HOW TO CO-EXIST WITH NONEXISTENT EXPECTATIONS

RANDALL G. MCCUTCHEON

Abstract. Dozens of articles have addressed the challenge that gambles having undefined expectation pose for decision theory. This paper makes two contributions. The first is incremental: we evolve Colyvan's "Relative Expected Utility Theory" into a more viable "conservative extension of expected utility theory" by formulating and defending emendations to a version of this theory proposed by Colyvan and Hájek. The second is comparatively more surprising. We show that, so long as one assigns positive probability to the theory that there is a uniform bound on the utility of possible gambles (and assuming a uniform bound on the amount of utility that can accrue in a fixed amount of time), standard principles of anthropic reasoning (as formulated by Brandon Carter) place lower and upper bounds on the expected values of gambles advertised as having no expectation–even assuming that with positive probability, all gambles advertised as having infinite expected utility are administered faithfully. Should one accept the uniform bound premises, this reasoning thus dissolves (or nearly dissolves, in some cases) several puzzles in infinite decision theory.

1. Introduction

We address, in this paper, the challenge that gambles having infinite or undefined expectation create for decision theory. Our investigation splits into two separate questions. The first question is how to respond to offers of gambles advertised as having infinite or no expectation, conditional on their being genuine, i.e. faithfully administered. In the case of infinite expectation, there is no real problem here; one should pay any (finite) price for a gamble having infinite expected payoff. However, we've found that one must forebear on ranking any no-expectation gamble relative to the status quo, lest one succumb to a Dutch Book. In the first part of the paper, we explain these matters and summarize our recommendations in an emended version of Mark Colyvan's "Relative Expectation Theory".

Later, we address the question of how to respond to an offered gamble X that is advertised as infinite or no expectation in the face of natural doubts that such gambles are possible at all. The challenge is that if one assigns positive probability to the notion that the offer of X is genuine, it may seem that one must assign X an infinite or non-existent expectation. According to the most successful version of anthropic reasoning (Brandon Carter's anthropic principle), however, self-selection effects may serve to mitigate this response. For it seems that, if one assigns any positive probability to the theory that infinite or no expectation gambles are impossible, one must on reasoning anthropically assign X a bounded posterior expectation. If one accepts the premises of this argument (one of which is that there is some upper bound on the rate at which utility may accrue), it would seem to dissolve many (if not most) puzzles in infinite decision theory at a stroke. Accordingly, we believe that it deserves careful scrutiny.

2. The Dutch Book Argument Against Assigning Values to No Expectation Gambles

As point of departure to the strange and "unexpected" results glossed above we consider in the next two sections a no expectation gamble that has been given a great deal of recent attention–Harris Nover and Alan Hájek's *Pasadena Game* (2004). Suppose a fair coin is tossed just until it first lands *heads*. Letting n be the number of tosses required, the Pasadena gamble pays $X = \frac{(-1)^{n-1}2^n}{n}$ dollars.

Since the expectation series for X,

$$\sum_{r\in\mathbf{R}} rP(X=r) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}2^n}{n} \cdot 2^{-n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots = \log 2$$

converges conditionally, X has no defined expectation.[1] Expected utility theory is, therefore, silent concerning the value of this game.[2] Nover and Hájek write:

> It is an uncomfortable silence. For intuition tells us...that we can make meaningful comparisons between the Pasadena game and other games. It is clearly worse than the St. Petersburg game [pays $Y = 2^n$ dollars], for starters. It is clearly worse than a neighbouring variant of the game –call it the Altadena game–in which every pay-off is raised by a dollar. (...) And the Pasadena game is clearly better than a 'negative' St. Petersburg game, in which all the pay-offs of the St. Petersburg game are switched in sign. Yet expected utility theory can say none of this.

Yet "Dutch Book reasoning" does seem to show that one cannot coherently value the Pasadena gamble relative to the null gamble. For suppose that you value it at $\log 2$. (It's easily checked that any value would lead to the same trouble.) Then you can be Dutch Booked as follows. First you are offered the chance to pay $.69 < \log 2$ to play the game. You accept and are put to sleep. While you sleep, your payoff will be determined as follows. First, a fair 4-sided die is rolled repeatedly until something less than a 4 is rolled. Let N be the number of rolls it takes. Repeat the procedure and let M be the number of rolls it takes the second time. Finally let Y be the result of a rolling of a fair 3-sided die. If $Y = 3$, let $n = 2M$. Otherwise, let $n = 2N - 1$. Your payoff is now $\frac{(-1)^{n-1}2^n}{n}$ dollars (a Pasadena gamble). Note: if $Y < 3$ and $n = 2N - 1$ you are winning money, whereas if $Y = 3$ and $n = 2M$ you are losing money.

[1]Like any conditionally convergent series, the expectation series can thus be made to diverge (or converge to any finite value whatsoever) by rearrangement of its terms. On the other hand, as noted by Kenny Easwaran (2008), X does have a weak expectation of $\log 2$. Since weak expectations are invariant under rearrangements, $\log 2$ therefore has some claim to be the presumptive value of the game, if it has one.

[2]We take utility to be linear with respect to currency, and in particular unbounded.

When all that's completed, you are awakened and told the value N (but not M and not Y). So you don't know whether or not you've won, nor how much you've lost if you've lost. You do however know that if you've won, you've won 2^{2N-1} dollars. At this point, you are given the opportunity to annul the gamble (but you don't get your .69 back). The expectation series for your payoff is now

$$\frac{2}{3} \cdot 2^{2N-1} + \sum_{M=1}^{\infty} \frac{1}{2^{2M}} \cdot \frac{-2^{2M}}{2M} = \frac{2}{3} \cdot 2^{2N-1} - \frac{1}{2} - \frac{1}{4} - \frac{1}{6} - \cdots = -\infty,$$

so you are compelled to accept this offer (indeed, you were let off easy–you would have paid any finite amount to annul the gamble). You have thus lost a sure .69.[3]

The lesson is this: if you assign faithful implementations of no expectation gambles values and believe you are being offered one then your decision theory is diachronically incoherent in a strong sense–you fall prey to a finitary Dutch Book.[4]

3. Other responses

If one can't place values on no expectation gambles, how should one respond to their offer? (Assume for the present that one takes them to be faithfully administered.) Hájek and Nover (2006) write:

> Here are...possible responses to the Pasadena game (...): 1. The game is coherent, and decision theory cannot handle it–too bad for decision theory. Compare: Russell's paradox was a decisive blow against Frege's set theory. Too bad for that set theory. 2. The game is incoherent, so it is not a black mark that decision theory cannot handle it–too bad for the game. Compare: the town barber, who shaves all and only those in the town who do not shave themselves, poses no problem for logic, or for anything else; he simply cannot exist, because the specification of him is incoherent. Too bad for the barber. (...) We maintain response (1)....

The admonition we perceive here is somewhat different. In fact, Russell's paradox is a decisive blow against *any* set theory that chooses to "handle" (treat as a set) the Russell object (the collection of sets that don't contain themselves). The standard response to it, accordingly, has been to deny such collections "sethood" while retaining such sets as are required by those who employ them sincerely. The issue isn't that such collections are incoherent, but that they are too exotic to be coherently treated in the manner of more conventional collections. (In some versions of modern set theory, they are termed "proper classes".) As to the Pasadena game, then, we advocate for a third response running as follows: "The

[3] What makes this Dutch Book work is that the positive and negative parts of the Pasadena variable each have infinite expectation, and the argument can be generalized to show that one cannot coherently place a value on any such variable. It is probably fair to say that the argument is implicit in the "Two Envelopes" literature; see especially Broome (1995) and Chalmers (2002).

[4] By *Finitary Dutch Book* we intend an almost surely finite sequence $X_1, \cdots, X_N$ of gambles (here N is a "stopping time"), each deemed individually favorable at time of offer (information may be obtained between gambles), but entailing an almost sure net loss, such that either (a) N is uniformly bounded, or (b) $\sum_{i=1}^{N} E(\min\{X_i, 0\}) > -\infty$ almost surely. (See the appendix.)

game is coherent, but it is too exotic to be coherently treated in the manner of more conventional gambles. So it is not a black mark that decision theory cannot handle it." We would say, in fact, that the above Dutch Book argument is a black mark against any decision theory that purports to handle it.

Succinctly, we hold that decision theory should incorporate some form of restriction on gamble formation. We don't think, however, that Hájek and Nover (2006) are thorough enough in seeking the form such restriction might take. They write: "There are two possible lines of attack–neither satisfactory, in our opinion. (...) *Restrict decision theory to finite state spaces* (...) *Restrict decision theory to bounded utility functions*." Neither of these moves is, strictly speaking, justified: if A contracts to pay B one dollar per year so long as both live, the expected gain to B in the nth year of this contract decays rapidly in n, but isn't *zero* for *any* n.

Of course, one might say that when n is large, the expected gain to B in the nth year is *negligible*. Nicholas J.J. Smith (2014) defends a decision rule that attempts to formalize this idea. As we'll see, however, the results remain unsatisfactory.

Rationally negligible probabilities (RNP): For any lottery featuring in any decision problem faced by any agent, there is an $\epsilon > 0$ such that the agent need not consider outcomes of that lottery of probability less than ϵ in coming to a fully rational decision.

For variables of finite expectation, **RNP** is harmless enough; as $\epsilon \to 0$, the relative effect of employing **RNP** instead of standard expectation tends to zero with ϵ. In this sense, **RNP** is an extension of (finite) expected utility theory. This observation can be used to respond to Hájek (2014), who seeks to discredit **RNP** using a zero expectation gamble (credited to John Matthewson) which **RNP**'s vanishing error term causes to look favorable (if infinitesimally). But although finite expectation gambles fail to discredit this more robust limit version of **RNP**, even it sanctions some embarrassing preferences.

Game: A dime is tossed until it comes up heads (on the nth toss). Then a nickel is tossed. If the nickel comes up heads, you win 2^n dollars. If it comes up tails, you lose 2^n dollars.

RNP instructs us to ignore outcomes having probability below some threshold $\epsilon > 0$. So there is an n depending on ϵ such that **Game's** value is:

$$(.25)(2)-(.25)(2)+(.125)(4)+(-.125)(4)+\cdots+(2^{-n-1})(2^n)+(-2^{-n-1})(2^n)=0.$$

Thus the value of **Game** is zero, independently of ϵ.

Now consider a variant of **Game** in which, if the nickel lands heads, a penny is also tossed. If the penny lands heads, it is added to your winnings. Otherwise everything is as before. This variant's value is:

$$(-.25)(2)+(.125)(2.01)+(.125)(2)+(-.125)(4)+(.0625)(4.01)+(.0625)(4) \\ +(.0625)(8)+\cdots+(2^{-n-1})(2^{n-1}+.01)+(2^{-n-1})(2^{n-1})+(2^{-n-1})(2^n)<-.49.$$

This inequality holds independently of ϵ.

So according to **RNP**, **Game** has value $V_1 = 0$ and the variant $V_2 < -.49$, independently of ϵ. But the only difference between **Game** and the variant is that in the latter there is a penny that you might get to keep.

4. Relative Expected Utility Theory

The failures of **RNP** and other approaches involving bounded utility functions or credence functions supported on finitely many points notwithstanding, we believe that there *is* a good way to avoid paradox by restriction. Indeed, Mark Colyvan has been championing an approach along these lines that is promising. We'll outline the evolution of Colyvan's proposal and indicate how we think it needs to be altered in order to yield a satisfactory theory.

First Colyvan pointed out (see Colyvan 2006) that a preference for Altadena over Pasadena can be established by dominance reasoning. Next (see Colyvan 2008) he formulated a "relative expected utility theory", a joint extension of finite expected utility theory and just this sort of dominance reasoning:

> **REU Dominance:** Let X be any random variable taking values in the natural numbers $\mathbf{N} = \{1, 2, \ldots\}$. Let Game A pay a_i and let Game B pay b_i when $X = i$, $i \in \mathbf{N}$. Put $REU(A, B) = \sum_{i=1}^{\infty} P(X = i)(a_i - b_i)$ if the right hand side converges or diverges to infinity. Then Game A is preferable to Game B if and only if $REU(A, B) > 0$.

In this original formulation, **REU Dominance** has several undesirable features. Among these are (a) its recommendations depend upon the way in which one orders the alternatives; and (b) it may indicate a preference for a gamble X over an identically distributed gamble Y[5]. Colyvan and Hájek (2016) accordingly make two emendations to **REU dominance**. The first is to insist that $REU(A, B)$ only be defined when the sum $\sum_{i \in \mathbf{N}} P(X = i)(a_i - b_i)$ is insensitive to the ordering of the indices.[6] (See also footnote 12 in Bartha 2016.) This fixes problem (a), and as we'll presently see (Theorem 1 below), it also fixes problem (b).

Colyvan and Hájek don't stop there, however. They proceed to introduce a second emendation. We will make a digression to discuss this second emendation (which we will ultimately reject). They motivate it by a comparison of the following bets:

Bet 1: Pays 5 if a fair coin toss lands heads; nothing otherwise;

---

[5]For example, let the state probabilities $p_i = P(X = i)$ be given by $(p_1, p_2, \ldots,) = (\frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{4}, \frac{1}{32}, \frac{1}{64}, \frac{1}{128}, \frac{1}{8}, \frac{1}{256}, \frac{1}{512}, \frac{1}{1024}, \frac{1}{16}, \frac{1}{2048}, \ldots)$. (The pattern is $p_{4n+4} = \frac{1}{2}p_{4n}$, $p_{4n+i} = \frac{1}{8}p_{4n+i-4}$, $i = 1, 2, 3$.) Next let $(a_1, a_2, \ldots,) = (2, 4, 8, 0, 16, 32, 64, 0, 128, 256, 512, 0, 1024, \ldots)$ and $(b_1, b_2, \ldots,) = (0, 0, 0, 2, 0, 0, 0, 4, 0, 0, 0, 8, 0, \ldots)$. a_X and b_X are identically distributed, but REU dominance judges a_X preferable to b_Y; in particular, $\sum_{i=1}^{4n} p_i(a_i - b_i) = n \to \infty$. What makes this example work is of course an ordering under which the states associated with positive outcomes for a_X sufficiently precede the corresponding states having positive outcomes for b_Y.

[6]What Colyvan and Hájek actually say is that $REU(A, B) = \sum_{i=1}^{\infty} P(X = i)(a_i - b_i)$ "where the right-hand side absolutely converges, or diverges to infinity or negative infinity." That generates sensitivity to order, as any conditionally convergent series has rearrangements tending to (positive or negative) infinity. We believe that our formulation corresponds to their intention.

Bet 2: Pays 6 if a fair die toss lands even; nothing otherwise.

They write:

> As things currently stand, RET is silent on this case because there are no states in common across the two bets: 'heads on a toss of a fair coin' and 'even number on a roll of a fair die' are, on the face of it, different states. Yet, we want to say that (Bet 2) is preferable to (Bet 1) by compelling dominance-like reasoning. The obvious move to make here is to stipulate that we can identify 'heads on a toss of a fair coin' and 'even on a roll of a fair die' because they have the same probability, and that's all that matters. We thus supplement RET with this probabilistic identification of states: the states under one action can be identified with the states under a different action in the same decision problem iff there is a one-to-one correspondence between the two sets of states that maps each state under one action to a state of equal probability under the second action.

This second emendation is somewhat imprecise: the one-to-one correspondence Colyvan and Hájek have in mind can't be between $\{heads, tails\}$ and $\{1, 2, 3, 4, 5, 6\}$ (these sets have different cardinalities), but rather between $\{heads, tails\}$ and $\{even, odd\}$, i.e. $\{heads, tails\}$ and $\{\{2, 4, 6\}, \{1, 3, 5\}\}$. Minimally, then, the intention here must be to allow that families of states on which the outcomes of a certain action are constant might be amalgamated prior to "probabilistic identification"–so if the state space of an action is "unnecessarily fine-grained", that won't be an impediment to probabilistic identification.[7]

To understand our quarrel with the resulting theory (call it "supplemented RET"), consider the following desideratum (proposed by Seidenfeld et. al. 2009):

Coherent Indifference: The difference of two indifferent variables should be indifferent from the zero variable.

We accept **Coherent Indifference**. For suppose that X and Y are indifferent variables. One ought to be indifferent between the status quo and an arrangement where one borrows X (interest free) and repays it immediately. But to borrow X, exchange it for Y and then repay the X simply yields a net gain of $Y - X$.

But let W be a St. Petersburg gamble, and independently toss a fair coin. Let

$X = W$ if *heads*; otherwise $X = 0$.

$Y = W$ if *tails*; otherwise $Y = 0$.

[7]Colyvan and Hájek probably want to allow more comparisons than those provided by the "minimal" suggestion in the text. Indeed, that suggestion won't even allow one to compare Bet 3, which pays 6 if the coin lands *heads* and nothing otherwise, with Bet 4, which pays 6 if the die lands on anything other than *one* (and nothing otherwise). That limitation would not sit well with Easwaran (2014), for example, who writes: "The central observation in the development of my theory is that if one ought to prefer act A to act B, and one ought to be indifferent between acts B and C, then one ought to prefer A to C." (Taking $A =$"accept Bet 4", $B =$"accept Bet 2" and $C =$"accept Bet 3", this principle calls for Bet 4 to be deemed preferable to Bet 3.)

Supplemented RET is indifferent between X and Y, but silent regarding the comparison of $Y - X$ with 0. (It regards $Y - X$ and 0 as *incommensurable.*) It therefore violates **Coherent Indifference**.

The first thought that a defender of supplemented RET will have is that its domain of comparison might be extended in some way. (Restoring **Coherent Indifference**, in particular.) Indeed, such moves were explicitly left open by Colyvan and Hájek, who wrote "It is worth seeing how far we can proceed with...state identification, perhaps supplementing it with something a little more sophisticated." Our claim, however, is that in allowing identification of states, Colyvan and Hájek have already proceeded too far. To see why, consider the following desideratum:

Transitivity of Indifference: If X is indifferent from Y and Y is indifferent from Z then X should be indifferent from Z.

We accept **Transitivity of Indifference** for reasons very similar to those grounding our acceptance of **Coherent Indifference**. For if one is indifferent to trading X for Y and indifferent to trading Y for Z, it seems to us that one is thereby indifferent to trading X for Z (the net result of conducting both trades).

Accepting both desiderata, our objection to supplemented RET is as follows:

Main Objection: There is no consistent extension of supplemented RET that satisfies both **Coherent Indifference** and **Transitivity of Indifference**.

Our proof of this goes by way of an example from Seidenfeld et. al. (2009). Suppose that **CH** is an extension of supplemented RET satisfying both **Coherent Indifference** and **Transitivity of Indifference**; we will show that **CH** must be inconsistent. To this end we let (following Seidenfeld et. al. 2009) X be a geometric random variable with $P(X = n) = 2^{-n}$, $n = 1, 2, \ldots$. Let a fair coin be tossed independently of X, and let $W = 2^X$ (a St. Petersburg variable). Next let

$W_1 = 2W$ if *heads* and $W_1 = 2$ if *tails*;

$W_2 = 2$ if *heads* and $W_2 = 2W$ if *tails*

Employing state identifications consistent with Colyvan and Hájek's practice, one may check that W, W_1 and W_2 are mutually indifferent under supplemented RET, and hence under **CH**. By **Coherent Indifference**, then, we get that both $W_1 - W$ and $W - W_2$ are indifferent from 0 under **CH** and so, by **Transitivity of Indifference**, from each other. Then (again by **Coherent Indifference**),

$$2 = (W_1 + W_2) - 2W = (W_1 - W) - (W - W_2)$$

is indifferent from zero–an inconsistency. qed

The objection may alternatively be formulated as vulnerability to a finitary Dutch Book. Suppose that a supplemented RET-obeying agent starts with a long position in W_1. She trades this for a long position in W; by state identification, she regards this trade as fair. Now she learns the valuation $W = k$, and accepts the following fair bet. If a previous independent toss of a fair coin landed *heads*, she wins $k - 1$ for a net of $2k - 1$; if the toss landed *tails*, she loses $k - 1$ for a net of 1. As it happens, this is the very coin toss used in the determination of W_1 from

W, i.e. $W_1 = 2k$ if *heads* and $W_1 = 2$ if *tails*. Therefore the agent's net is exactly $W_1 - 1$, meaning that she is 1 unit of currency worse off than when she started.

The culprit here is plainly the identification of states–which, it's worth noting, Colyvan and Hájek weren't fully comfortable with themselves.[8] They wrote: "RET needed...state identification to deliver various compelling dominance-like verdicts. But perhaps we've jumped from the frying pan into the fire here."

But RET does *not* require state identification to deliver the sought-for verdicts. Indeed, when comparing experiments defined on distinct spaces, some join of these spaces may be chosen to model each. In the above example involving Bet 1 and Bet 2 (where one may realistically assume that the tosses of coin and die are independent, even if one is merely counterfactual), the independent join is appropriate. This has four states: *heads even*, *heads odd*, *tails even* and *tails odd*. Bet 1 will have payout vector $(b_1, b_2, b_3, b_4) = (5, 5, 0, 0)$ and Bet 2 will have payout vector $(a_1, a_2, a_3, a_4) = (6, 0, 6, 0)$ with $P(X = i) = \frac{1}{4}$, $i = 1, 2, 3, 4$. An easy computation now gives $REU(A, B) = \frac{1}{2} > 0$, so A is preferable to B.

If that's right then Colyvan and Hájek's account of state identification isn't necessary for RET after all, and there is nothing to prevent one from offering as extensions of RET theories not subscribing to any such account. Indeed, the cost of Colyvan and Hájek's second emendation would appear to preclude state identification so far as "conservative" extensions of RET are concerned.

The implied directive is clear: keep Colyvan and Hájek's emendation to **REU Dominance** about order independence, but jettison the one about state identification. The resulting theory is a joint extension of expected utility theory and dominance reasoning accomplishing much in the way of Colyvan's initial concerns:

Order-independent Relative Expectation Theory (ORET): Let X be any random variable taking values in the natural numbers $\mathbf{N} = \{1, 2, \ldots\}$. Let Game A pay a_i and let Game B pay b_i when $X = i$. Define $REU(A, B) = \sum_{i \in \mathbf{N}} P(X = i)(a_i - b_i)$ provided this sum's value is independent of the order of the indices. (I.e. whenever $E(a_X - b_X)$ has a value in the extended reals.) Then Game A is preferable to Game B if and only if $REU(A, B) \in (0, +\infty]$. Game A and Game B are *incommensurable* if $REU(A, B)$ is undefined.

ORET recovers the "clear" comparisons (Altadena preferable to Pasadena, etc.) cited by Nover and Hájek. A further virtue is that it never indicates a preference for a gamble X over an identically distributed gamble Y.

Theorem 1. Suppose that a_X and b_X are identically distributed and $REU(A, B)$ is defined. Then $REU(A, B) = 0$.

[8]The decision theory described in the third section of Easwaran (2014) (entitled "One Version of the Theory", i.e. of the general type described in the first two sections) advocates for state identifications, and so is vulnerable to the current objections. Easwaran however hedges as follows: "I think that the relations presented here are normative for decision theory, but if some of them are not, then they can be switched out for others that might do some of the same work."

Proof.[9] We may assume without loss of generality that $REU(A, B) \geq 0$. Let A and B be the random variables a_X and b_X, respectively. Suppose for *reductio* that $E(A - B) > 0$. Fix k large enough that $E(\min\{k, A - B\}) > 0$. Let A_n and B_n be the truncations of A and B at $[-n, n]$, respectively. (I.e. $A_n = A$ if $|A| \leq n$, but $A_n = n$ when $A > n$ and $A_n = -n$ when $A < -n$, etc.) One easily checks that $\min\{A_n - B_n, k\} \to \min\{A - B, k\}$ almost surely. Moreover, $|\min\{A_n - B_n, k\}| \leq |\min\{A - B, k\}|$. So by the dominated convergence theorem,

$$E\big(\min\{A_n - B_n, k\}\big) \to E\big(\min\{A - B, k\}\big) > 0.$$

But A_n and B_n are identically distributed, so for all large enough n

$$0 = E\big(A_n - B_n\big) \geq E\big(\min\{A_n - B_n, k\}\big) > 0,$$

a contradiction. qed

5. Objections to **ORET**

As noted, agents subscribing to **ORET** do not ever prefer a gamble X to an identically distributed gamble Y. They also possess immunity from finitary (see footnote 4 and the appendix) Dutch Books. These are apt properties for a decision theory to have; still, there are objections that can be made against **ORET**.

First Objection: Group Dutch Books

Utility-pooling groups of agents subscribing to **ORET** are vulnerable to "Group Dutch Books" if the agents comprising the group can make unilateral decisions and believe that it's possible to confer a good of unbounded expected utility:

San Marino Game: Stanley and Stella are **ORET** subscribers married in the state of Louisiana, where they have what is known as the Napoleonic Code (according to which what belongs to the wife belongs to the husband also and vice versa). Stanley (together with a lawyer acquaintance) has devised a plan capitalizing on the fact that a gift of money on Stella's birthday is theoretically free under the Code. To liven things up, he presents to Stella a Huntington Library postcard with an enclosed coupon reading "Happy Birthday Stell. Luck is believing you're lucky! This coupon good for one Pasadena gamble, payable in dollars." Stella complains to her sister (Blanche) that although she has accepted the "gift" she's realized she could end up owing Stanley money under its terms. Blanche (a sometime adjunct scholar) sees an opportunity and offers to administrate the gamble. First, however, Blanche shows Stella a partition $\{P_i : i = 1, 2, \ldots\}$ of the naturals such that, for every i, the expectation of the Pasadena gamble in question exists and is equal to $-\infty$ conditional on the gamble paying from a state $n \in P_i$, and shows Stanley a partition $\{Q_j : i = 1, 2, \ldots\}$ of the naturals such that, for every j, the expectation of the Pasadena gamble in question exists and is equal to $+\infty$ conditional on the gamble paying from a state $n \in Q_j$. She explains to them both that the Pasadena gamble pays P_n with probability p_n, $n = 1, 2, \ldots$

[9] Our original proof was needlessly complicated; this simplification is due to Máté Wierdl.

and that they will learn which cell from their own partition contains n and be given a chance to cancel their position (for a price) after receiving this information but before learning the value of n. Blanche now puts Stanley and Stella to sleep in separate rooms. While they are asleep she rolls dice to determine n, then wakes them up. Blanche now goes to Stella and tells her the unique value i for which $n \in P_i$. At this point Stella realizes that the expected value of the gamble, from her perspective, is $-\infty$. Blanche now offers, as she promised, to sell her a short position in the same gamble–for a mere \$5. Stella gives Blanche the money, effectively cancelling her long position. Blanche then goes to Stanley and tells him the unique value j for which $n \in Q_j$. At this point Stanley realizes that the expected value of the gamble, from his perspective, is $+\infty$. Blanche now offers, as she promised, to sell him a long position in the same gamble–for \$150. He agrees to the transaction, which effectively cancels both his own short position and her long. The Stanley/Stella team has lost a sure \$155–a "Group Dutch Book".

Second Objection: Infinitary Dutch Books

Vann McGee (1999) used an "airtight Dutch Book" in an attempt to show that the combination of infinite state space and unbounded utility function leads expected utility theory subscribers to decision theoretic incoherence. McGee's Dutch Book is *infinitary*; in particular it consists in an infinite sequence of payoffs $(w_i)_{i=1}^{\infty}$ such that $\sum_i E\big(\min\{w_i, 0\}\big) = -\infty$. So while expected utility theory sanctions each bet considered by itself, simultaneous acceptance of them would appear to violate the spirit of **ORET**. Indeed, it's an easy matter to express any no-expectation wager as an infinite series of finite expectation wagers, so clearly **ORET** must be taken to implicitly sanction against simultaneous acceptance (or acceptance within any bounded window of time) of wagers constituting such a series.

Even with this clarification, however, individual **ORET**-subscribing agents who believe themselves to have lifespans of infinite expected duration are still subject to infinitary Dutch Books administered across time. To see this, consider the following twist on an experiment from Arntzenius, Elga and Hawthorne (2004):

Trumped. Donald Trump has just arrived in Purgatory. God explains that the duration X of his afterlife will be an instance of the St. Petersburg random variable (equal to 2^n with probability 2^{-n}), in days. Variety is the spice of the afterlife, however, and Trump will have the option, on Day 1, of spending that day in Hell in exchange for Days $4, 8, 12, \ldots, 1020$ in Heaven (each contingent on his being around). This is an expected two days and he finds Heaven to be as pleasant as Hell is unpleasant, so he takes the deal. On Day 2 he agrees to spend that day in Hell in exchange for Days $1024, 1028, 1032, \ldots, 2^{18}-4$ in Heaven (again an expected two days). And so forth...each day numbered $4n$ Trump spends in Heaven, but he spends the other days in Hell in exchange for contingent days $2^{8k+2}, 2^{8k+2}+4, \ldots, 2^{8k+10}-4$ in Heaven, $k = 2, 3, \ldots$. Expected utility theory recommends each bet, but the result of accepting them all is that Trump spends at least three-fourths of his afterlife in Hell.

Though Trump only accepts one wager per day, though their negative payoffs are bounded below and though there are almost surely only finitely many such wagers, the door is left open to the Dutch Bookie by the fact that their negative expected payoffs nevertheless sum to minus infinity, due to the fact that the expected number of wagers is infinite.[10]

Third Objection: Missed Arbitrages

Adam Elga (2010) has an arbitrage argument against imprecise credences that can be turned against the **ORET** subscriber. Suppose you are offered a dollar to take a long position in X, a no-expectation random variable. If you subscribe to **ORET** then X and the dollar are not commensurable, so if you believe that the offer is made in good faith then you'll presumably decline it. Moments later you are offered a dollar to take a short position in X. Again you decline. Nothing changes if we assume that you have prior knowledge of the protocol. That's apparently irrational, as accepting both offers strictly dominates rejecting them.

6. How Anthropic Reasoning Lowers Expectations

The problems raised in the previous section look, at first anyway, quite hopeless. One solution suggests itself: deny that infinite expectation (or no expectation) gambles are possible. Richard Jeffrey (1983) has taken this line. Most philosophers will be unsatisfied with this solution, however. McGee (1999), for example, writes that "a global plan cannot afford to ignore exotic possibilities". He opines: "If it were somehow assured to me that, for the price of one licorice jellybean I could guarantee that, if there is indeed an afterlife, my place in it would be one of boundless bliss, I would give up the bean." Having constructed his infinitary Dutch Book against agents having such attitudes, McGee concludes that there is simply no way to "avoid being defeated by our own ill-planned actions."

Hájek (2006) concurs. Citing "regularity" assumptions on which one should not assign probability zero to events of discrete type (e.g. *this St. Petersburg gamble is genuine*) that are not logically contradictory, Hájek writes:[11]

> Suppose I offer you the St. Petersburg game. You don't believe me; in fact you assign probability one in a trillion to the offer being genuine. Still, the paradox has a hold on you: for now the expectation of the game is a trillionth of infinity, which is still infinity.

For **ORET**, which would become a vacuous extension of classical expected utility theory should infinite expectation variables be banished from decision theory, the point is especially pressing. There is, however, a way to allay Hájek's concerns less radically–arrange that the expectation of the variable equal to Hájek's dubious St. Petersburg gamble, if it be genuine, and equal to zero otherwise be finite, yet non-zero.[12] This is apt to sound *ad hoc* and desperate. However, we shall argue

---

[10]This sort of disaster can't befall Trump if he knows his afterlife to have expected duration (in days) $E(X) = L < \infty$. This is a consequence of the optional stopping theorem (thanks to an anonymous referee for this point). It is also a consequence of Theorem 2 in the appendix, which establishes that finitary Dutch Books are precluded under **ORET** more generally.

[11]Cf. the "contagion" issue raised in Hájek and Smithson (2012).

[12]By, for example, setting the probability of veracity equal to some infinitesimal hyperreal ϵ and setting the expectation, conditional on veracity, equal to a finite non-zero multiple of ϵ^{-1}.

that it's a rationally mandated consequence of the anthropic principle. Brandon Carter (1983; see also A. Lewis 2001), formulates this principle as follows:

> "In a typical application of the anthropic (self-selection) principle, one is engaged in a scientific discrimination process of the usual kind in which one wishes to compare the plausibility of a set of alternative hypotheses, $H(T_i)$, say, to the effect that respectively one or other of a corresponding set of theories $T_1, T_2, \ldots$ is valid for some particular application in the light of some observational or experimental evidence E, say. Such a situation can be analysed in a traditional Bayesian framework by attributing *a priori* and *a posteriori* plausibility values (i.e. formal probability measures), denoted by p_E and p_S, say, to each hypothesis respectively before and after the evidence E is taken into account, so that for any particular result X one has
>
> $$p_E(X) = p_S(X|E),$$
>
> the standard symbol $|$ indicating conditionality. According to the usual Bayesian formula, the relative plausibility of two theories A and B, say, is modified by a factor equal to the ratio of the corresponding conditional *a priori* probabilities $p_S(E|A)$ and $p_S(E|B)$ for the occurrence of the result E in the theories, i.e.
>
> (1) $$\frac{p_E(A)}{p_E(B)} = \frac{p_S(E|A)}{p_S(E|B)} \frac{p_S(A)}{p_S(B)}."$$

The "Selected" or "Subjective" probability function p_S in (1) is related to an "Original" or "Objective" probability function p_O by $p_S(\cdot) = p_O(\cdot|S)$: "$S$ denotes...the selection conditions that are implied by the hypothesis of application of the theory to a concrete experimental or observational situation, but which are not necessarily included in the abstract theory" on which p_O is based.

It's implicit from Carter's own usage of the principle that a "theory" meanwhile is a family of measures on the set of universe histories rather than a chance or ineliminably indexical event such as "this toss of this coin lands *heads*". Examples of "theories" from Carter (1983) include the hypotheses: life is very rare, even in favorable conditions; gravitational coupling strength is fixed across time; and, the expected average time $\bar{t}$ intrisically most likely for the evolution of a system of observers intelligent enough to comprise a scientific civilization such as our own is geometrically small relative to the main sequence lifetime τ of a typical star.

Indeed, where A is a theory your nomologically accessible evidential counterparts (possible beings with thoughts indiscriminable from your own) should intend by "A" the exact same proposition you intend by "A"; in particular, the two utterances should be associated with the same truth value. (Your counterparts aren't actual beings contemplating a different coin or counterfactual beings for whom the coin landed otherwise than it actually did.) The importance of this restriction cannot be overstated, for Carter wishes to employ the identity

(2) $$\frac{p_S(A)}{p_S(B)} = \frac{p_O(A)}{p_O(B)},$$

which will not in general be valid for "non-theory" events A and B.[13]

Consider now the following two premises.

First premise. The theory that there are *no* infinite expectation gambles should be assigned positive probability.

Second premise. There is an upper bound on the rate at which utility can accrue.

Accepting these, the expectation one ought to assign to an encountered St. Petersburg gamble after applying Carter's anthropic principle isn't infinite, but finite. For suppose that Trump, on his first (and only) day in Purgatory, is offered a St. Petersburg variable X of days in Heaven[14] in exchange for Y days in Hell, where Y is either 10 or 300 based on the toss of a fair coin. (Trump is told the value of Y prior to making his choice.) We again suppose that Trump values x days in Heaven at $+x$, x days in Hell at $-x$, and x days in Purgatory at 0. If Trump refuses the offer, his afterlife terminates at the end of the day. If he accepts, it terminates upon settlement. Trump is typical; every other conscious being in the universe faces an analogous afterlife gamble (we assume their pre-afterlife lifetimes are negligible) and that they are all rational and know the relevant protocols.

Suppose further that Trump entertains exactly two competing theories about the universe. Theory A says that the St. Petersburg offers X that one encounters in the afterlife are genuine. Theory B says that they are not genuine; in fact, their true expectations are precisely 50 days. Prior to invoking anthropic reasoning, Trump assigns Theory A positive probability $\epsilon < \frac{1}{2}$. According to the sort of reasoning Hájek implicitly cites, then, the expected number of days Trump will spend in Heaven, should he accept, is $(\epsilon)\infty + (1-\epsilon)50 = \infty$. He should therefore accept, regardless of the value Y.

[13]Cf. Sleeping Beauty, where the majority intuition is that (2) fails for $A =$ *heads*, $B =$ *tails*, vs. Bostrom's "Presumptuous Philosopher", where the majority intuition is that (2) holds for $A =$ *trillion trillion persons* and $B =$ *trillion trillion trillion persons*. See Bostrom (2007).

[14]A referee suggests a line of resistance to this argument: assume instead (against the second premise) that Trump is offered one day in "St. Petersburg Heaven", i.e. a Heaven offering a single-day experience which, with probability 2^{-n}, will be as gratifying as 2^n days in Hell is ungratifying. This method of realizing an infinite expectation payoff requires one to countenance the notion of satisfaction (or dissatisfaction, presumably) "singularities". (*In an instant's compass, great hearts sometimes condense to one deep pang, the sum total of those shallow pains kindly diffused through feebler men's whole lives.*) Note however that for the objection to work, unbounded quantities of satisfaction would have to be–to put the matter somewhat crudely– "crammed into a bounded quantity of consciousness". Put to the choice, this is where we would draw the line. Wherever satisfaction is singular with respect to spacetime, the currency of anthropic reasoning ("consciousness", if you like) is singular with respect to spacetime as well. (The referee responds "...the argument...seems to assume that utility is hedonic rather than just a representation on the amount one cares about outcomes in the world". But if the idea is that one might "care about a world one will not see", this just reverts to the situation treated in the main text. It doesn't matter there that Trump experiences the days in Heaven himself; if they are experienced by others and Trump is, e.g., a utilitarian, nothing relevant changes.)

If we take Carter's anthropic principle into account, however, this computation breaks down. Assume for the moment that every being accepts the offer of X days in Heaven for Y days in Hell when $Y = 10$, but refuses when $Y = 300$. Let P be the event "I am now in Purgatory". Conditional on Theory B, half of all agents encounter $Y = 10$ and accept the ensuing offer; these agents spend an average of 61 days in the afterlife (1 in Purgatory, 10 in Hell, and an expected 50 in Heaven). The other half encounter $Y = 300$ and refuse the ensuing offer; these agents spend 1 day in the afterlife (in Purgatory). It follows that the expectation of Trump's afterlife conditional on B is 31 days, so that $p_S(P|B)$ is the multiplicative inverse of this expectation, i.e. 1/31.

Next assume, for the time being, that $E(X|A)$ is large but finite. Then

$$L = \frac{1}{2}(1) + \frac{1}{2}\big(11 + E(X|A)\big) = 6 + \frac{1}{2}E(X|A)$$

is the expectation of Trump's afterlife conditional on A. $p_S(P|A)$, meanwhile, is equal to L^{-1}. One therefore has

$$\frac{p_P(A)}{p_P(B)} = \frac{p_S(P|A)}{p_S(P|B)}\frac{p_S(A)}{p_S(B)} = \frac{L^{-1}}{1/31}\frac{\epsilon}{1-\epsilon},$$

from which it follows that $p_P(A) = \frac{31\epsilon}{31\epsilon + L(1-\epsilon)}$ and $p_P(B) = \frac{L(1-\epsilon)}{31\epsilon+L(1-\epsilon)}$. Trump now computes the posterior expectation of X as follows:

$$E(X) = p_P(A)E(X|A) + p_P(B)E(X|B) \approx \frac{31\epsilon(2L-12)}{31\epsilon + L(1-\epsilon)} + \frac{L(1-\epsilon)(50)}{31\epsilon + L(1-\epsilon)}.$$

Letting $L \to \infty$, we get $E(X) = \frac{62\epsilon}{1-\epsilon} + 50 \leq 112$ when $E(X|A) = \infty$.[15]

Suppose next that the agents do accept when $Y = 300$. Then the expectation of afterlife duration conditional on B would be $\frac{1}{2}(61 + 351) = 206$ and the expectation of afterlife duration conditional on A would be $L = 156 + E(X|A)$, so that $p_S(P|B) = \frac{1}{206}$ and $p_S(P|A) = L^{-1}$. So

$$\frac{p_P(A)}{p_P(B)} = \frac{p_S(P|A)}{p_S(P|B)}\frac{p_S(A)}{p_S(B)} = \frac{L^{-1}}{1/206}\frac{\epsilon}{1-\epsilon},$$

from which it follows that $p_P(A) = \frac{206\epsilon}{206\epsilon+L(1-\epsilon)}$ and $p_P(B) = \frac{L(1-\epsilon)}{206\epsilon+L(1-\epsilon)}$. So:

$$E(X) = p_P(A)E(X|A) + p_P(B)E(X|B) \approx \frac{206\epsilon(L-156)}{206\epsilon + L(1-\epsilon)} + \frac{L(1-\epsilon)(50)}{206\epsilon + L(1-\epsilon)}.$$

Letting $L \to \infty$, $E(X) = \frac{206\epsilon}{1-\epsilon} + 50 \leq 256$ when $E(X|A) = \infty$. So $E(X) < 300$ whether the agents accept when $Y = 300$ or not. (Whence they don't.)

[15] This seems a fair way to compute $E(X)$; as alluded to in footnote 12, one could also come to this conclusion via nonstandard analysis by letting L be an appropriate infinite hyperreal (so that L^{-1} is infinitesimal).

If that's right, then the claim that assigning even a tiny positive probility to a claimed St. Petersburg variable's veracity requires you to assign the variable an infinite unconditional expectation can be resisted. In fact premise 2, together with Carter's anthropic principle, establishes something like the opposite. Namely, that assigning positive probability to the theory that faithful St. Petersburg gambles are impossible requires you to assign gambles advertised as St. Petersburg gambles finite unconditional expectation whenever they are encountered.

Note however that the above computation cannot be adapted to sharply evaluate gambles that have no expectation conditional on their being genuine; one might obtain finite upper and lower bounds on the value of such gambles by this method (and could bring these to within the distance of McGee's bean by setting one's prior credence in the impossibility of infinite expectation gambles near 1)[16], but precise evaluation would continue to render agents vulnerable to a finitary Dutch Book. Such observations support Nover and Hájek's (2004) contention that "the Pasadena game is more paradoxical than the St. Petersburg game in several respects." Indeed, they suggest that such no-expectation gambles must be given their due, at least by anyone superstitious enough to entertain the theory that they are possible. Ideal agents, to whom the complete theory of everything is transparent, will apparently (with high probability) be unwilling to entertain such a theory. Those whose epistemic limitations oblige them to assign it probability some $\epsilon > 0$, meanwhile, must simply accept the implied vulnerability.[17]

7. Appendix

In Section 5 we saw that ORET subscribers are vulnerable to group Dutch Books and to infinitary Dutch Books. The content of the following theorem is that they needn't worry about *finitary* Dutch Books.

Theorem 2. Individual ORET subscribers are not vulnerable to finitary Dutch Books.

Proof. Let (Ω, μ) be a probability space and let $(\mathcal{F}_n)_{n=0}^{\infty}$ be a filtration on Ω. That is, for each n, $\mathcal{F}_n$ is a σ-algebra of measurable subsets of Ω, with $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots$. For $n = 0, 1, 2, \ldots$, let X_n be a real valued random variable defined on Ω. We assume that $X_0 = 0$ a.e. and $E(X_{n+1} - X_n | \mathcal{F}_n) \geq 0$ a.e., $n = 0, 1, 2, \ldots$.

The idea here is that X_n represents the agent's bankroll at timestep n. Also at time step n (or shortly after), the agent is assumed to learn which cell of $\mathcal{F}_n$ obtains. The agent then consents (between time n and time $n+1$) to a gamble having payoff $X_{n+1} - X_n$. Since the agent knows which cell of $\mathcal{F}_n$ obtains and $E(X_{n+1} - X_n | \mathcal{F}_n) \geq 0$ a.e., the agent is at least indifferent to these gambles (she may view them as favorable), so we can assume that she accepts them.

The challenge for the would-be finitary Dutch Bookie is to construct a random variable T such that the gambles stop at time T (that is, the agent's final bankroll

[16]Replacing the genuine St. Peterburg variables by genuine Pasadena variables in the first calculation of the Trump example yields bounds $E(X) \in [50 - \frac{62\epsilon}{1-\epsilon}, 50 + \frac{62\epsilon}{1-\epsilon}]$.

[17]Thanks to Máté Wierdl, the editors of *Synthese*, and the anonymous referees.

is X_T) and $X_T < 0$ a.e. We require that for every n, $\{\omega \in \Omega : T(\omega) \leq n\} \in \mathcal{F}_n$. ($T$ is a *stopping time*; the bookie is allowed to know no more than the agent–when the gambles have stopped, the agent will know this.) For the Dutch Book to be finitary, one of following two additional conditions must be met (cf. footnote 4).

First Condition: $T(\omega) \leq K$ a.e. for some $K < \infty$.

Note that $E(X_1|\mathcal{F}_0) = E(X_1 - X_0|\mathcal{F}_0) \geq 0$ a.e., which implies in turn that $P\big(E(X_1|\mathcal{F}_1) \geq 0\big) > 0$.

Next note that $E(X_2|\mathcal{F}_1) \geq E(X_1|\mathcal{F}_1)$ a.e. (both may $= +\infty$), so that

$$P\big(E(X_2|\mathcal{F}_1) \geq E(X_1|\mathcal{F}_1) \geq 0\big) > 0.$$

This, in turn, implies that

$$P\big(E(X_2|\mathcal{F}_2) \geq E(X_1|\mathcal{F}_1) \geq 0\big) > 0.^{18}$$

Having shown that

$$P\big(E(X_n|\mathcal{F}_n) \geq E(X_{n-1}|\mathcal{F}_{n-1}) \geq \cdots \geq E(X_1|\mathcal{F}_1) \geq 0\big) > 0,$$

note that $E(X_{n+1}|\mathcal{F}_n) \geq E(X_n|\mathcal{F}_n)$ a.e. This implies that

$$P\big(E(X_{n+1}|\mathcal{F}_n) \geq E(X_n|\mathcal{F}_n) \geq E(X_{n-1}|\mathcal{F}_{n-1}) \cdots \geq E(X_1|\mathcal{F}_1) \geq 0\big) > 0,$$

which implies in turn that

$$P\big(E(X_{n+1}|\mathcal{F}_{n+1}) \geq E(X_n|\mathcal{F}_n) \geq E(X_{n-1}|\mathcal{F}_{n-1}) \cdots \geq E(X_1|\mathcal{F}_1) \geq 0\big) > 0.$$

By induction, then, one has

$$P\big(E(X_K|\mathcal{F}_K) \geq E(X_{K-1}|\mathcal{F}_{K-1}) \geq E(X_{K-2}|\mathcal{F}_{K-2}) \cdots \geq E(X_1|\mathcal{F}_1) \geq 0\big) > 0.$$

There is, therefore, a positive measure set $F \in \mathcal{F}_K$ such that

$$E(X_K|\mathcal{F}_K) \geq E(X_{K-1}|\mathcal{F}_{K-1}) \geq E(X_{K-2}|\mathcal{F}_{K-2}) \cdots \geq E(X_1|\mathcal{F}_1) \geq 0$$

a.e. on F. We may, moreover, assume that $T(\omega)$ is constant on F, say $T(\omega) = n \leq K$, $\omega \in F$. We therefore have

$$F \subseteq \big\{\omega : E(X_n|\mathcal{F}_n)(\omega) \geq 0\big\} \cap T^{-1}(n) = F_n \in \mathcal{F}_n.$$

Now if $X_{T(\omega)}(\omega) < 0$ a.e. then $X_n(\omega) < 0$ for a.e. $\omega \in F_n$, which implies that $E(X_n|\mathcal{F}_n)(\omega) < 0$ for a.e. $\omega \in F_n$, contradicting the definition of F_n. So $P\big(X_{T(\omega)}(\omega) \geq 0\big) > 0$ and there is no Dutch Book.

Second Condition: $\sum_{n=0}^{\infty} E[(X_{n+1} - X_n)_- \cdot 1_{\{T>n\}}] < \infty$.

Let $(X_n^T)_{n=0}^{\infty}$ denote the stopped process, defined by $X_n^T = X_n$ if $n \leq T$ and $X_n^T = X_T^T$ when $n > T$. Notice that $E(X_{n+1}^T - X_n^T|\mathcal{F}_n) \geq 0$ a.e., $n = 0, 1, 2, \ldots$ and $X_n^T(\omega) \to X_{T(\omega)}(\omega)$ a.e. Let now

$$M = \sum_{n=0}^{\infty} (X_{n+1} - X_n)_- \cdot 1_{\{T>n\}}.$$

---

[18] There is a positive measure set $F \in \mathcal{F}_1$ such that $E(X_2|\mathcal{F}_1) \geq E(X_1|\mathcal{F}_1)$ a.e. on F. To assume that $E(X_2|\mathcal{F}_2) < E(X_2|\mathcal{F}_1)$ a.e. on F leads to an immediate contradiction, as integrating both sides of this equation over F yields $E(X_2 \cdot 1_F) < E(X_2 \cdot 1_F)$. So, there is a positive measure set $F' \in \mathcal{F}_2$ (with $F' \subseteq F$), on which $E(X_2|\mathcal{F}_2) \geq E(X_2|\mathcal{F}_1) \geq E(X_1|\mathcal{F}_1) > 0$.

Then $E(M) = \sum_{n=0}^{\infty} E[(X_{n+1} - X_n)_- \cdot 1_{\{T>n\}}] < \infty$ by the monotone convergence theorem. Moreover one has

$$-M \leq X_n^T \leq X_{T(\omega)} + M \tag{1}$$

a.e. Since $E(X_T) \geq E(-2M) > -\infty$, either $E(X_T) = +\infty$, or $E(X_T)$ is finite. In the latter case, (X_n^T) is uniformly integrable by (1) and so

$$E[X_T] = \lim_{n\to\infty} E[X_n^T] \geq E(X_0^T) = 0$$

by the dominated convergence theorem, since $E(X_{n+1}^T) \geq E(X_n^T)$ in the integrable case. So in either case, $E[X_T] \geq 0$ and there is no Dutch Book. qed

References

Arntzenius, Frank, Adam Elga and John Hawthorne. 2004. Bayesianism, Infinite Decisions and Binding. *Mind* 113:251-283.

Bartha, Paul. 2016. Making Do Without Expectations. *Mind* 125:799-827.

Bostrom, Nick. 2007. Sleeping Beauty and Self Location: A Hybrid Model. *Synthese* 157:59-78.

Broome, J. 1995. The 2-Envelope Paradox. *Analysis* 55:6-11.

Carter, Brandon. 1983. The Anthropic Principle and its Implications for Biological Evolution. *Philosophical Transactions of the Royal Society of London. Series A, Mathmatical and Physical Sciences.* 310:347-363.

Chalmers, David. 2002. The St. Petersburg Two Envelope Paradox. *Analysis* 62:155-57.

Colyvan, Mark. 2006. No Expectations. *Mind* 115:695-702.

Colyvan, Mark. 2008. Relative Expectation Theory. *Journal of Philosophy* 105(1):37-44.

Colyvan, Mark and Alan Hájek. 2016. Making Ado Without Expectations. *Mind* 125:829-857.

Easwaran, Kenny. 2008. Strong and Weak Expectations. *Mind* 117:633-641.

Easwaran, Kenny. 2014. Decision Theory Without Representation Theorems. *Philosopher's Imprint* 14(27):1-30.

Hájek, Alan. 2006. In Memory of Richard Jeffrey: Some Reminiscences and Some Reflections on *The Logic of Decision. Philosophy of Science* 73:947-958.

Hájek, Alan. 2014. Unexpected Expectations. *Mind* 123:533-567.

Hájek, Alan and Harris Nover. 2006. Perplexing Expectations. *Mind* 115:703-720.

Hájek, Alan and Michael Smithson. 2012. Rationality and Indeterminate Probabilities. *Synthese* 187:33-48.

Jeffrey, Richard C. 1983. *The Logic of Decision.* Chicago: University of Chicago Press; 2nd ed. 1983.

Lewis, Antony. 2001. Comparing Cosmological Theories. Online. Available at http://cosmologist.info/anthropic.html.

McGee, Vann. 1999. An Airtight Dutch Book. *Analysis.* 59:257-265.

Nover, Harris and Alan Hájek. 2004. Vexing Expectations. *Mind* 113:237-249.

Seidenfeld, Teddy, Mark J. Schervish and Joseph B. Kadane. 2009. Preference for Equivalent Random Variables: A Price for Unbounded Utilities. *Journal of Mathematical Economics* 45:329-340.

Smith, N. 2014. Is Evaluative Compositionality a Requirement of Rationality? *Mind* 123:457-502.

Randall McCutcheon, Dept. of Mathematical Sciences, University of Memphis
E-mail address: rmcctchn@memphis.edu