Shifted Level Set Integral does not depend on the particular choice of the shift M. For suppose that U is non-negative. Then for any non-negative m we have:

$$\begin{aligned}
\mathrm{LSI}_P(m + U) &= \int_0^\infty P(m + U > y)\,dy \\
&= \int_0^\infty P(U > y - m)\,dy \\
&= \int_{-m}^\infty P(U > y)\,dy \\
&= \int_{-m}^0 P(U > y)\,dy + \int_0^\infty P(U > y)\,dy \\
&= mP(\Omega) + \mathrm{LSI}_P\,U = m + \mathrm{LSI}_P\,U
\end{aligned}$$

since $P(U > y) = P(\Omega)$ for $y < 0$ as U is non-negative. It follows that if $M + U$ and $M' + U$ are both non-negative, with $M' \geq M$, then we can let $m = M' - M$ and then:

$$\begin{aligned}
\mathrm{LSI}_P(M' + U) - M' &= \mathrm{LSI}_P(m + M + U) - M' \\
&= \mathrm{LSI}_P(M + U) + m - M' = \mathrm{LSI}_P(M + U) - M,
\end{aligned}$$

so that it doesn't matter whether the Shifted Level Set Integral is defined with a shift of M or a shift of M'.

We can then extend the Split and Shifted Level Set Integrals to wagers by specifying that $\mathrm{LSI}^{\pm}_P W = \mathrm{LSI}^{\pm}_P U_W$ and $\mathrm{LSI}^{\uparrow}_P W = \mathrm{LSI}^{\uparrow}_P U_W$ for a wager W.

In the case of consistent credences, the Split and Shifted Level Set Integrals will be equal, but we cannot expect equality in the general case. The following are some useful facts whose simple proofs are left as an exercise to the reader:

Lemma 1. *Suppose that P satisfies Zero and Normalization. Then:*

(a) $\mathrm{LSI}^{\pm}_P U$ *and* $\mathrm{LSI}^{\uparrow}_P U$ *are defined and finite for any simple function U*
(b) $\mathrm{LSI}^{\pm}_P U = \mathrm{LSI}^{\uparrow}_P U = c$ *if U is a constant function equal to c everywhere*
(c) $\mathrm{LSI}^{\uparrow}_P(c + U) = c + \mathrm{LSI}^{\uparrow}_P U$ *for any simple function U and real c*
(d) $\mathrm{LSI}^{\uparrow}_P cU = c\,\mathrm{LSI}^{\uparrow}_P U$ *for any non-negative constant c*
(e) $\mathrm{LSI}^{\pm}_P cU = c\,\mathrm{LSI}^{\pm}_P U$ *for any real constant c*
(f) *If $U \geq 0$ everywhere, then* $\mathrm{LSI}_P U = \mathrm{LSI}^{\pm}_P U = \mathrm{LSI}^{\uparrow}_P U$
(g) *If $U \leq 0$ everywhere, then* $-\mathrm{LSI}_P(-U) = \mathrm{LSI}^{\pm}_P U = \mathrm{LSI}^{\uparrow}_P U$
(h) *For any event A,* $\mathrm{LSI}_P 1_A = P(A)$.

In particular, the Shifted Level Set Integral behaves well under positive affine transformations $x \mapsto a + bx$ for $b > 0$. We include the Split Level Set Integral in our discussions, however, because it is usual for mathematicians to define Lebesgue integrals by first defining them for the positive case and then splitting, and hence the Split Level Set Integral may seem more mathematically natural. The Split Level Set Integral has the property that it behaves well under mirroring, namely $\mathrm{LSI}^{\pm}_P -U = -\mathrm{LSI}_P U$, but mirroring is not a very natural transformation when considering utilities, except in the context of zero-sum games.

4.2. **Dutch Books.**

Theorem 1. *Suppose that P satisfies Zero, Non-negativity and Normalization. Then neither $\mathrm{LSI}^{\pm}_P$ nor $\mathrm{LSI}^{\uparrow}_P$ is subject to single-shot weak Dutch Books, or to multi-shot weak Dutch Books given cumulative decision theory.*

Proof. We omit the subscript P when no confusion will result. Suppose U is a utility function that is strictly negative everywhere. By Lemma 1, $\mathrm{LSI}^{\uparrow} U = \mathrm{LSI}^{\pm} U = -\mathrm{LSI}(-U)$. Note that $-U > 0$ everywhere. I now claim that if f is strictly positive everywhere, then $\mathrm{LSI}\, f > 0$. It will follow that $\mathrm{LSI}^{\uparrow} U = \mathrm{LSI}^{\pm} U < 0$ and hence we do not have a weak Dutch Book, as both previsions will require one not to choose the wager with utility function U.

To see that $\mathrm{LSI}\, f > 0$, let c be the smallest of the values of our strictly positive function f. Then $c > 0$ and:

$$\mathrm{LSI}\, f = \int_0^\infty P(f > y)\, dy \geq \int_0^c P(f > y)\, dy = \int_0^c 1\, dy = c,$$

where the first inequality follows from Non-negativity and the last equality from Normalization and the fact that $\{\omega : f(\omega) > y\} = \Omega$ if $y < c$. The claim about multi-shot Dutch Books follows from Proposition 1. □

We thus see that, given Zero, Non-negativity and Normalization, our level-set previsions allow us to escape Dutch Books.

4.3. **Single-Shot Domination Failure.** If E assigns a real number to every non-negative simple function, then we will say that E is weakly (strongly) monotonic on non-negative functions provided that if $0 \leq f < g$ everywhere, then $Ef \leq Eg$ ($Ef < Eg$). Being weakly (strongly) monotonic simply comes to not suffering from single-shot strong (weak) domination failure for non-negative utilities.

Here is a useful fact:

Theorem 2. *Suppose that P satisfies Zero, Non-negativity and Normalization. Then the following conditions are equivalent:*

(a) LSI_P *is strongly monotonic on non-negative functions*
(b) LSI_P *is weakly monotonic on non-negative functions*
(c) *if* $0 \leq f \leq g$ *everywhere, then* $\mathrm{LSI}_P\, f \leq \mathrm{LSI}_P\, g$
(d) *if* $f \leq g$ *everywhere, then* $\mathrm{LSI}_P^{\uparrow} f \leq \mathrm{LSI}_P^{\uparrow} g$
(e) $\mathrm{LSI}_P^{\uparrow}$ *is strongly monotonic*
(f) $\mathrm{LSI}_P^{\uparrow}$ *is weakly monotonic*
(g) $\mathrm{LSI}_P^{\pm}$ *is weakly monotonic*
(h) P *satisfies Monotonicity*

Additionally, the following are equivalent:

(i) $\mathrm{LSI}_P^{\pm}$ *is strongly monotonic*
(j) P *satisfies Monotonicity and Binary Non-Disappearance.*

The proof is given in the Appendix.

Corollary 2. *If P satisfies Zero, Non-negativity, Normalization and Monotonicity, then cumulative decision theory based on the prevision $\mathrm{LSI}_P^{\uparrow}$ is not subject to single-shot weak domination failure. The same is true for $\mathrm{LSI}_P^{\pm}$ if we additionally assume that P satisfies Binary Non-Disappearance.*

Finally, note that the conjunction of Zero, Non-negativity, Normalization, Monotonicity and Binary Non-Disappearance is strictly weaker than consistency. One way to see this is to note that if P is any consistent credence satisfying the non-triviality condition that there is an event A such that $0 < P(A) < 1$, and for α a positive number other than one we define

$P^\alpha(B) = (P(B))^\alpha$, then $P^\alpha(A) + P^\alpha(\Omega - A) \neq 1$,[4] so P^α is not additive. But P^α clearly satisfies Zero, Non-negativity, Normalization, Monotonicity and Binary Non-Disappearance. More generally given a consistent credence P, and any monotonic function $\phi : [0,1] \to [0,1]$ such that $\phi(0) = 0$, $\phi(1) = 1$ and $0 < \phi(x)$ for all x, the credence function $\phi \circ P$ (defined by $(\phi \circ P)(B) = \phi(P(B))$) satisfies Zero, Non-negativity, Normalization and Binary Non-Disappearance.

4.4. **Continuity.** *Continuity* will be the very plausible property of a prevision E that tells us that wagers that are "close" to each other have prevision values that are close. More precisely, it will say that for every simple function f and any $\varepsilon > 0$, there is a $\delta > 0$ such that whenever g is a simple function such that $|f - g| < \delta$ everywhere, then $|Ef - Eg| \leq \varepsilon$.

Technically, this means that E is a continuous function on the space of simple functions with the maximum-difference metric $d(f,g) = \max_\omega |f(\omega) - g(\omega)|$.

Proposition 2. *If P satisfies Zero, Non-negativity, Normalization and Monotonicity, then $\mathrm{LSI}_P^{\uparrow}$ and $\mathrm{LSI}_P^{\pm}$ are continuous.*

The proof is given in the Appendix.

4.5. **Sensitivity to details and Independence.** Observe that the decision theories based on the previsions $\mathrm{LSI}_P^{\uparrow}$ and $\mathrm{LSI}_P^{\pm}$ are sensitive to all the details of P, given Zero and Normalization. For we recover the probability of an event E from binary decisions recommended by LSI_P (for wagers with non-negative utility functions) as follows. Given that $\mathrm{LSI}_P\, 1_A = P(A)$ and $\mathrm{LSI}_P(c \cdot 1_\Omega) = c$ (see Lemma 1), we are required to choose the wager $\langle\langle \Omega, c\rangle\rangle$ over the wager $\langle\langle A, 1\rangle\rangle$ if and only if $c > P(A)$, so the value of $P(A)$ can be read off from the pattern of decision-theoretic requirements based on LSI_P. Hence, no details of the credences are lost in using LSI_P as the prevision, and we have:

Proposition 3. *If P and P' both satisfy Zero and Normalization, and LSI_P and $\mathrm{LSI}_{P'}$ impose the same single-shot decision-theoretic requirements for decisions involving non-negative utility functions, then $P = P'$.*

Proposition 3 has the following consequence. If P' is consistent and P is not consistent, then $\mathrm{LSI}_{P'}$ and LSI_P impose different single-shot decision-theoretic requirements on non-negative utility functions. But $\mathrm{LSI}_{P'}$ for consistent P' is the same as the classical expected utility $E_{P'}$. Thus:

Corollary 3. *If P satisfies Zero and Normalization, but is not consistent, then there is no consistent P' such that LSI_P imposes the same single-shot decision-theoretic requirements for decisions involving non-negative simple functions as $E_{P'}$ does.*

[4]If $\alpha < 1$, then $P^\alpha(A) + P^\alpha(\Omega - A) > P(A) + P(\Omega - A) = 1$ and if $\alpha > 1$, then the inequality reverses.

And note that Shifted and Split Level Set Integrals are the same as Level Set Integrals for non-negative simple functions. Thus, in using either sort of Level Set Integrals, we aren't simply cheating and in effect replacing an inconsistent credence with a consistent one. We really have a fairly natural decision that is sensitive to the details of the inconsistent credences.

It follows that decision theory based on $\mathrm{LSI}^{\pm}_P$ or $\mathrm{LSI}^{\uparrow}_P$, for inconsistent P satisfying Zero and Normalization, must reject one or more of the axioms of the von Neumann-Morgenstern (1953) Representation Theorem.

Which axiom or axioms? Well, it is trivial to see that Completeness (either f is at least as good as g or vice versa or both) and Transitivity (if h is better than g and g is better than f, then h is better than f) hold.

If we further assume Monotonicity of P, then we have Continuity. For Continuity requires that if $E_P f < E_P g < E_P h$ for simple f, g and h, then there is a number $0 \le p \le 1$ such that $E_P(pf + (1-p)h) = E_P g$. To see that this is true, let $\phi(p) = E_P(pf + (1-p)h)$. Given that f and h are simple and hence bounded, $p \mapsto pf + (1-p)h$ is a continuous function from the interval $[0,1]$ to the metric space of utility functions with the maximum-distance norm. Since E_P is a continuous function on that metric space by Proposition 2 given Monotonicity (which with Zero implies Non-negativity), it follows that ϕ is continuous. Since $\phi(0) = E_P f$ and $\phi(1) = E_P h$ and $E_P g$ lies between these two values, then by the Intermediate Value Theorem, there is a $0 \le p \le 1$ such that $\phi(p) = E_P g$, which implies von Neumann-Morgenstern's Axiom of Continuity.

By the Representation Theorem, it follows that the remaining axiom, namely Independence, must be violated when we have Monotonicity but lack Finite Additivity. The Independence Axiom says that for any simple f, g and h and any $0 < p \le 1$, the utility f is at least as choiceworthy as g if and only if $pf + (1-p)h$ is at least as choiceworthy as $pg + (1-p)h$. In other words, we have shown:

Proposition 4. *If P satisfies Zero and Normalization, then the von Neumann-Morgenstern Axioms of Completeness and Transitivity hold for $\mathrm{LSI}^{\pm}_P$ and $\mathrm{LSI}^{\uparrow}$. If P further satisfies Monotonicity, then $\mathrm{LSI}^{\pm}_P$ and $\mathrm{LSI}^{\uparrow}$ satisfy Continuity, and if, even further, P does not satisfy Finite Additivity, then $\mathrm{LSI}^{\pm}_P$ and $\mathrm{LSI}^{\uparrow}$ do not satisfy Independence.*

4.6. **Multi-Shot Domination Failure.** Matters are more complicated for multi-shot domination for individual or cumulative decision theory. First, let's define some terms. Suppose that there is a fixed sequence of choices that you are being offered, each choice being to select one of a non-empty finite set of wagers, with the set of wagers available at step n not being dependent on what one chose earlier (but perhaps different from the sets of wagers available at other steps). Then under individual decision theory corresponding to a prevision E, at each step you are permitted to choose any wager that maximizes the E-prevision of the wager from among all the ones available at that step, and required to choose some such maximizing wager.

Under cumulative decision theory, at each step n instead of maximizing the prevision of the wagers available at that step, you maximize $U_{n-1} + U_W$ where U_{n-1} is the sum of the utility functions of the wagers accepted at previous steps and W ranges over the wagers available at the current step.

It would be very nice to have the following property for cumulative decision theory: given a fixed sequence of choices, you are never permitted to select a finite sequence of wagers that turns out to be dominated by some other finite sequence of wagers you could have chosen. (This property is a stronger version of the negation of what Pettigrew 2019 calls "finite exploitability".)

Unfortunately, our Level Set previsions do not have this property. In fact, we have a general theorem on which any individual or cumulative decision theory that escapes multi-step domination is very close to the classical expected value prevision based on a consistent probability function.

Say that a *qualitative prevision* is a total preorder $\preceq$ on the set of simple functions on Ω (i.e., $\preceq$ is transitive and reflexive, and for any f and g we have the totality condition that $f \preceq g$ or $g \preceq f$). Write $f \prec g$ just in case $f \preceq g$ but not $g \preceq f$. For a single-shot choice between finite non-empty set S of utility functions (representing wagers), we can say that $\preceq$ permits the choice of U from among the members of S provided that $U' \preceq U$ for all $U' \in S$.

We can then extend this to individual and cumulative decision theory for sequences of choices. If at step n we are to choose from the utility functions in S_n, then on individual decision theory, $\preceq$ permits the choice of U if and only if $U' \preceq U$ for all $U' \in S_n$, and on cumulative decision theory, it permits the choice of U if and only if $U_{n-1} + U' \preceq U_{n-1} + U$ for all $U' \in S_n$, where U_{n-1} is the sum of the utility functions already chosen at steps prior to n.

Now, define the quantitative prevision $E_{\preceq}$ corresponding to $\preceq$ as follows: $E_{\preceq} f = \sup\{c : c \cdot 1_\Omega \preceq f\}$.

The following is easy to check:

Proposition 5. *If $f \preceq g$, then $E_{\preceq} f \leq E_{\preceq} g$. If $E_{\preceq} f < E_{\preceq} g$, then $f \prec g$.*

This means that a binary single-shot decision permitted by $\preceq$ is permitted by $E_{\preceq}$, and a binary single-shot decision required by $E_{\preceq}$ is required by $\preceq$. However, in general, $\preceq$ might be stricter than $E_{\preceq}$. For instance, suppose we start with classical expected utility E for a fair coin flip case, and say that $f \prec g$ for two utility functions f and g just in case either (a) $Ef < Eg$ or (b) $Ef = Eg$ but g assigns a higher value to tails than f does. Then it is easy to see that $E_{\preceq}$ is our old classical expected utility, but a decision theory based on $\preceq$ is more discriminating than one based on $E_{\preceq}$, as it has a tie-breaker rule for cases of equal expected utility, namely it favors the utility function that assigns a higher value to tails.

Now, suppose we have a (quantitative) prevision E. Define a qualitative prevision by saying that $f \preceq_E g$ if and only if $Ef \leq Eg$. Then the following is very easy to check:

Proposition 6. *For any quantitative prevision E such that $E(c \cdot 1_\Omega) = c$ for all c, we have $E_{\preceq_E} f = Ef$ for all simple f.*

A *binary choice* is a choice between exactly two wagers.

Theorem 3. *The following are equivalent for a qualitative prevision $\preceq$:*

(a) *$E_\preceq$ is expected utility with respect to the finitely-additive probability measure $P_{E_\preceq}$ defined by $P_{E_\preceq}(A) = E_\preceq(1_A)$*

(b) *there is no sequence of three binary choices such that $\preceq$ with individual decision theory permits one to make a sequence of selections that is dominated by another sequence of selections one could have made*

(c) *there is no sequence of three binary choices such that $\preceq$ with cumulative decision theory permits one to make a sequence of selections that is dominated by another sequence of selections one could have made.*

The proof is given in the Appendix.

Corollary 4. *Suppose P satisfies Zero and Normalization but not Finite Additivity. Let E be $\mathrm{LSI}^\uparrow_P$ or $\mathrm{LSI}^\pm_P$. Then there is a sequence of three binary choices such that E with individual decision theory permits one to make a sequence of selections that is dominated by another sequence of selections one could have made, and the same is true for cumulative decision theory (with perhaps a different sequence of choices).*

Proof. The prevision E satisfies the condition in Proposition 6. Thus $P_{E_{\preceq_E}}(A) = E_{\preceq_E}(1_A) = E(1_A) = P(A)$ by Lemma 1. Since P does not satisfy Finite Additivity, the conclusions of the Corollary follow from Theorem 3 applied to $\preceq_E$. □

That said, if E is either $\mathrm{LSI}^\uparrow$ or $\mathrm{LSI}^\pm$, and P satisfies Zero, Normalization, Monotonicity and, in the $\mathrm{LSI}^\pm$ case, Binary Non-Disappearance, by Theorem 2 there are no weak failures of single-shot domination. Because of this, there is a decision procedure for multi-shot cases that is not subject to domination failures as long as the agent knows ahead of time what choices she will be offered in the future. For in that case, the agent can simply ahead of time figure out a sequence of wagers with maximal E-prevision of the sum of the corresponding utility functions, and simply stick to the wagers of this sequence once the choices are to be made. The monotonicity of E then guarantees that no other sequence of wagers would have a dominating sum of utility functions. We can stipulate that *global decision theory* is the theory that permits one to choose any such E-maximizing wager sequence and requires one to choose some such.

This won't work if the agent does not know what choices will be offered in the future. Thus, there is a definite advantage for the expected utility maximizer working with consistent credences: she has a procedure that assures her of non-domination in a sequence of choices even when she does not know ahead of time what the future options will be. In fact, her expected utility maximization procedure does not even require her to remember what wagers she already accepted (in the way that cumulative decision theory requires), since in the classical case independent and cumulative decision theory are equivalent. Theorem 4 and the proof of the Corollary tell us that there isn't going to be a quantitative prevision E satisfying the reasonable condition $E(c \cdot 1_\Omega) = c$ that allows cumulative or independent deciders to avoid domination in multi-shot cases, unless E just is mathematical expectation with respect to some consistent probability measure.

This advantage could be used in place of the invocation of Dutch Books in arguments for probabilism. However, the resulting arguments would not, I think, be very compelling. The expected utility maximizer with consistent credences is shown to have an advantage only in a very special case, that of sequences of choices such that:

(a) what future choices are offered is not statistically dependent on what the agent has opted for so far, but
(b) the agent does not know what future choices will be offered.

For brevity, let us now say that P is a monotonic credence assignment if and only if it satisfies Zero, Normalization and Monotonicity. If one drops condition (b), an inconsistent reasoner can avoid domination by adopting global decision theory for inconsistent but monotonic credence assignments P, and using $\mathrm{LSI}_P^\uparrow$. In that case, rather than us having an argument for consistency, we have an argument for monotonicity. On the other hand, if one drops condition (a), then the situation becomes a two-player game, and single-step expected utility maximization, even in the classical consistent case, is too simplistic a strategy to guarantee doing well in such a game. For instance, we could imagine that we are playing chess with a tyrant, where we are killed if we lose, but also at earlier steps we get minor prizes or punishments depending on what move we select. In that case, we definitely should not decide what to do by maximizing the expected utility of the immediate payoff of the current step or even of the current step plus the earlier ones: rather, we need a holistic strategy that maximizes our chance at winning the whole game.

Admittedly, our holder of inconsistent monotonic credences will need to have some holistic strategy even when what future choices are offered one does not depend on one's past selection, while the holder of consistent credences will need such a strategy only in cases of such dependence. But this does not seem to be a sufficiently glaring difference to make it intuitively clear that the inconsistency is irrational.[5]

[5]It is worth noting that while Pettigrew (2019) discusses something similar to multi-shot domination failure in his critique of Hedden (2013), what makes his case against Hedden most compelling is the single-shot domination failure that Pettigrew identifies.

It is worth noting at this point that there is a similarity between our decision theory for monotonic credences and Lara Buchak's (2013) risk-based decision theory, which also rejects Independence. Indeed, Buchak's way of calculating risk-modified expected utilities is very similar to Shifted Level Set Integrals, and likewise strongly favors thinking about decisions in a temporally global way.

There is an alternative to the temporally-global decision-making, however.[6] Instead of thinking of the *decisions* as between life-long strategies to stick to, one can think of the *utilities* as including all future outcomes, not just the immediate future outcomes. Indeed, that is a very common way to think about things. Thus, in the chess game with the tyrant, at every step, one can take into account the disutility of being killed should one lose, and calculate one's expected utilities taking into account the probability distribution over one's and one's opponent's future moves conditionally on the various moves one is choosing between now. This yields another decision theory that one can plug various previsions into while asking whether unfortunate consequences for multi-step decisions follow given probabilistic inconsistencies. That is a very interesting question for future investigation. But at least we now know that there is *a* fairly natural strategy—the global decision one—that avoids many of the problems that have been alleged for inconsistent probabilities given some assumptions on the credences.

5. Lebesgue Sums

There is another prevision that might seem to some readers more natural than LSI or its variants. Often, but not always[7], the Lebesgue integral mathematical expectation of a non-negative measurable function is defined by approximating the measurable function with simple functions, and evaluating their expectation using the rule

$$\mathrm{Leb}_P\, f = \sum_{y \in Rf} y P(f = y),$$

where $Rf = \{y : \exists\omega(f(\omega) = y)\}$ is the range of f. If f is simple, Rf is finite and hence the sum is well-defined. This sum remains well-defined for simple f (even if it's not everywhere non-negative) and an inconsistent P. Thus, we have a Lebesgue Sum prevision on the space of simple functions that *prima facie* competes with our Level Set previsions.

The Lebesgue Sum prevision is equivalent to Pettigrew's (2019) extension of Hedden's (2013) $\mathrm{MSEU}_{\mathrm{bet}}$ to more complex wagers than those considered by Hedden.

Pettigrew (2019) has shown that Lebesgue Sums suffer from single-shot domination failure unless the credence is a scaled version of a consistent credence. That may look like a good reason to think that our Level Set

[6] I am grateful to a referee for pointing this out.

[7] Lieb and Loss (2013) proceed via Level Set Integrals, as we have already noted.

previsions are superior, since for a broad class of inconsistent credences they do not suffer from such a failure. But if that was the only advantage that Level Set Integrals had over Lebesgue Sums, we might think that we are *ad hoc* opting for Level Set Integrals to avoid domination failure. But there is another intuitive advantage that Level Set Integrals have, namely that Proposition 2 shows that they are continuous for all monotonic credence functions, whereas:

Proposition 7. *Suppose P fails Finite Additivity. Then* Leb_P *fails continuity.*

The proof of Proposition 7 is given in the Appendix.

6. Conclusions

Level Set Integral decision procedures based on inconsistent but monotonic credence assignments nonetheless escape Dutch Book arguments and single-shot domination failure. Thus, arguments for consistency on the basis of Dutch Books and single-shot domination failure are mistaken, in that they wrongly assume one particular way of computing the value of wagers—the De Finetti prevision. But an agent with inconsistent credence assignments is apt to see the De Finetti prevision as inappropriate, because it assigns different values to equivalent wagers, and the Level Set previsions are superior to the De Finetti prevision given inconsistency. There is also good reason to think they are superior to a prevision defined in terms of Lebesgue Sums.

The Level-Set Integral procedures—and any procedures like them—will still fall prey to a certain kind of multi-stage domination failure. But that domination failure does not seem striking enough to provide a powerful argument for probabilism. It appears that Dutch Book and domination arguments are not successful in establishing that agents should have consistent credences.[8]

Appendix: Some proofs

Proof of Theorem 2. Clearly, (a)⇒(b). We now prove the converse. Assume (b). Suppose f and g are simple functions such that $0 \le f < g$ everywhere. Then there is an $\varepsilon > 0$ such that $\varepsilon + f < g$ everywhere (e.g., let $\varepsilon = (1/2)\min(g - f)$). By weak monotonicity, $\mathrm{LSI}(\varepsilon + f) \le \mathrm{LSI}\, g$. By Lemma 1, we have $\mathrm{LSI}(\varepsilon + f) = \varepsilon + \mathrm{LSI}\, f$. Thus, $\varepsilon + \mathrm{LSI}\, f \le \mathrm{LSI}\, g$, so $\mathrm{LSI}\, f < \mathrm{LSI}\, g$ as $\varepsilon > 0$. Hence, we have shown (a)⇔(b).

Clearly, (c)⇒(b). For the converse, suppose (b) and suppose that $0 \le f \le g$ everywhere. Then for any fixed $\varepsilon > 0$, we have $0 \le f < \varepsilon + g$ everywhere. By the equivalence of (a) and (b), LSI is strongly monotonic on non-negative functions, so $\mathrm{LSI}\, f \le \mathrm{LSI}(\varepsilon + g) = \varepsilon + \mathrm{LSI}\, g$ (by Lemma 1). Since this is

[8] I am grateful to Kenny Easwaran, Yoaav Isaacs and Ian Slorach for discussion of these topics, to two anonymous readers for improving the presentation and fixing a number of errors, and to one of these referees for pointing out interesting philosophical issues.

true for any $\varepsilon > 0$, it must be true that $\mathrm{LSI}\, f \le \mathrm{LSI}\, g$, completing the proof of (b)$\Leftrightarrow$(c).

Observe that (c) and (d) are equivalent, since $\mathrm{LSI}^{\uparrow} f = \mathrm{LSI}(M + f)$ and $\mathrm{LSI}^{\uparrow} g = \mathrm{LSI}(M + g)$ if M is sufficiently large that both $M + f$ and $M + g$ are non-negative.

Now, assume (a). Let f and g be any simple functions such that $f < g$ everywhere and suppose $M + f$ is non-negative. Then $M + g$ is non-negative and $M + f < M + g$ everywhere. Hence $\mathrm{LSI}(M + f) < \mathrm{LSI}(M + g)$ by (a). Hence, $\mathrm{LSI}^{\uparrow} f < \mathrm{LSI}^{\uparrow} g$, and so we have shown (a)$\Rightarrow$(e).

Clearly, (e)$\Rightarrow$(f)$\Rightarrow$(b) (we are using Lemma 1 again). Hence, (a)–(f) are all equivalent.

Next, (g)$\Rightarrow$(b) by Lemma 1. Now assume (c), and we will prove (g). Suppose $f < g$ everywhere. Then $f^+ \le g^+$ and $f^- \ge g^-$ everywhere, so by (c) we have $\mathrm{LSI}\, f^+ \le \mathrm{LSI}\, g^+$ and $\mathrm{LSI}\, f^- \ge \mathrm{LSI}\, g^-$. Since $\mathrm{LSI}^{\pm} f = \mathrm{LSI}\, f^+ - \mathrm{LSI}\, f^-$ and likewise for g, it follows that $\mathrm{LSI}^{\pm} f \le \mathrm{LSI}^{\pm} g$, so $\mathrm{LSI}^{\pm}$ is weakly monotonic, which completes the proof of (g). Hence, (a)–(g) are all equivalent.

Now suppose that P fails Monotonicity. Then there are events $A \subseteq B$ such that $P(A) > P(B)$. Note that $1_A \le 1_B$ everywhere, and $\mathrm{LSI}_P\, A = P(A) > P(B) = \mathrm{LSI}_P\, B$, so condition (c) fails. Hence we have shown that (c)$\Rightarrow$(h).

Next, assume P satisfies Monotonicity and $0 \le f \le g$. Then $\{\omega : f(\omega) > y\} \subseteq \{\omega : g(\omega) > y\}$, so by Monotonicity $P(f > y) \le P(g > y)$ for y. Integrating over y, we get $\mathrm{LSI}_P\, f \le \mathrm{LSI}_P\, g$. Hence (h)$\Rightarrow$(c), and so we have by now shown that (a)–(h) are equivalent.

It remains to show that (i) and (j) are equivalent. First, assume (i). That P satisfies Monotonicity follows by Lemma 1 from the equivalence of (a) and (h). Now suppose that Binary Non-Disappearance fails so we have $P(A) = P(B) = 0$ for disjoint A and B whose union is Ω. Let $f = 1_A - 1_B$ and let $g = \frac{1}{2} + f$. Then $f < g$ everywhere, but it is easy to see that $P(f > 0) = P(g > 0) = P(A) = 0$ and $P(f < 0) = P(g < 0) = P(B) = 0$. By Monotonicity and Non-negativity, it follows that $P(f < -y) = P(g < -y) = P(f > y) = P(g > y) = 0$ for all $y \ge 0$, and so $\mathrm{LSI}_P^{\pm} f = 0 = \mathrm{LSI}_P^{\pm} g$, contrary to strong monotonicity of $\mathrm{LSI}_P^{\pm}$. Hence, (i)$\Rightarrow$(j).

Finally, assume (j). Demonstrating (i) will be the hardest part of the proof of the Theorem. To obtain a contradiction, suppose that P satisfies Monotonicity and Binary Non-Disappearance, but $\mathrm{LSI}_P^{\pm}$ fails to be strongly monotonic. Thus, there will be f_1 and f_2 such that $f_1 < f_2$ everywhere but $\mathrm{LSI}_P^{\pm} f_1 \ge \mathrm{LSI}_P^{\pm} f_2$. Then:

$$\int_0^\infty P(f_1 > y)\, dy - \int_0^\infty P(f_1 < -y)\, dy \\ \ge \int_0^\infty P(f_2 > y)\, dy - \int_0^\infty P(f_2 < -y)\, dy.$$

Hence,

$$(2)\quad \int_0^\infty [P(f_1 > y) - P(f_2 > y)]\,dy \geq \int_0^\infty [P(f_1 < -y) - P(f_2 < -y)]\,dy.$$

Now, as $f_1 < f_2$, if $f_1(\omega) > y$ then $f_2(\omega) > y$ so by Monotonicity $P(f_1 > y) - P(f_2 > y) \leq 0$, while if $f_2(\omega) < -y$, then $f_1(\omega) < -y$, so by Monotonicity $P(f_1 < -y) - P(f_2 < -y) \geq 0$. Thus the left-hand-side of (2) is non-positive and the right-hand-side is non-negative. Hence both sides must be zero. But the integrands on both sides are everywhere non-positive and non-negative respectively, so the integrands must be zero for almost all values of y (i.e., for all y other than in some Lebesgue null set). But it is easy to see that the functions $y \mapsto P(f_i > y)$ and $y \mapsto P(f_i < -y)$ are right-continuous when f_i is simple, hence the integrands in (2) are right-continuous and zero almost everywhere, which implies they are zero everywhere (if g is right-continuous and zero almost everywhere, then for any any x there is a sequence of x_n approaching x from the right such that $g(x_n) = 0$; by right-continuity, $g(x) = \lim_{n\to\infty} g(x_n) = 0$).

Thus, $P(f_1 > y) = P(f_2 > y)$ and $P(f_1 < -y) = P(f_2 < -y)$ for all $y \geq 0$.

Let $\varepsilon = \min_{\omega\in\Omega}(f_2(\omega) - f_1(\omega))$. (This is positive and defined as $f_2 > f_1$ and both functions are simple.) Then, $f_2(\omega) \geq f_1(\omega) + \varepsilon$. Hence, by Monotonicity of P, $P(f_2 > y) \geq P(f_1 > y - \varepsilon) \geq P(f_1 > y)$ for any $y \geq 0$. But $P(f_2 > y) = P(f_1 > y)$, so for all $y \geq 0$, we have $P(f_1 > y) = P(f_1 > y - \varepsilon)$ and hence for all $y \geq -\varepsilon$ we have $P(f_1 > y) = P(f_1 > y + \varepsilon)$. Iterating starting with $y = -\varepsilon$ we get:

$$P(f_1 > -\varepsilon) = P(f_1 > 0) = P(f_1 > \varepsilon) = P(f_1 > 2\varepsilon) = \cdots .$$

Thus, $P(f_1 > -\varepsilon) = P(f_1 > n\varepsilon)$ for any natural number n. Choosing n large enough that f_1 be nowhere greater than $n\varepsilon$, we will have $P(f_1 > n\varepsilon) = 0$ and hence $P(f_1 > -\varepsilon) = 0$. By Monotonicity and Non-negativity of P, we have $P(f_1 \geq 0) = 0$.

Similarly, $P(f_2 < -y) \leq P(f_1 + \varepsilon < -y) \leq P(f_1 < -y)$ for all $y \geq 0$. Since $P(f_2 < -y) = P(f_1 < -y)$, it follows that $P(f_1 < -y - \varepsilon) = P(f_1 < -y)$ for all $y \geq 0$. Again, iterating:

$$P(f_1 < \varepsilon) = P(f_1 < 0) = P(f_1 < -\varepsilon) = P(f_1 < -2\varepsilon) = \cdots .$$

Thus, $P(f_1 < 0) = P(f_1 < -n\varepsilon)$ for any natural number n. Choosing n large enough, by Zero we can ensure $P(f_1 < -n\varepsilon) = 0$, and hence $P(f_1 < 0) = 0$.

Letting $A = \{\omega : f_1(\omega) \geq 0\}$ and $B = \{\omega : f_1(\omega) < 0\}$, we have A and B disjoint, with union Ω and with $P(A) = P(B) = 0$, contradicting Binary Non-Disappearance. □

Proof of Proposition 2. First, observe that if f and g are non-negative and $|f - g| < \varepsilon/2$ everywhere, then

$$P(f - \varepsilon/2 > y) \leq P(g > y) \leq P(f + \varepsilon/2 > y)$$

by Monotonicity, since if $f(\omega) - \varepsilon/2 > y$, then $g(\omega) > y$ and if $g(\omega) > y$, then $f(\omega) + \varepsilon/2 > y$. But

$$\int_0^\infty P(f - \varepsilon/2 > y)\,dy = \int_{\varepsilon/2}^\infty P(f > y)\,dy \geq \int_0^\infty P(f > y) - \varepsilon/2,$$

where the last inequality used Normalization and Monotonicity which together show that $P(A) \leq 1$ always. Moreover

$$\int_0^\infty P(f + \varepsilon/2 > y)\,dy = \int_{-\varepsilon/2}^\infty P(f > y)\,dy \leq \int_0^\infty P(f > y) + \varepsilon/2.$$

It follows that

$$\operatorname{LSI} f - \varepsilon/2 \leq \operatorname{LSI} g \leq \operatorname{LSI} f + \varepsilon/2$$

and hence that $|\operatorname{LSI} f - \operatorname{LSI} g| \leq \varepsilon$.

Now suppose f and g are any simple functions such that $|f - g| < \varepsilon/2$ everywhere. Let c be a constant such that $c + f$ and $c + g$ are both non-negative. Then

$$|\operatorname{LSI}^\uparrow f - \operatorname{LSI}^\uparrow g| = |\operatorname{LSI}(c + f) - \operatorname{LSI}(c + g)| \leq \varepsilon,$$

by what we have proved above, since if $|f - g| < \varepsilon/2$, likewise $|(c + f) - (c + g)| < \varepsilon/2$. Letting $\delta = \varepsilon/2$, we see that the definition of continuity is satisfied in the case of $\operatorname{LSI}^\uparrow$.

Now, suppose that f and g are any simple functions such that $|f-g| < \varepsilon/4$ everywhere. Then it is easy to see that $|f^+ - g^+| \leq |f - g| < \varepsilon/4$ and $|f^- - g^-| \leq |f - g| < \varepsilon/4$ everywhere. It follows from what was proved above and by the triangle inequality that:

$$\begin{aligned}
|\operatorname{LSI}^\pm f - \operatorname{LSI}^\pm g| &= |\operatorname{LSI} f^+ - \operatorname{LSI} g^+ - (\operatorname{LSI} f^- - \operatorname{LSI} g^-)| \\
&\leq |\operatorname{LSI} f^+ - \operatorname{LSI} g^+| + |\operatorname{LSI} f^- - \operatorname{LSI} g^-| \\
&\leq \varepsilon/2 + \varepsilon/2 = \varepsilon.
\end{aligned}$$

Letting $\delta = \varepsilon/2$, we see that the definition of continuity is satisfied in the case of $\operatorname{LSI}^\pm$. □

Suppose $\preceq$ is a qualitative prevision and let $P(A) = E_\preceq(1_A)$.

Lemma 2. *Suppose that $\preceq$ satisfies strong monotonicity. Then $E_\preceq(c \cdot 1_\Omega) = c$ for all c, and P satisfies Zero, Non-negativity and Normalization. Moreover, for any simple f and any real $\varepsilon > 0$, we have:*

$$(-\varepsilon + E_\preceq f) \cdot 1_\Omega \prec f \prec (\varepsilon + E_\preceq f) \cdot 1_\Omega.$$

Proof. Let $f = c \cdot 1_\Omega$ and $g = c' \cdot 1_\Omega$. If $f \preceq g$, we don't have $f > g$ everywhere, and hence we must have $c \leq c'$. The converse is also true. For if $c \leq c'$, then by totality of $\preceq$, either $g \preceq f$ or $f \preceq g$. If we have $g \preceq f$, then by the above argument with g and f swapped, we conclude that $c' \leq c$ and hence $c = c'$ and $f = g$ and so $f \preceq g$ by reflexivity. So either way, $f \preceq g$.

It follows from this and the definition of $E_\preceq$ that $E_\preceq(c \cdot 1_\Omega) = c$. Now, $1_\varnothing = 0 \cdot 1_\Omega$, so $P(\varnothing) = 0$, and $P(\Omega) = E_\preceq(1 \cdot 1_\Omega) = 1$. This yields Zero and Normalization. To prove Non-negativity, choose an event A and fix any $\varepsilon > 0$. Observe that $1_A > -\varepsilon \cdot 1_\Omega$. Therefore, by strong monotonicity, we cannot have $1_A \preceq -\varepsilon \cdot 1_\Omega$, and so we must have $-\varepsilon \cdot 1_\Omega \preceq 1_A$ by totality. Therefore, $E_\preceq(1_A) \geq -\varepsilon$. Since this is true for all $\varepsilon > 0$, it follows that $E_\preceq(1_A) \geq 0$ and hence we have Non-negativity.

Now, to obtain a contradiction suppose that we do not have $(-\varepsilon + E_\preceq f) \cdot 1_\Omega \prec f$. Then by totality $f \preceq (-\varepsilon + E_\preceq f) \cdot 1_\Omega$. Choose c such that $c \cdot 1_\Omega \preceq f$ and $c > E_\preceq f - \varepsilon/2$ (such a c exists by definition of the supremum). Then $c \cdot 1_\Omega \preceq (-\varepsilon + E_\preceq f) \cdot 1_\Omega$. Hence, by strong monotonicity, $c \leq -\varepsilon + E_\preceq f$. Thus, $E_\preceq f - \varepsilon/2 < c \leq -\varepsilon + E_\preceq f$, which is impossible. Thus, indeed, $(-\varepsilon + E_\preceq f) \cdot 1_\Omega \prec f$.

Finally, suppose we do not have $f \prec (\varepsilon + E_\preceq f) \cdot 1_\Omega$. Then, $(\varepsilon + E_\preceq f) \cdot 1_\Omega \preceq f$, and hence $\varepsilon + E_\preceq f \leq E_\preceq f$ by definition of $E_\preceq f$. And that's impossible. □

Say that $\preceq$ satisfies the *sixtuple condition* if there are no f_1, f_2, f_3, g_1, g_2 and g_3 such that $f_i \preceq g_i$ for $i = 1, 2, 3$, but $f_1 + f_2 + f_3 > g_1 + g_2 + g_3$ everywhere. Say $\preceq$ is strongly monotonic provided that if $f \preceq g$, then we cannot have $f > g$ everywhere. It is clear that the sixtuple condition implies strong monotonicity (just let $f_i = g_i = 0$ for $i = 2, 3$).

Lemma 3. *Suppose that $\preceq$ satisfies the sixtuple condition. Then, P is a finitely additive probability.*

Proof. Lemma 2 gives us Non-negativity and Normalization. Finite Additivity remains to be proved. Suppose A and B are disjoint. Fix any $\varepsilon > 0$. Let $f_1 = 1_A$, $f_2 = 1_B$ and $f_3 = -1_{A\cup B}$. Then by Lemma 2 and the definition of P:

$$(-\varepsilon + P(A)) \cdot 1_\Omega \prec f_1 \prec (\varepsilon + P(A)) \cdot 1_\Omega$$
$$(-\varepsilon + P(B)) \cdot 1_\Omega \prec f_2 \prec (\varepsilon + P(B)) \cdot 1_\Omega$$
$$(-\varepsilon - P(A \cup B)) \cdot 1_\Omega \prec f_3 \prec (\varepsilon - P(A \cup B)) \cdot 1_\Omega.$$

Observe that $f_1 + f_2 + f_3 = 0$ everywhere. By two applications of the sixtuple condition, it follows that

$$-3\varepsilon + P(A) + P(B) - P(A \cup B) \leq 0 \leq 3\varepsilon + P(A) + P(B) - P(A \cup B).$$

The only way this can be true for all $\varepsilon > 0$ is if $P(A) + P(B) - P(A \cup B) = 0$. Hence Finite Additivity holds. □

Proof of Theorem 3. If (a) is true, then (b) and (c) are true. The individual and cumulative decision methods are equivalent given the mathematical expectation prevision and consistent credences. Mathematical expectation is additive, i.e., $E(f+g) = Ef + Eg$, and so if one maximizes utility expectation in each choice, one maximizes the expectation of the sum of the utilities. But if $E(f_1 + f_2 + f_3) \leq E(g_1 + g_2 + g_3)$, then we cannot have $g_1 + g_2 + g_3 < f_1 + f_2 + f_3$ everywhere. For if we had that, then there would be an $\varepsilon > 0$ such that $\varepsilon + g_1 + g_2 + g_3 < f_1 + f_2 + f_3$ everywhere (this follows from the fact that we are dealing solely with functions that take on a finite number of values). But then we would have $E(\varepsilon + g_1 + g_2 + g_3) \le E(f_1 + f_2 + f_3)$ and hence $\varepsilon + E(g_1 + g_2 + g_3) \le E(f_1 + f_2 + f_3)$ which is impossible if $E(f_1 + f_2 + f_3) \le E(g_1 + g_2 + g_3)$.

It remains to show that each of (b) and (c) implies (a). By Lemma 3 all we need to show is that each of (b) and (c) implies the sixtuple condition.

Assume (b). Suppose that we have a violation of the sixtuple condition, so $f_i \preceq g_i$ for $i = 1, 2, 3$ but $f_1 + f_2 + f_3 > g_1 + g_2 + g_3$ everywhere. Consider an agent choosing between f_1 and g_1, then between f_2 and g_2, and then between f_3 and g_3. Individual decision theory permits the agent to choose g_i in each case. But since $f_1 + f_2 + f_3 > g_1 + g_2 + g_3$ everywhere, this would result in the agent getting dominated, which would violate (b).

The case of (c) is a little trickier. To obtain a contradiction, assume (c) and suppose that we have a violation of the sixtuple condition, so $f_i \preceq g_i$ for $i = 1, 2, 3$ but $f_1 + f_2 + f_3 > g_1 + g_2 + g_3$ everywhere. Now, let

$$\begin{aligned} F_1 &= f_1 \\ G_1 &= g_1 \\ F_2 &= f_2 - G_1 \\ G_2 &= g_2 - G_1 \\ F_3 &= f_3 - G_1 - G_2 \\ G_3 &= g_3 - G_1 - G_2. \end{aligned}$$

Suppose our agent is offered binary choices between F_i and G_i, for $i = 1, 2, 3$, at the three successive stages. Cumulative decision theory then permits the choice of G_1 over F_1 at the first step, since $F_1 = f_1 \preceq g_1 = G_1$. At the next stage, given the prior choice of G_1, it permits the choice of G_2 over F_2 since $G_1 + F_2 = f_2 \prec g_2 = G_1 + G_2$. Given the agent's having chosen G_1 and G_2, at the final stage the agent will permitted to choose G_3 over F_3 since $G_1 + G_2 + F_3 = f_3 \preceq g_3 = G_1 + G_2 + G_3$.

But

$$f_1 + f_2 + f_3 = F_1 + (F_2 + G_1) + (F_3 + G_1 + G_2) = F_1 + F_2 + F_3 + 2G_1 + G_2$$

and

$$g_1 + g_2 + g_3 = G_1 + (G_2 + G_1) + (G_3 + G_1 + G_2) = G_1 + G_2 + G_3 + 2G_1 + G_2.$$

Since we have $f_1 + f_2 + f_3 > g_1 + g_2 + g_3$ everywhere, we must also have $F_1 + F_2 + F_3 > G_1 + G_2 + G_3$, and so the sum of the $E_{\preceq}$-permissibly chosen (under cumulative decision theory) utility functions $G_1 + G_2 + G_3$ is dominated by $F_1 + F_2 + F_3$, which contradicts (c). □

Proof of Proposition 7. Suppose A and B are disjoint. Let $C = \Omega - (A \cup B)$. Suppose ε is a non-zero real number. Let $f = 1_A + (1 + \varepsilon) \cdot 1_B$ and $g = \varepsilon + 1_{A\cup B}$. Then

$$\mathrm{Leb}_P\, f = P(A) + (1 + \varepsilon)P(B)$$

and

$$\mathrm{Leb}_P\, g = (1 + \varepsilon)P(A \cup B) + \varepsilon P(C).$$

Let $\eta = \frac{1}{4}|P(A) + P(B) - P(A \cup B)|$. Let $h = 1_{A\cup B}$. Suppose that we have continuity, so there is a $\alpha > 0$ such that for any function u such that $|h - u| \leq \alpha$, we have $|\operatorname{Leb}_P h - \operatorname{Leb}_P u| \leq \eta$. Let ε be any number strictly between 0 and α. Then $|f - h| < \alpha$ and $|g - h| < \alpha$ everywhere. Then $|\operatorname{Leb}_P h - \operatorname{Leb}_P f| \leq \eta$ and $|\operatorname{Leb}_P h - \operatorname{Leb}_P g| \leq \eta$. By the triangle inequality, $|\operatorname{Leb}_P f - \operatorname{Leb}_P g| \leq 2\eta$. But

$$\operatorname{Leb}_P f - \operatorname{Leb}_P g = P(A) + (1+\varepsilon)P(B) - (1+\varepsilon)P(A \cup B) - \varepsilon P(C).$$

We saw the left-hand-side can be no bigger than 2η in absolute value for every ε strictly between 0 and α. Hence:

$$|P(A) + (1+\varepsilon)P(B) - (1+\varepsilon)P(A \cup B) - \varepsilon P(C)| \leq 2\eta.$$

Taking the limit as ε approaches 0 from above, we conclude that

$$|P(A) + P(B) - P(A \cup B)| \leq 2\eta = \frac{1}{2}|P(A) + P(B) - P(A \cup B)|.$$

The only way this can be true is if $P(A) + P(B) - P(A \cup B) = 0$. Thus, from continuity we get Finite Additivity. □

References

[1] Buchak, Lara (2013). *Risk and Rationality*, Oxford: Oxford University Press.

[2] de Finetti, Bruno (1937). "Foresight: Its logical laws, its subjective sources." In Henry E. Kyburg and Howard E.K Smokler (eds.), *Studies in Subjective Probability*, Huntington, NY: Kreiger Publishing.

[3] Hájek, Alan (2008). "Arguments for, or against, probabilism?" *British Journal for the Philosophy of Science* **59**: 793–819.

[4] Hedden, B. (2013). "Incoherence without exploitability", *Noûs* **47**: 482—295.

[5] Howson, Colin, and Peter Urbach (1993). *Scientific Reasoning: The Bayesian Approach*, second edition, La Salle, Illinois: Open Court.

[6] Kemeny, John (1955). "Fair bets and inductive probabilities", *Journal of Symbolic Logic*, 20: 263–273.

[7] Lehman, R. Sherman (1955). "On confirmation and rational betting", *Journal of Symbolic Logic* **20**: 251–262.

[8] Lieb, Elliott H. and Loss, Michael (2001). *Analysis*, second edition, Providence, RI: American Mathematical Society.

[9] Pettigrew, Richard (2019). "On the expected utility objection to the Dutch Book argument for probabilism", *Noûs* Online First.

[10] Ramsey, F. P. (1931). "Truth and probability." In: *The Foundations of Mathematics and Other Lgoical Essays*, Routledge and Kegan Paul, 156–198.

[11] von Neumann, John and Morgenstern, Oskar (1953). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.