Infinite Prospects

Jeffrey Sanford Russell
University of Southern California

Yoaav Isaacs
Baylor University

Draft of April, 2020

Abstract

People with the kind of preferences that give rise to the St. Petersburg paradox are problematic—but not because there is anything wrong with infinite utilities. Rather, such people cannot assign the St. Petersburg gamble any value that any kind of outcome could possibly have. Their preferences also violate an infinitary generalization of Savage's Sure Thing Principle, which we call the *Countable Sure Thing Principle*, as well as an infinitary generalization of von Neumann and Morgenstern's Independence axiom, which we call *Countable Independence*. In violating these principles, they display foibles like those of people who deviate from standard expected utility theory in more mundane cases: they choose dominated strategies, pay to avoid information, and reject expert advice. We precisely characterize the preference relations that satisfy Countable Independence in several equivalent ways: a structural constraint on preferences, a representation theorem, and the principle we began with, that every prospect has a value that some outcome could have.

Here's the plan. (Section 1) The St. Petersburg gamble cannot be assigned any *fair value.* (2) Standard arguments for orthodox expected utility theory, involving money pumps or regret, are also arguments against the rationality of preferences that give rise to the St. Petersburg gamble. (3) These arguments are doing something very different from familiar arguments for bounded utilities. (4) The structural requirement the arguments support is compatible with *infinite* utilities. (5) We describe a representation theorem, and explain how money pumps and regret are connected to the first point about fair values. (6) Some misgivings are expressed.

Thanks to Andrew Bacon, Adam Elga, Brian Hedden, John Hawthorne, and an anonymous referee for helpful comments.

1 Don't Cost a Thing

Meet Lin. Lin values money *linearly:* two dollars is worth twice as much as one to her, and so on, however large the amount may be. (Lin is quirky that way.) Lin is also *risk-neutral*: she values an uncertain prospect according to its expectation—the usual probability-weighted average—at least when only finitely many different possible outcomes are in play. Lin also *respects dominance*: whenever one prospect is sure to be at least as good as another, and it might be better, that's the one she prefers. (This holds even for prospects that involve infinitely many different possible outcomes.)

Surfing eBay, Lin notices an auction posted by a seller in St. Petersburg for a ticket to play a famous game. A fair coin will be flipped until it comes up heads. If it's heads on the first flip, the buyer gets \$2. If the first heads is on the second flip, the buyer gets \$4, if on the third, \$8, and so on. How much should Lin bid?

She reasons as follows. For any number n, there is a length-n *truncation* of the St. Petersburg gamble. If heads doesn't come up in the first n flips, the n-truncated gamble pays nothing. Working out the expectation of this finite gamble, Lin sees that its dollar value is

$$\frac{1}{2} \cdot 2 + \frac{1}{4} \cdot 4 + \cdots + \frac{1}{2^n} \cdot 2^n = n$$

The St. Petersburg gamble *dominates* this truncated gamble, so it's worth strictly more than n dollars. So the St. Petersburg gamble is worth more to Lin than any finite amount of money. (So far, Lin's reasoning has followed Hájek and Nover 2006, 4.)

Here's a new question. Is the St. Petersburg gamble perhaps worth something like an *infinite* amount of money? No. The St. Petersburg gamble is sure to pay only a finite amount of money. Suppose there is something which is worth *more* than each finite amount of money—such as an infinite amount of money (whatever that might come to), or a priceless artwork, or true love. If Lin has something like that, then the prospect of keeping it (with certainty) will *dominate* giving it up in exchange for the St. Petersburg gamble; thus the St. Petersburg gamble is not worth so much.

Of course, nothing is worth more than *each* finite amount of money, yet not worth more than *every* finite amount of money. So the conclusion of Lin's reasoning is that nothing she could bid—monetary or otherwise—would be the right price. (That is, no *sure* thing she could bid. If she already had *another* ticket for a St. Petersburg gamble, that might do.) If she pays anything, she will either underpay or overpay. The St. Petersburg ticket has no *fair value.* To put it another way, however much the St. Petersburg gamble is worth, no particular *outcome* could be worth exactly that much.

This isn't because the space of outcomes has "gaps" that could be filled. The bet "one dollar if a fair die comes up six, otherwise nothing" is worth a sixth of a dollar to Lin—but maybe there doesn't happen to be any sure thing worth exactly one sixth of a dollar. (Nobody mints the right fractional coins.) If so, this is a mere accidental omission that could be filled, in principle, by filling in outcome space with new "extended outcomes" (like an "ideal" one-sixth-dollar-bill). The St. Petersburg ticket isn't like that. Even if we added new "extended outcomes," the St. Petersburg ticket still wouldn't be worth the same as any of them. Lin's preferences violate what we will call the *Extended Outcome Principle.* Here is the basic idea. (See Appendix A for details.) We start with a certain set of *outcomes*, and a preference relation over uncertain *prospects* constructed from these outcomes. The Extended Outcome Principle says that the set of outcomes can be mapped into a larger set of "extended outcomes," and likewise the preference ordering can be extended to a preference ordering over prospects constructed from these extended outcomes, in a way which is still risk-neutral and respects dominance, such that every extended prospect is equivalent in value to some extended outcome. Lin's preferences violate the Extended Outcome Principle: every *extended* outcome is still either no better than some finite amount of money, and thus worse than the St. Petersburg gamble, or else better than every finite amount of money, and thus better than the St. Petersburg gamble.[1]

An upshot of this is that Lin's preferences cannot be *represented* in a certain way. Just as she can't put a dollar value on every prospect, she also can't put a *utility* value on every prospect. This is an obstacle to a familiar general strategy for doing decision theory, which goes like this. Step one: assign each possible outcome a *utility.* Step two: do some mathematical operation or other on probabilities and utilities, to come up with a utility for each prospect—call it the *general expected utility.*[2] Step three: rank

[1] We assume the Completeness axiom throughout this essay: for any two options A and B, $A \lesssim B$ or $B \lesssim A$. We take strict preference and indifference to be defined in terms of weak preference in the usual way.

[2] To be clear, the *general expected utility* operation we are here imagining is supposed to be defined for *every* prospect, unlike the standard mathematical expectation, which in some infinite cases is ill-defined.

prospects according to their general expected utilities. This strategy won't work for Lin. If it did work, then we could think of utilities themselves as "extended outcomes," and use general expected utilities to construct an extended preference ordering in accordance with the Extended Outcome Principle. But this is impossible. Whatever utility we might assign the St. Petersburg gamble—whether we think of "utilities" as corresponding to real numbers, or extended reals, or hyperreals, or surreals, or abstract vectors, or whatever—dominance reasoning tells us that a prospect that *surely* had this utility either would be strictly worse than the St. Petersburg gamble or else strictly better than it. So the St. Petersburg gamble does not have any utility, even in this abstract sense.

This observation is bad news for any future decision theory that extends the standard theory with exotic new utilities (for example, Chen and Rubio 2018). No such theory can represent Lin's preferences. Still, we might think, so much the worse for that kind of decision theory. (It's not as if there are no alternative brands to try: see for example Bartha 2007; Colyvan 2008; Easwaran 2014; Lauwers and Vallentyne 2016, 2017) The St. Petersburg gamble has no fair value, but Lin's preferences between the St. Petersburg gamble and other material (and immaterial) goods seem to be in perfectly good order, for all we've seen so far. While violating the Extended Outcome Principle is bad news for certain decision theorists, it's not obviously bad news for Lin.

2 A Losing Strategy

Here comes bad news for Lin. Despite underbidding, Lin still lucked out and won the St. Petersburg ticket on eBay. But then she notices an intriguing post on Craigslist.[3]

> Want to trade one St. Petersburg ticket for another St. Petersburg ticket. (Mine has bad emotional associations.) Will throw in $100. Must pick up in Long Beach.

Sure, Long Beach is annoyingly far away, but the round trip Lyft fare will only be $50. This seems like a pretty good deal!

[3] The following story is a variant of puzzles presented by Arntzenius and McCarthy (1997, 49–50) and Chalmers (2002). To construct this variant we also drew on ideas from Buchak (2013 ch. 6); Briggs (2015); Buchak (2015); see also Machina (1989).

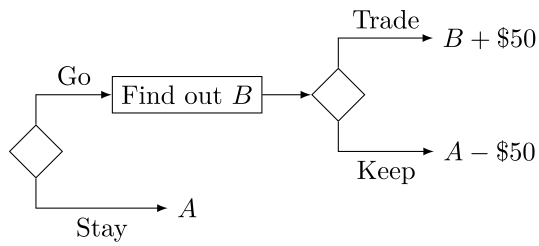


Figure 1: Lin goes for the dominated strategy (Go, Keep)

To be clear, Lin and the Craigslist poster have tickets for two *different* isomorphic St. Petersburg gambles, played using two different coins. Let's call Lin's ticket the A ticket and the Craigslist one the B ticket.

Lin shows up at the address, and is ready to make the swap.[4] But as she stands on the porch checking her phone, she sees that the results of the B coin flips have just been reported. In fact, the first heads was on the fourth flip, and so it turns out that the B ticket is worth just \$16. This whole trip doesn't seem worth it anymore: after all, as she already worked out, the A ticket she has in her pocket is worth more than any finite amount of money to her, and thus more than \$116. Without even knocking, she gets back in the car. She still has her A ticket, but she's out \$50 for the Lyft ride, with no upside.

Let's look at the structure of the decision situation Lin faced (Figure 1). When Lin first opted to go to Long Beach, this was because there was a strategy available to her—(Go, Trade)—which was associated with a prospect (B + \$50) that she preferred over the prospect associated with Stay, which was in turn better than (Go, Keep). But after gaining new information, her preference between these strategies was guaranteed to be upended. Whatever the result of the B gamble might have been, once Lin learned it she would prefer not to trade. Then (Go, Keep) would look better to her than (Go, Trade). But (Go, Keep) was sure all along to turn out *worse* than another strategy she could have chosen in the first place—Stay. In short, by always playing according to her preferred strategy, Lin has ended up playing by a strategy which is *dominated*—and indeed, dominated by an alternative strategy which is not itself dominated. That seems pretty bad.[5]

[4] The following reasoning assumes $B + \$50 > A$. This assumption can be dropped: otherwise, by Completeness, we have $B < B + \$50 \lesssim A < A + \$50$, in which case we could run exactly the same story, just switching A and B.

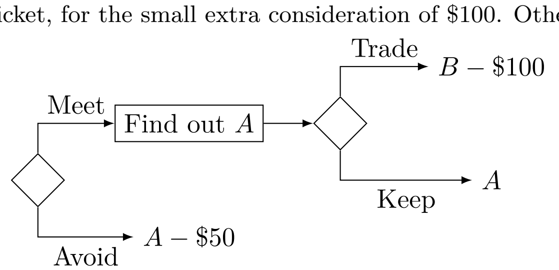
[5] You might think this is just a problem with the policy of always playing according

A number of philosophers have made this observation before: the St. Petersburg gamble—as well as other closely related cases involving similar underlying preferences, like the two-envelope paradox—can give rise to sequences of choices that lead to sure losses.[6] One common diagnosis of this kind of problem ascribes it to a mismatch between conditional and unconditional expected utilities.[7] This is basically on the right track, but not exactly perspicuous in this case—because, as we have discussed, the St. Petersburg gamble has no expected utility, even in a very general sense. We can sharpen our diagnosis. Lin's basic problem is that her *conditional preferences* given new information fail to mesh with her unconditional preferences. So Lin falls afoul of a natural principle of decision theory inspired by Savage (1954).

First a little set-up. As is standard, we can think of Lin's *options* as represented by functions from *states* to *outcomes*. Lin has some preference ordering over options. An *event* is a set of states; an event E is *non-null*

to your preferred strategy, in cases like these (see Buchak 2013, 175ff.). What if at the outset Lin took her future preference switch explicitly into account? After all, knowing herself pretty well, she can predict her disappointment upon learning the outcome of B. It's true that in this particular decision set-up, this *sophisticated choice* strategy will keep her from taking a dominated strategy. But there is a very close variant of this set-up that gets the sophisticated chooser in trouble instead (compare Briggs 2015).

This time, the Craigslist poster is coming to meet Lin, unless she pays \$50 to avoid the meeting. If she takes the meeting, then the poster will first tell her the outcome of the A gamble was. After that, Lin will have the opportunity to trade her A ticket for the B ticket, for the small extra consideration of \$100. Otherwise she can keep the A ticket.



In this case, the sophisticated chooser reasons that, upon finding out A, she will choose $B - \$100$ rather than A (however A turns out). She currently regards this as a worse prospect than $A - \$50$, so she will pay to avoid the meeting. But this strategy is dominated by the alternative strategy (Meet, Keep), which still gets A and saves \$50. Moreover, this alternative strategy is undominated.

Another approach is *resolute choice*: Lin might settle on a strategy at the outset, and stick to it even if it starts to look bad. This approach strikes us as giving up. It amounts to saying that Lin should not be guided by decision theory at all, except when it comes to one very special decision.

[6] See Footnote 3. See also Arntzenius, Elga, and Hawthorne (2004).

[7] See Lee (2013, 6–7); see also Broome (1995); Arntzenius and McCarthy (1997), p. 44; Clark and Shackel (2000); Chalmers (2002).

iff Lin assigns E positive probability. Then for any non-null event E, Lin also has a preference ordering over options *conditional on* E, which we write $A <_E B$ (and likewise $A \sim_E B$ and $A \lesssim_E B$). A *regular partition* is a set of non-null events which are pairwise disjoint and jointly exhaustive.

Lin's preferences violate the following principle.

The Countable Sure Thing Principle. Let A and B be options, and let $\mathcal{E}$ be a regular partition. If $A \lesssim_E B$ for every $E \in \mathcal{E}$, then $A \lesssim B$. If furthermore $A <_E B$ for some $E \in \mathcal{E}$, then $A < B$.

The Countable Sure Thing Principle captures the idea that unconditional preferences should "reflect" conditional preferences: if the conditional preferences all go the same way, no matter how things turn out, then the unconditional preferences should follow them. We call this the *Countable* Sure Thing Principle because all regular partitions are countable. We call this the Countable *Sure Thing* Principle because of its close relationship to another more familiar principle. If we simply insert the word "finite" before "partition," we get something which is equivalent to Savage's famous Sure Thing Principle (Savage 1954, 21–22).[8] So the Countable Sure Thing Principle is basically an infinite generalization of Savage's principle.

The Countable Sure Thing Principle is analogous to dominance: indeed, it has a version of statewise dominance as a special case.[9] But it is stronger. The Countable Sure Thing Principle doesn't require that B is *sure* to turn out at least as well as A. Rather, it requires that B and A can be split up into corresponding uncertain *prospects*, where each B-prospect is at least as good as its corresponding A-prospect.[10]

Lin's preferences violate the Countable Sure Thing Principle. Consider the A and B gambles again. The events

The first heads for the A gamble is on the nth flip

[8] To be more careful, it is equivalent to the principle that the Sure Thing Principle holds for each conditional preference relation. The equivalence turns on some background assumptions: in particular, Completeness, as well as the assumption sometimes called "Consequentialism," after Machina (1989): if A and A' have the same outcome for each state in E, then $A \sim_E A'$.

[9] Suppose A and B are discrete options, in the sense that they each have only countably many different possible outcomes, each with positive probability. In this case, there is a regular partition $\mathcal{E}$ such that A and B are both constant on each event $E \in \mathcal{E}$. If B statewise dominates A, then we have $A \lesssim_E B$ for each $E \in \mathcal{E}$, and $A <_E B$ for some $E \in \mathcal{E}$, and thus by the Countable Sure Thing Principle, $A < B$.

[10] For careful and illuminating discussion of this contrast, see Buchak (2013, 162ff).

(for each number n) form a regular partition. Conditional on each of these events, Lin prefers the B gamble to the A gamble. But similarly, the events

The first heads for the B gamble is on the nth flip

form another regular partition, and conditional on each of *these* events, Lin prefers the A gamble to the B gamble. But Lin can't *both* strictly prefer the A gamble to the B gamble, and *also* strictly prefer the B gamble to the A gamble. So her preferences cannot satisfy the Countable Sure Thing Principle.

As we have mentioned, others have argued that St. Petersburg gambles (as well as two-envelope gambles) provide counterexamples to some principles about conditional expectations that are closely related to the Countable Sure Thing Principle. (See Footnote 7.) Some might think that they similarly provide counterexamples to the Countable Sure Thing Principle itself. We suspect that this gets things the wrong way around. The Sure Thing Principle is a centerpiece of orthodox decision theory—for powerful reasons, we think. As we will now argue, some of the best arguments that support the Sure Thing Principle straightforwardly generalize to support the *Countable* Sure Thing Principle as well. That's not to say these arguments are irresistible. There are *heterodox* decision theorists who reject the Sure Thing Principle and the arguments in its favor (Buchak 2013 is a prominent example). Those people are not our target. But for those who are persuaded by certain familiar arguments in support of the Finite Sure Thing Principle, there are nearly identical arguments that support the Countable Sure Thing Principle as well.

The first kind of argument arises from the same kind of strategic considerations that we have just discussed: preferences that violate the Sure Thing Principle display a kind of *dynamic inconsistency* (see Buchak 2013, ch. 6; Machina 1989; see also Markowitz 1952; Raiffa 1997). Lin's predicament dramatized this kind of inconsistency: by choosing to act according to her preferred strategy at each juncture, she ended up going for a *dominated* strategy overall. One influential motivation for the Sure Thing Principle is to avoid exactly this. In sufficiently idealized choice situations, if your preferences *satisfy* Savage's principle then you won't go for a dominated strategy; meanwhile, if your preferences *violate* Savage's principle, then there are (similarly idealized) decision situations in which you *will* go for a dominated strategy.[11] This seems like a strong point in favor of Savage's Sure Thing

[11] One of the idealizations involved is the assumption that you are sure you will be able to tell exactly what new information you have received (see Williamson 2002, sec. 10.6).

Principle.

An exactly parallel thing can be said in favor of the Countable Sure Thing Principle (see also Hammond 1998, sec. 8.7). The only difference is that we consider decision situations where you can learn the answer to a question that has infinitely many possible answers. And that kind of question is entirely normal. Before you look out the window, it isn't as if there is some finite list of all the things that might be outside, or exactly how they might look. The Finite Sure Thing Principle is supported by the maxim "Don't choose dominated strategies just because you gain new information." The Countable Sure Thing Principle is supported by the same maxim.

Not everyone accepts this non-dominance maxim. For example, some theorists are guided by a "time-slice rationality" picture, according to which there are simply no diachronic norms on decision-making, and thus no norms constraining one's choice of *strategy* in a multiple-stage game (see Moss 2015; Hedden 2015; Buchak 2015, 853–56.). The idea is that your acts at two different times are analogous to the acts of two different players in a game like the Prisoner's Dilemma (Hedden 2015, 13). Each player proceeds rationally, but the combination of two individually rational moves can lead to ruin. But insofar as this view makes trouble for the dynamic consistency argument for the Countable Sure Thing Principle, it makes just as much trouble for the dynamic consistency argument for the *Finite* Sure Thing Principle.[12] So this is not a way of driving a wedge between them.

Arntzenius, Elga, and Hawthorne (2004) do argue specifically, "In infinite cases, rationality does not require one to choose one's dominant options" (p. 262. They even put it in a box!) But in fact, they give two quite different arguments for this, which it is instructive to separate. One of these arguments is importantly tied to infinity, but it does not apply to the kind of decision set-up that gets Lin in trouble; the other argument does apply to this kind of set-up, but it has nothing to do with infinity.

The first argument they give turns on the observation that, if there are *infinitely many available strategies*, then it can turn out that *every* strategy is dominated by some other strategy. In this case it doesn't seem especially blameworthy to go for a dominated strategy. (As they discuss, this is what is going in certain well-known puzzles such as that of McGee (1999).) But this is not what is going on in Lin's case: our argument for the Countable

[12]For what it's worth, we don't think that the "dynamic consistency" argument has to turn on diachronic norms at all—instead, we can appeal to a norm about *hypothetical* preferences. What matters isn't what you *will* do, but what you *endorse* doing.

Sure Thing Principle only turns on it being a bad idea to choose a strategy which is dominated *by a strategy which is not itself dominated*.[13]

The second argument Arntzenius, Elga, and Hawthorne offer, against certain other infinite Dutch books (such as a variant of the two-envelope game, p. 272), turns on a completely different idea. They claim that in a two-round game, even if your sequence of choices leads to ruin,

> The situation is not one where a failure of rationality is displayed. For how can the reasoning at either round be faulted? The situation is rather one where the rational person is punished. Lacking the ability to control his future self, the earlier self launches a sequence of actions that can foreseeably be improved upon but is foreseeably unavailable given the rational dispositions of his future self. (2004, 273)

They go on to draw the analogy between diachronic choices and a two-player game: in fact, this is precisely the "time-slice" idea we considered above. But *this* kind of consideration has nothing special to do with *infinite* cases. If it is correct that choosing a dominated *strategy* cannot display a rational failing that does not reduce to a rational failing at some stage or other, then this is a problem for diachronic choice arguments even when infinity *isn't* involved—and thus it undermines the argument for the Finite Sure Thing Principle just as well.

Here is another style of argument for the Sure Thing Principle, which turns on very closely related technical facts. Good's Theorem tells us that the orthodox expected utility maximizer will always prefer to base her decisions on more information, rather than less: the *value of information* for her is always non-negative (again, in suitably ideal circumstances).[14] In contrast—as Buchak concedes—violating Savage's Sure Thing Principle has the "somewhat unpalatable upshot that more information is not always better for decision-making": sometimes you assign a question negative value of information (Buchak 2013, 199–200; see also Machina 1989). Here are two different ways of drawing out why this is an uncomfortable situation.

Unlike orthodox expected utility maximizers, people who violate the Sure Thing Principle can be guaranteed to reverse their preferences when they

[13] Compare Pettigrew (2013, 587, discussing accuracy dominance rather than utility dominance). Pettigrew (2019) calls this principle "Undominated Dominance".

[14] See Good (1966). The connection to the Sure Thing Principle does depend a bit delicately on details of formulation and background assumptions (see Schlee 1997).

Again, one of the idealizations is that you have perfect "access" to what your new information is (see Das, forthcoming; Dorst, forthcoming).

learn new information. In such cases, the choices you make now are sure to look bad to you in hindsight. This violates a principle that Arntzenius (2008) calls "Piaf's maxim": "a rational person should not be able to foresee that she will regret her decisions" (p. 277). As Arntzenius argues, one reason this is bad is because it leads to dynamic inconsistency of the sort we have already discussed. But it also seems bad on its own terms.

Arntzenius also argues for Piaf's maxim another way.

> Suppose you have a friend who has the same initial degrees of belief and the same utilities as you have. Suppose your friend acquires additional information, but he is not allowed to give you the information. He is only allowed to advise you how to act. Surely you should follow his advice. (2008, 279–80)

Again, orthodox expected utility maximizers follow this *expert deference* principle (as long as we are careful about some idealizing assumptions, such as that the pieces of information the friend might have form a partition). But again, those who violate the Sure Thing Principle do not, in general. People who violate the Countable Sure Thing Principle have the same problems. At the beginning of the story, the question of how the B gamble turns out has negative value of information for Lin. By her lights, committing to trading in the A ticket for $B + \$50$ is a better option than allowing her choice to be guided by the outcome of the B gamble. Thus she also violates the "no regret" and expert deference principles.

For our money, the dynamic consistency and value of information arguments provide some of the strongest reasons for believing the Finite Sure Thing Principle, and they are just about as weighty when it comes to the Countable version.[15] We are certainly *not* claiming that the Countable Sure Thing Principle is beyond question. We have already mentioned some reasons for doubt worth serious contemplation—decision theory is hard. But these reasons for doubt carry over to the *Finite* Sure Thing Principle as well. And despite these cogent doubts, the Finite Sure Thing Principle is a key part of orthodox decision theory. Those many decision theorists who *are* moved by the kinds of arguments we have discussed to accept the orthodoxy

[15] We can't survey all of the standard arguments for the Finite Sure Thing Principle here. We think some other arguments would probably carry over straightforwardly to the countable case, including the "small world" argument from Briggs (2015; drawing on Savage 1954, sec. 5.5) and the "reasons for betterness" argument of Harsanyi (1977). Others do not seem so straightforward, including "long run" arguments and (synchronic) "bookmaking" arguments (for discussion see Buchak 2013, ch. 7).

in finite cases should also be moved toward its analogue in infinite cases. If these arguments reveal something wrong with preferences that are not expectational in finite cases, then they also reveal something wrong with preferences like Lin's.

The prevailing methodology for investigating infinite decision theory is to take the orthodox theory of expected utility for granted in cases involving just finitely many possible outcomes, and then to attempt to extend this in some reasonable way to handle infinite cases like the St. Petersburg gamble (as well as other puzzles like the Pasadena game and the two-envelope paradox). The moral of this section is that the normative foundations of this project are in disrepair. Powerful reasons for accepting orthodox finite decision theory in the first place are also reasons *against* permitting preferences with the St. Petersburg structure.

In any case, it seems clearly worthwhile to investigate what decision theory will look like if the Countable Sure Thing Principle is a genuine normative requirement. This is our next task.

3 Bounded Utilities and Infinite Utilities

A natural diagnosis is that Lin cares too much about large amounts of money. Many historic and contemporary decision theorists insist that decision theory requires *bounded* utilities. We will argue that this response does not quite strike to the heart of the matter: while bounded real utilities are sufficient for well-behaved preferences over infinite prospects, they are not necessary.

The arguments that have been offered for bounded utilities are numerous, and generally underwhelming. First, there are considerations of psychological realism: do actual people have bounded utility functions?[16] We agree with Nover and Hájek (2004, 248) that this is far from clear. (See Section 6.) But more importantly, we also agree that, even if flesh and blood folks have only limited mundane ends, the preferences of merely possible people still matter for normative decision theory, insofar as those preferences can be rational.

We are also unpersuaded by direct appeals to intuition on these matters. It does seem counterintuitive to pay dearly for a gamble with a tiny chance

[16] "Bounded utility ... is more psychologically realistic than unbounded utility" (Sprenger and Heesen 2011, 826–27; they cite Aumann 1977).

of huge rewards (compare Bostrom 2009). But we are inclined to distrust these intuitions: humans' intuitive judgments are distorted by well-known biases. A classic experiment demonstrating *scope neglect* showed no difference in what one group was willing to pay to save 2,000 birds and what another group was willing to pay to save 20,000 (Desvousges et al. 1992; see also related research on "psychic numbing", such as Västfjäll et al. 2014). Similarly, our judgments are susceptible to *framing effects*—for example, we care more about losses than about unrealized gains (Kahneman, Knetsch, and Thaler 1991). St. Petersburg style gambles create the conditions for a perfect storm of cognitive distortion.

Another kind of theorist might restrict attention to bounded utility functions for the sake of technical convenience: after all, these are the conditions under which the standard theory of expected utility can be straightforwardly applied (compare for example Colyvan 2006, 697). But we're not in the business of making theorists' work technically easier—as will be evident. Our goal is to understand the normative limits on preference; this goal is not advanced by artificially truncating the space of possibilities (compare Hájek and Nover 2006, 708–10).

This brings us to what we are really looking for: *normative* arguments for bounded utilities. Joyce (1999, 37) offers one such argument, in two stages. In the first stage, Joyce presents the standard argument that if you have unbounded *finite* utilities (and preferences over a rich enough space of prospects), then there is a St. Petersburg style prospect that you will assign *infinite* utility. In the second stage, drawing on Hájek (2003), Joyce argues that assigning infinite utility to any prospect leads to violations of dominance reasoning.

We reject the second stage of this argument. Indeed, as we will show, while there *is* something wrong with St. Petersburg style prospects, it is *not* that they have infinite utility. Hájek and Joyce's dominance argument really only applies to a specific way of modeling infinite utilities—*extended real utilities,* which augment the standard real numbers with just two infinite values $+\infty$ and $-\infty$. This model collapses many important distinctions: for example, it assigns the very same value to a gamble that has infinite value for sure as a gamble that just has a chance of giving infinite value. The correct moral of this argument (which Hájek (2003) already recognized) is not that no prospect can have infinite value, but rather that infinite values should not be modeled by the extended reals. Finer distinctions between different infinite prospects are required. In Section 4 we will show how to model prospects with infinite utility while respecting dominance reasoning—and moreover, the generalized kind of dominance reasoning expressed by the Countable Sure Thing Principle.

Here is another way of putting things. The claim that no prospect has infinite utility is captured by the *Archimedean Axiom*: we will give a more precise statement in the next section, but the intuitive idea is that the values of any two prospects stand in some finite proportion to one another. So the basic structure of Joyce's argument goes like this: (1) the St. Petersburg prospect violates the Archimedean Axiom; (2) the Archimedean Axiom is a requirement of rationality (see also Fine 2008). We reject (2).

It's true that the Archimedean Axiom is a standard assumption; this is because it has the advantage of technical convenience: "the only motivation ... is that [it makes] the math easier" (Isaacs 2014, 12; see also Hájek and Nover 2008, 649ff.) Making the math easier is not our goal. Our goal is to understand the limits on rational preference.

If the Archimedean Axiom is not a rational requirement, then *bounded real utilities* are not rationally required either—for in fact, if the Archimedean Axiom is violated, then some prospects do not have utilities which are representable by real numbers at all. The correct characterization of the normative requirement imposed by the Countable Sure Thing Principle is *not* that preferences must be representable by bounded finite utilities; rather, it is something a bit subtler.

4 Limited Preferences

In fact, there is a natural constraint on the structure of preferences, which is broadly analogous to having bounded cardinal utility, but which does not require the Archimedean Axiom. This constraint is both necessary and sufficient for avoiding the problems Lin faced in Section 2. We will start by introducing a convenient technical framework based on von Neumann and Morgenstern's lotteries. In this context, the Countable Sure Thing Principle takes a new form as an infinitary generalization of von Neumann and Morgenstern's *Independence* axiom—we call it *Countable Independence*. Next, we will explain the basic structural feature of the St. Petersburg gamble that clashes with Countable Independence. Finally, we will use this insight to state the appropriate generalization of the "boundedness" idea, and sketch how it follows from Countable Independence, while ordinary boundedness does not.

First, the framework. So far we have been thinking about *options* as functions from states to outcomes (following Savage). But if we take for granted that the state space has a fixed probability measure, and we also take for granted that all that ultimately matters for preference is the probability that is assigned to each outcome, then we can substantially simplify our framework, going back to the classic "lotteries" of von Neumann and Morgenstern (1944)—modestly generalized to allow lotteries with countably infinitely many outcomes.[17] Let a *prospect* be a function from outcomes to probabilities with values that sum to one.

In this setting, the Countable Sure Thing Principle can be restated in terms of *mixtures* rather than *partitions.* Given a probability measure on the state space, we can associate each Savage-style option with a von Neumann-and-Morgenstern-style prospect, which simply tells you the probability of getting each outcome from that option. A partition is a way of splitting up an option into conditional options—one associated with each event in the partition. We can analogously split up the corresponding prospect into "conditional prospects." If we do this, the unconditional prospect is a *mixture* of the conditional prospects: that is, a probability-weighted average, where the weights are given by the probabilities of the events in the partition.[18]

Now suppose that preferences between options are determined by preferences

[17] While these assumptions are standard, Seidenfeld, Schervish, and Kadane (2009) prove a result that casts some suspicion on them: these assumptions are incompatible with unbounded utilities, given a few additional modest-looking assumptions. But here we have already taken the Countable Sure Thing Principle for granted, and this principle independently rules out unbounded real utilities. So their result, while extremely interesting, is no threat to our present project. (See also Lauwers and Vallentyne 2016, 2017.) Meacham (forthcoming) turns this same argument from Seidenfeld et al against the principle that strict preference is negatively transitive, which amounts to an argument against Completeness. Again, though, this argument is no threat to our use of Completeness in this context, and for the same reason: we have already independently ruled out the unbounded real utilities on which the argument relies.

[18] More precisely, if A is a function from states to outcomes with countably many different values, the associated prospect $[A]$ assigns each outcome x the probability of the set of states $A^{-1}(x)$. Similarly, for any non-null event E, the *conditional* prospect, $[A \mid E]$ assigns each outcome x the *conditional* probability of $A^{-1}(x)$ given E. If $E_1, E_2, \ldots$ is a regular partition, then it follows from the probability calculus that for each outcome x,

$$[A](x) = \sum_i p_i \cdot [A \mid E_i](x)$$

between *prospects.*[19] Then the Countable Sure Thing Principle amounts to the following principle about discrete mixtures (see Blackwell and Girshick 1954; Hammond 1998, sec. 8.4):[20]

Countable Independence. For any prospects $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$, and any probabilities $p_1, p_2, \ldots$ that sum to one, if $X_1 \lesssim Y_1$, $X_2 \lesssim Y_2$, $\ldots$, then

$$\sum_i p_i \cdot X_i \lesssim \sum_i p_i \cdot Y_i$$

If furthermore $X_j < Y_j$ for some j such that $p_j > 0$, then

$$\sum_i p_i \cdot X_i < \sum_i p_i \cdot Y_i$$

The distinctive axiom of von Neumann and Morgenstern's decision theory is called *Independence.* Countable Independence stands to von Neumann and Morgenstern's Independence axiom just as the Countable Sure Thing Principle stands to Savage's Sure Thing Principle. It amounts to generalizing the principle from *finite* mixtures of prospects to *countably infinite* mixtures of prospects. In the framework where we assume that probabilities are given, and that the probabilities of outcomes determine the choiceworthiness of each option, the very same arguments that support the Countable Sure Thing Principle also support Countable Independence.

From here on, we will treat Countable Independence as a constraint. What does this tell us about what preferences should be like?

The fundamental peculiarity of the St. Petersburg prospect is that it is worth *more* than any outcome it could have. (As Hájek and Nover (2006, 705) quip, we have a reverse Lake Wobegon effect: "all the payoffs are below average.") Call a prospect X *proper* iff it has some possible outcomes x^- and x^+ such that $x^- \lesssim X \lesssim x^+$; otherwise X is *improper.*[21]

Improper prospects clash directly with Countable Independence. Suppose X is a prospect that assigns probabilities $p_1, p_2, \ldots$ to outcomes $x_1, x_2, \ldots$.

[19] That is, $A \underset{E}{\lesssim} B$ iff $[A \mid E] \lesssim [B \mid E]$.

[20] The implication from Countable Independence to the Countable Sure Thing Principle is straightforward; the other direction requires a further "plenitude" assumption to guarantee that each mixture of prospects really does correspond to a partition of some option.

[21] We don't bother to distinguish outcomes from single-outcome prospects when this won't lead to confusion.

We can think of X as a countable mixture in two different ways. First, it is a mixture of the one-outcome prospects $x_1, x_2, \dots$ in the obvious way. Second, it is also a mixture of infinitely many copies of X itself. If X is improper, this means that X is strictly better than each outcome x_i. But then Countable Independence would require that X is strictly better than X. (The argument proceeds the same way if X is strictly worse than each outcome x_i instead.)

So Countable Independence rules out St. Petersburg-like improper prospects. Which preferences give rise to these? The standard St. Petersburg gamble uses a *geometrically* growing sequence of outcomes $x_1, x_2, \dots$. The outcome x_2 is at least twice as good as x_1, and x_3 is at least twice as good again, and so on. To spell out what this means in this context—where all we are given directly is ordinal preferences over prospects—first we need to choose a baseline. Then we can say that (for example) a trip to the aquarium is twice as good as a trip to the park, relative to the baseline of staying home, iff a fair coin flip that takes you to the aquarium on heads and otherwise leaves you home is as good as going to the park for sure. In general, for prospects X, Y, and Z (where $X > Z$), say that Y *is at least twice as good as* X *with respect to the baseline* Z iff $X \lesssim \frac{1}{2} \cdot Y + \frac{1}{2} \cdot Z$. Our key "anti-St.-Petersburg" condition says that there are no sequences of prospects that grow at least this fast.[22]

Limitedness. There is no infinite sequence of prospects $X_1, X_2, X_3, \dots$ such that X_2 is at least twice as good as X_1, X_3 is at least twice as good as X_2, and so on, with respect to some baseline Z.

Countable Independence implies Limitedness. The idea is that any "supergeometric" sequence of prospects can be used to construct a St. Petersburg-like improper prospect (see Russell 2020).

We should take special note of what Limitedness tells us about *infinite* prospects—since this is where it differs from ordinary boundedness. One way that a prospect might be at least twice as good as another is if it is infinitely better. So Limitedness also tells us that there is no infinite

[22]Note that it is consistent with Limitedness that there are geometrically *decreasing* chains of prospects, which get closer and closer in value to the baseline Z. On the other hand, Limitedness *does* rule out infinite sequences that are geometrically decreasing in a different sense, where each prospect is at least twice as *bad* as the previous, relative to a baseline Z. (See Appendix A.)

sequence of prospects that grows *infinitely fast*—for in fact, any sequence like that would also be St. Petersburg-like.

What does this even mean? Remember, if the aquarium is twice as good as the park relative to staying home, this means that a gamble that gives you chance $\frac{1}{2}$ of the aquarium or else staying home is at least as good as going to the park for sure. We can likewise say that the aquarium is n times as good as the park (relative to staying home) iff the gamble which gives you probability $\frac{1}{n}$ of the aquarium and otherwise home is as good as the park. Extending this idea, we can say that the aquarium is *infinitely* better than the park iff a gamble with $\frac{1}{n}$ chance of the aquarium is better than the park for *every* positive value of n (compare Bartha 2007, 18).

What the *Archimedean Axiom* says is that this can never happen: for any prospects X and Y and any baseline Z, there is some n such that X is no more than n times as good as Y relative to Z. But we do not take this principle for granted. Our project is to explore the limits on decision theory imposed by Countable Independence: as we will show, some non-Archimedean preferences are well within these limits.

Suppose there is a "baseline" outcome z, and outcomes $x_1, x_2, \ldots$, such that for each n, x_{n+1} is *infinitely* better than x_n with respect to z. Then consider a prospect X that assigns positive probability to each of these outcomes (and no others).[23] Then X must be improper. For each n we can consider a gamble X_n, which gives you probability p_n of x_n, and otherwise gives you the baseline z. Since x_{n+1} is infinitely better than x_n (with respect to z), X_{n+1} is better than x_n. And by dominance, X_{n+1} is worse than X: its outcome is strictly worse than X's in every case except the $(n+1)$th, where it is the same. By transitivity, X is strictly better than each of its outcomes, which means X is improper.

Limitedness rules out this kind of case: we cannot have an infinite sequence of ever-infinitely-better outcomes. But this is *not* because it rules out infinite values altogether. In fact, Limitedness, unlike bounded real utilities, does not require the Archimedean Axiom: infinite values are allowed.

Here is a simple model to illustrate this. Suppose that the aquarium, the park, and home are the only three outcomes. Then we can represent each prospect as a pair of two numbers q and p such that $q + p \leq 1$, where q represents the probability of going to the aquarium and p represents the

[23] Unlike in the original St. Petersburg gamble, the precise assignment of probabilities does not matter in this case.

probability of going to the park. (Intuitively, we can think of staying home as the "zero" outcome.) Suppose that for Maggie, the aquarium is infinitely better than the park, with respect to staying home. Her preferences are modeled by this *lexical* order:

$$(q,p) \lesssim (q',p') \quad \text{iff either } \begin{cases} q < q', \text{or} \\ q = q' \text{ and } p \le p' \end{cases}$$

Maggie's preferences satisfy Limitedness,[24] but they cannot be represented by bounded real utilities; indeed, they cannot be represented by real utilities at all. So Limitedness is a weaker constraint than that imposed by bounded real utilities.[25]

Furthermore, unlike ordinary bounded utilities, Limitedness is supported by a powerful normative argument. As we have discussed, preferences that violate Limitedness thereby violate Countable Independence; such violations lead to all the interrelated problems we discussed in Section 2: preferring dominated strategies, forseeable regret, failing to defer to experts, and assigning negative value to information. These are analogous to the problems that come with violating von Neumann and Morgenstern's standard Independence axiom. Only Limited preferences avoid these problems.

5 The Equivalence Theorem

In Section 1 we reported some bad news for decision theorists: preferences like Lin's are incompatible with what we called the *Extended Outcome Principle*, and thus they do not admit any reasonable *utility* representation. If

[24] The basic reason is that there isn't very much room for doubling along either axis. For simplicity consider the baseline $Z = (0,0)$. (The argument easily generalizes.) Call a prospect (q,p) *big* if $q > 0$, and otherwise call it *small.* Each time you go from one prospect to another which is at least twice as good (with respect to Z), you have to do one of three things: (1) go from a small prospect to another small prospect which at least doubles p; (2) go from a small prospect to a big prospect; or (3) go from a big prospect to another big prospect which at least doubles q. Each of these three kinds of step can only be taken finitely many times within the model.

[25] Limitedness can be split up into two properties, which taken together are equivalent to it (given Finite Independence). The first property is a straightforward boundedness condition on (relatively) *finite* prospects, and the second property is a constraint on the structure of (relatively) *infinite* prospects. (1) For any pair of prospects Z and X, the set of numbers n such that some prospect Y is at least n times as good as X with respect to Z, but not infinitely better than X, is bounded. (2) There are no prospects Z and $X_1, X_2, \ldots$ such that each X_{n+1} is infinitely better than X_n with respect to Z.

Lin's preferences were rational, this would deprive normative decision theory of a useful tool. But in Section 2 we gave arguments that Lin's preferences are *not* rational, since they violate the Countable Sure Thing Principle and Countable Independence. If that's right, the useful tool may still be available.

We have good news: it is. In fact, Countable Independence *guarantees* that an abstract utility representation is possible, because Countable Independence implies the Extended Outcome Principle. Countable Independence guarantees that we can enrich the space of outcomes (to fill in any gaps), in a way that still satisfies Countable Independence, such that every prospect is equivalent in value to some extended outcome. We can think of these extended outcomes as abstract "utilities."[26]

These very abstract "utilities" aren't especially useful on their own: you can't do convenient calculations with them, or easily tell whether a given preference relation is amenable to being thus represented. But we have more good news. In addition to this very abstract representation, we can also provide you with something much more concrete and tractable: Russell (2020) proves a representation theorem of a more familiar kind for preferences that satisfy Countable Independence. This lets you apply decision theory using a straightforward and familiar kind of recipe. First, we present you with a menu of acceptable utilities: your job is to freely choose a utility from this menu for each possible outcome. Second, we give you a formula that lets you calculate what the *expected* utility must be for each prospect over these outcomes. Then you know that your preferences are coherent (by the lights of these axioms) as long as they order prospects according to their expected utilities.

All of this sounds familiar and comforting. The key difference between this representation theorem and the ones you know is what the menu of utilities looks like. Utilities are not real numbers. We took the first step in Section 4: Maggie's preferences over aquarium-park-home prospects were represented by pairs of *two* bounded real numbers, lexically ordered. There's no need to stop there. You can similarly represent an outcome with a *sequence* of real numbers from the interval $[-1, 1]$—even an infinite sequence. But there's no

[26]Further details are in Appendix A, but the technical idea is simple: we can treat prospects themselves as "utilities." The "utility" of an outcome is just its sure prospect. We then have to extend the preference relation to apply to "higher-order prospects"—probability distributions over probability distributions over outcomes. There is an elegant way of doing this, by taking mixtures—but this depends on Countable Independence to work.

need to stop *there*, either. Simple omega-sequences of real components don't do justice to all coherent preferent relations. In general, each outcome can be represented by a *transfinite* sequence of real numbers from the interval $[-1, 1]$: that is, a sequence of numbers whose length can be any ordinal number. As before, these sequences are ordered lexically.

This space of utilities is truly vast: in fact, it forms a proper class, rather than a set. But in practice this is not important. For any set of outcomes, the sequences that are assigned as utilities to outcomes in that set will all have lengths bounded by some particular ordinal α. So you can get away with restricting the menu of utilities to sequences of length α.

Once you have picked a utility for each outcome from this grand space of alternatives, calculating the *expected* utility for a prospect is actually very simple. If we have assigned probabilities to some α-sequences, then we can first take the expected value of their first components with respect to these probabilities (this is guaranteed to be well-defined), then likewise the second component, and so on down the ordinals. Usually you don't have to go on all the way to the end. In order to *compare* the expected utilities of two prospects, you just have to go as far as the first component where their expectations differ. Since the order is lexical, whichever prospect comes out ahead at this point comes out ahead overall.

Notice that, while in some respects this is a very permissive theory, it puts interesting constraints on the structure of value. One constraint is that you can't have infinitely many upward-ascending regimes of infinite value. That is, while it's fine for some things to be infinitely better than others, and for other things to be infinitely better still, any such sequence must come to an end. The ordinals, after all, are *well-ordered*: this means that any non-empty set of ordinals includes some earliest ordinal. Since earlier ordinals represent *lexically more important* regimes of value, this means that any non-empty set of ascending regimes of value has to max out at some *highest* regime. On the other hand, there is no such constraint in the other direction. It is perfectly permissible (for all our axioms say) to lexically prioritize today's happiness over tomorrow's, and tomorrow's over the next day's, and so on into the infinite future. The reverse, however, is ruled out.[27]

[27]More exotic structures are also possible. For example, the axioms permit a kind of particularly bizarre "future Tuesday indifference" (Parfit 1984, 123–24), where each future Tuesday is lexically prioritized over those that follow it, and likewise each future Monday is lexically prioritized over those that follow it, but also *all* future Mondays are lexically more important than *all* future Tuesdays. The sequence-utilities that we would use to represent these preferences would be double-omega-sequences. This illustrates that two

We find this structure surprising. We have no hope of motivating it by direct appeal to intuitions. Rather, it looks like this is just how things have to go, given Countable Independence. We didn't choose a natural-looking conclusion and then attempt to argue for it. This strange place is where the argument leads.

The representation theorem of Russell (2020) shows that the preference orders that satisfy Countable Independence (as well as transitivity and completeness) are in fact precisely those that are representable by lexicographically ordered transfinite sequences of bounded real utilities. It also follows from results in that paper that these are also precisely the orders that satisfy both Limitedness and Finite Independence.[28] In Appendix A we also show that the Extended Outcome Principle belongs in this circle of equivalent statements, thus establishing the following.

> **The Equivalence Theorem.** For any preference order on the set of prospects,[29] the following are equivalent:
>
> 1. Countable Independence
> 2. The Extended Outcome Principle
> 3. Limitedness and Finite Independence
> 4. Representability by lexicographically ordered ordinal sequences of bounded real utilities.

The Equivalence Theorem tells us that the different threads we have followed are tightly intertwined. When we first met the Extended Outcome Principle

regimes of value (like next Monday and next Tuesday) can have infinitely many other regimes that fall in between them in importance (all the later Mondays). This is fine, as long as the infinite sequence only goes "downward" in importance.

[28]A technical point worth calling attention to here is that Limitedness speaks of *prospects*, not just *outcomes*. As it turns out, the analogous condition that merely rules out geometric sequences of *outcomes* is too weak. Here is a model in which there are no supergeometric sequences of outcomes (with respect to any baseline), but which is not Limited. Let outcomes be represented by pairs of real numbers of the form $(x, 1/x)$ for $0 < x < 1$. We order prospects by comparing the expectations of each coordinate lexically. There are no supergeometric sequences of outcomes, because the first coordinate is bounded and comparisons on the first coordinate are lexically prioritized. (The model also includes no improper prospects.) Even so, in this model we can construct a geometrically growing sequence of *prospects*: we can consider mixtures of one outcome of the form $(2^{-n}, 2^n)$ together with the outcome $(1, 1)$, with the probabilities in each mixture carefully chosen so that each of them has the same expectation $\frac{1}{2}$ for its first coordinate (in particular, $p_n = \frac{2^n}{2^{n+1}-2}$). While the expectations of the first coordinate of these mixtures are fixed, the expectations of the *second* coordinate zoom off geometrically to infinity.

[29]We assume transitivity and completeness.

in Section 1, it seemed like the sort of thing that would be theoretically convenient to have, but by no means indispensable for normative decision theory. But that's not how it turns out. For in fact, the Extended Outcome Principle and Countable Independence are equivalent. So the possibility of a theoretical project generalizing standard expected utility is intimately tied to a *coherence* constraint on choices. Similarly, in Section 4 we sketched why Limitedness (unlike ordinary bounded utilities) is necessary for Countable Independence. The Equivalence Theorem tells us that Limitedness is also sufficient.

The Equivalence Theorem means we can characterize and apply decision theory equally well using any of three tools: a coherence constraint relating conditional and unconditional preferences, a "qualitative" structural constraint on preference relations, or a concrete utility representation. The theorem tells us just what it takes for preferences over countable von Neumann-Morgenstern lotteries to be well-behaved—and escape Lin's tribulations from Section 2—in a way that allows us to straightforwardly extend core ideas of standard decision theory to countably infinite non-Archimedean lotteries.

6 Antithesis

The Equivalence Theorem presents a tidy picture. But we'll conclude with a confession: the situation seems messier. We are convinced that the arguments in favor of the Countable Sure Thing Principle and Countable Independence are quite strong—indeed, they are very close variations on what strike us as some of the strongest arguments decision theory has to offer in support of its central orthodoxy: expected utility theory. However, these are not all the arguments there are. There also seem to be some powerful things to be said on behalf of preferences that are not Limited—which, given the mathematics, amount to powerful arguments *against* the Countable Sure Thing Principle. To be honest, we're not sure what to think.

Many authors have pointed out that innocuous-seeming preferences cannot be represented by bounded real-valued utility functions. (It isn't hard to see that sequence-valued utility functions are no help with these cases.) For example, Vann McGee argues,

> [I]t seems altogether unreasonable to suppose that our utility scale ought to be bounded. To suppose so requires that there is some time

> t such that we should be willing to trade a thousand years of utter happiness starting at t for a Snickers bar today, and it is implausible to imagine that rationality not only allows but demands that there should be such a t. ... [I]t's hard to believe ... that reason is so opposed to prudence as to demand that, as we look to the farther and farther distant future, our attitude toward our long-term welfare should approach complete indifference (McGee 1999, 263; compare Nover and Hájek 2004, 248).

This argument takes for granted that a Snickers bar contributes the same cardinal utility boost no matter what your future pattern of happiness may be. Other arguments don't rely on assumptions about cardinal value, but instead turn on direct comparisons of prospects. For example:

> Suppose that Methuselah were to truly report, 'For any real number p greater than zero, there exists a natural number n such that I prefer chance p of extending my life by n years, to receiving $1 with certainty.' Methuselah's beliefs and preferences seem to be perfectly coherent. But his utility scale must be unbounded (Arntzenius, Elga, and Hawthorne 2004, 270).[30]

We agree: the Snickers preferences seem weird, and Methuselah's seem fine.[31]

We don't entirely know what to make of these arguments. Before working through the ideas in this paper, we were quite moved by them. But preferences like Methuselah's allow us to construct a St. Petersburg style improper prospect, which will grant him some unknown finite lifespan, and for which he would happily stake *any* finite lifespan. This in turn gets him into trouble, choosing dominated strategies, paying to avoid information, and so on. Cavalierly accepting such consequences as the cost of doing business threatens to undermine orthodox decision theory altogether. Some might welcome this result. We are troubled by it.

We have made a case for the Countable Sure Thing Principle and Countable Independence. But in light of the rather radical consequences these

[30] We assume that no finite lifespan is *infinitely* better for Methusaleh than a dollar; otherwise these preferences could be modeled by sequence-utilities.

[31] We are especially moved by variants of these arguments recast in terms of impartially saving large numbers of people from harm. Apparent clashes between rationality and prudence (as McGee put it) are troubling; apparent clashes between rationality and morality even more so. See also Basu and Mitra (2003).

principles have (particularly for the prudence and ethics of large numbers), we think it is entirely reasonable to treat them both with suspicion. More philosophy is called for. Still, figuring out how decision theory goes *if they are right* seems clearly worthwhile: this naturally extends von Neumann and Morgenstern's classic decision theoretic framework to infinite prospects. We have also advanced this project: as it turns out, decision theory with Countable Independence amounts to expected utility theory using bounded ordinal sequences as utilities. It also amounts to decision theory with the constraint that value cannot grow indefinitely (as expressed by Limitedness). And it also turns out to be the *only* decision theory for infinite prospects that has the same basic form as expected utility theory (as expressed by the Extended Outcome Principle). This isn't a decision theory we expected, but it's the one we get.

A Appendix

This appendix contains a more official statement of the *Extended Outcome Principle* from Section 1, and a sketch of the remaining parts of the proof of the Equivalence Theorem from Section 5 that are not already proved by Russell (2020): namely, that Countable Independence implies the Extended Outcome Principle, which in turn implies Limitedness and Finite Independence. This in turn implies representability by bounded ordinal sequences, which by the representation theorem of Russell (2020) is equivalent to Countable Independence.

Let X be a non-empty set of *outcomes*, and let PX be the set of probability mass functions on X: that is, functions $\lambda : X \to [0, 1]$ such that $\sum_{x \in X} \lambda(x) = 1$. We call this the set of X*-prospects* or *lotteries*. For $x \in X$, let $\overline{x} \in PX$ be the single-outcome lottery that assigns probability one to x.

A **preference order** is a total preorder on PX: that is, a relation $\lesssim$ which is reflexive, transitive, and complete in the sense that for any $\lambda, \mu \in PX$, either $\lambda \lesssim \mu$ or $\mu \lesssim \lambda$.

A preference order satisfies **Countable Independence** iff, for any probability distribution $p \in P\omega$ and any sequences of lotteries $\lambda_1, \lambda_2, \ldots$ and $\mu_1, \mu_2, \ldots$ in PX:

(a) if $\lambda_i \lesssim \mu_i$ for each i, then $\sum_i p_i \cdot \lambda_i \lesssim \sum_i p_i \cdot \mu_i$, and

(b) if furthermore $\lambda_i < \mu_i$ for some i such that $p_i > 0$, then $\sum_i p_i \cdot \lambda_i < \sum_i p_i \cdot \mu_i$.

The order satisfies **Finite Independence** iff (a) and (b) hold for each $p \in P\omega$ with finite support. The order satisfies **Dominance** iff (a) and (b) hold in the restricted case where each λ_i and μ_i is a single outcome. That is to say, Dominance says that for each $p \in P\omega$ and any sequences $x_1, x_2, \ldots$ and $y_1, y_2, \ldots$ of outcomes in X,

(a) if $\overline{x_i} \lesssim \overline{y_i}$ for each i, then $\sum_i p_i \cdot \overline{x_i} \lesssim \sum_i p_i \cdot \overline{y_i}$, and

(b) if furthermore $\overline{x_i} < \overline{y_i}$ for some i such that $p_i > 0$, then $\sum_i p_i \cdot \overline{x_i} < \sum_i p_i \cdot \overline{y_i}$.

A preference order on PX satisfies the **Extended Outcome Principle** iff there is a set X^* (of "extended outcomes"), a total preorder $\lesssim^*$ on PX^*, and a function $u : PX \to X^*$, such that

(a) Intuitively, u maps each lottery to an extended outcome that is equivalent to it. More precisely, for each lottery $\lambda \in PX$, let

$$u^*(\lambda) = \sum_{x \in X} \lambda(x) \cdot \overline{u(\overline{x})}$$

This is the "lifted" lottery in PX^* that assigns the same probability to each extended outcome $u(\overline{x})$ as λ assigns to the original outcome x. Then for each $\lambda \in PX$,

$$\overline{u(\lambda)} \sim^* u^*(\lambda)$$

(b) The extended order respects the original order of lotteries. That is, for $\lambda, \mu \in PX$,

$$\lambda \lesssim \mu \qquad \text{iff} \qquad \overline{u(\lambda)} \lesssim^* \overline{u(\mu)}$$

(c) The extended order $\lesssim^*$ satisfies Finite Independence and Dominance.

Lemma 1. Countable Independence implies the Extended Outcome Principle.

Proof Sketch. The basic idea is that we can think of lotteries themselves as "extended outcomes." Let $X^* = PX$, let $u : PX \to PX$ be the identity function. For $\lambda^* \in PX^* = P(PX)$, we can let

$$M\lambda^* = \sum_{\nu \in PX} \lambda^*(\nu) \cdot \nu$$

Then we can define the extended order so that for $\lambda^*, \mu^* \in PX^*$,

$$\lambda^* \lesssim^* \mu^* \qquad \text{iff} \qquad M\lambda^* \lesssim M\mu^*$$

It is straightforward to check condition (a). In fact,

$$Mu^*(\lambda) = \lambda = M\overline{u(\lambda)}$$

So $\overline{u(\lambda)} \sim^* u^*(\lambda)$. This likewise implies condition (b):

$$\lambda \lesssim \mu \qquad \text{iff} \qquad M\overline{u(\lambda)} \lesssim M\overline{u(\mu)} \qquad \text{iff} \qquad \overline{u(\lambda)} \lesssim^* \overline{u(\mu)}$$

For condition (c) we can check that, because the original order on PX satisfies Countable Independence, this extended order on PPX does as well. Finite Independence and Dominance both follow from this.

□

Lemma 2. The Extended Outcome Principle implies that for any lottery λ, there is some outcome x in the support of λ such that $\lambda \lesssim \overline{x}$.

Proof. Suppose that $\lambda > \overline{x}$ for each outcome x in the support of λ. So $\overline{u(\lambda)} > \overline{u(\overline{x})}$ for each such x, and thus by Dominance

$$\overline{u(\lambda)} = \sum_{x \in X} \lambda(x) \cdot \overline{u(\lambda)} > \sum_{x \in X} \lambda(x) \cdot \overline{u(\overline{x})} = u^*(\lambda)$$

This contradicts condition (a) of the Extended Outcome Principle.

□

Lemma 3. The Extended Outcome Principle implies Limitedness.

Proof Sketch. The core idea of this proof was sketched in Section 1: if a preference order is not Limited, then we can construct a St. Petersburg style lottery, which is better than any of its possible outcomes. This violates the Extended Outcome Principle.

Suppose there is a sequence $\lambda_1, \lambda_2, \ldots$ such that, with respect to baseline μ, λ_2 is twice as good as λ_1, λ_3 is twice as good as λ_2, and so on. We can let ω be the St. Petersburg style "extended" lottery in PX^* such that

$$\omega(u(\lambda_n)) = 2^{-n} \quad \text{for each } n$$

By essentially the same reasoning as in Section 1, we can argue that ω is better than each of its possible outcomes. We will skip the details (but compare Russell 2020, Lemma 4): basically, for each n, we can find a *truncated* lottery ω_n that agrees with ω for finitely many outcomes and assigns all the rest of its probability to the baseline value $u(\mu)$, and is such that

$$\omega_n \gtrsim \overline{u(\lambda_n)} \quad \text{for each } n$$

Since ω dominates each of these truncated lotteries ω_n,

$$\omega > \overline{u(\lambda_n)} \quad \text{for each } n$$

and these are all of the outcomes in its support, contradicting the previous lemma.

□

It is also straightforward to show that the Extended Outcome Principle implies Finite Independence (since it is part of the definition that this holds for "extended" lotteries).

To complete the proof of the Equivalence Theorem, the last thing to show is that Limitedness and Finite Independence suffice for the lexicographic representation presented by Russell (2020). This follows from results in that paper, as follows. First, any preference order that satisfies Finite Independence can be extended to a preference order defined on a *Banach space* (namely $\ell^1(X)$) (Russell 2020, Lemma 1). Second, if the original order also satisfies Limitedness, it straightforwardly follows that the extended order will also have the property (Russell 2020, Lemma 4) that there are no norm-bounded *supergeometric sequences* $0 < x_1, x_2, \ldots$, such that $2x_i < x_{i+1}$ for each i. From this it follows that every subspace of the Banach space has the property that each norm-bounded subset is order-bounded (Russell 2020, Lemma 5), and this property suffices for the remainder of the proof of the representation theorem.

References

Arntzenius, Frank. 2008. "No Regrets, or: Edith Piaf Revamps Decision Theory." *Erkenntnis* 68 (2). Springer: 277–97.

Arntzenius, Frank, Adam Elga, and John Hawthorne. 2004. "Bayesianism, Infinite Decisions, and Binding." *Mind* 113 (450). Oxford University Press: 251–83.

Arntzenius, Frank, and David McCarthy. 1997. "The Two Envelope Paradox and Infinite Expectations." *Analysis* 57 (1). Oxford University Press: 42–50.

Aumann, Robert. 1977. "The St. Petersburg Paradox: A Discussion of Some Recent Comments." *Journal of Economic Theory* 14 (2): 443–45. https://EconPapers.repec.org/RePEc:eee:jetheo:v:14:y:1977:i:2:p:443-445.

Bartha, Paul. 2007. "Taking Stock of Infinite Value: Pascal's Wager and Relative Utilities." *Synthese* 154 (1). Springer: 5–52.

Basu, Kaushik, and Tapan Mitra. 2003. "Aggregating Infinite Utility Streams with Intergenerational Equity: The Impossibility of Being Paretian." *Econometrica* 71 (5): 1557–63.

Blackwell, David, and M. A. Girshick. 1954. *Theory of Games and Statistical Decisions*. New York: John Wiley & Sons, Inc. https://onlinelibrary.wiley.com/doi/abs/10.1002/nav.3800010313.

Bostrom, Nick. 2009. "Pascal's Mugging." *Analysis* 69 (3). Oxford University Press: 443–45.

Briggs, R.A. 2015. "Costs of Abandoning the Sure-Thing Principle." *Canadian Journal of Philosophy* 45 (5): 827–40.

Broome, John. 1995. "The Two-Envelope Paradox." *Analysis* 55 (1). Oxford University Press: 6–11.

Buchak, Lara. 2013. *Risk and Rationality*. Oxford University Press.

———. 2015. "Revisiting Risk and Rationality: A Reply to Pettigrew and Briggs." *Canadian Journal of Philosophy* 45 (5): 841–62.

Chalmers, David J. 2002. "The St. Petersburg Two-Envelope Paradox." *Analysis* 62 (2). Oxford University Press: 155–57.

Chen, Eddy Keming, and Daniel Rubio. 2018. "Surreal Decisions." *Philosophy and Phenomenological Research*. https://doi.org/10.1111/phpr.12510.

Clark, Michael, and Nicholas Shackel. 2000. "The Two-Envelope Paradox." *Mind* 109 (435). Oxford University Press: 415–42.

Colyvan, Mark. 2006. "No Expectations." *Mind* 115 (459). Oxford University Press: 695–702.

———. 2008. "Relative Expectation Theory." *Journal of Philosophy* 105 (1). Journal of Philosophy Inc: 37–44.

Das, Nilanjan. Forthcoming. "The Value of Biased Information." *British Journal for the Philosophy of Science*, axaa003. https://doi.org/10.1093/bjps/axaa003.

Desvousges, William H., F. Reed Johnson, Richard W. Dunford, Kevin J. Boyle, Sara P. Hudson, and K. Nicole Wilson. 1992. *Measuring Nonuse Damages Using Contingent Valuation: An Experimental Evaluation of Accuracy*. Vol. 1992. Research Triangle Institute Research Triangle Park, NC.

Dorst, Kevin. Forthcoming. "Evidence: A Guide for the Uncertain." *Philosophy and Phenomenological Research*. https://doi.org/10.1111/phpr.12561.

Easwaran, Kenny. 2014. "Decision Theory Without Representation Theorems." *Philosophers' Imprint* 14.

Fine, Terrence L. 2008. "Evaluating the Pasadena, Altadena, and St Petersburg Gambles." *Mind* 117 (467). Oxford University Press: 613–32.

Good, I. J. 1966. "On the Principle of Total Evidence." *British Journal for the Philosophy of Science* 17 (4). Taylor & Francis: 319–21.

Hammond, Peter J. 1998. "Objective Expected Utility: A Consequentialist Perspective." In *Handbook of Utility Theory*, edited by Peter J. Hammond, Salvador Barberà, and Christian Seidl. Vol. I. Kluwer Academic.

Harsanyi, John C. 1977. "On the Rationale of the Bayesian Approach: Comments on Professor Watkins's Paper." In *Foundational Problems in the Special Sciences*, 381–92. Springer.

Hájek, Alan. 2003. "Waging War on Pascal's Wager." *Philosophical Review* 112 (1). Duke University Press: 27–56.

Hájek, Alan, and Harris Nover. 2006. "Perplexing Expectations." *Mind* 115 (459). Oxford University Press: 703–20.

———. 2008. "Complex Expectations." *Mind* 117 (467). Oxford University Press: 643–64.

Hedden, Brian. 2015. "Time-Slice Rationality." *Mind* 124 (494): 449–91.

Isaacs, Yoaav. 2014. "Duty and Knowledge." *Philosophical Perspectives* 28 (1): 95–110.

Joyce, James M. 1999. *The Foundations of Causal Decision Theory*. Cambridge University Press.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. 1991. "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias." *Journal of Economic Perspectives* 5 (1): 193–206.

Lauwers, Luc, and Peter Vallentyne. 2016. "Decision Theory Without Finite Standard Expected Value." *Economics and Philosophy* 32 (3): 383–407.

———. 2017. "A Tree Can Make a Difference." *Journal of Philosophy* 114 (1): 33–42.

Lee, Chunghyoung. 2013. "The Two-Envelope Paradox: Asymmetrical Cases." *Mind* 122 (485): 1–26.

Machina, Mark J. 1989. "Dynamic Consistency and Non-Expected Utility Models of Choice Under Uncertainty." *Journal of Economic Literature* 27 (4): 1622–68. http://www.jstor.org/stable/2727025.

Markowitz, Harry. 1952. "Portfolio Selection." *The Journal of Finance* 7 (1): 77–91.

McGee, Vann. 1999. "An Airtight Dutch Book." *Analysis* 59 (4). Oxford University Press: 257–65.

Meacham, Christopher. Forthcoming. "Difference Minimizing Theory." *Ergo: An Open Access Journal of Philosophy*.

Moss, Sarah. 2015. "Time-Slice Epistemology and Action Under Indeterminacy." In *Oxford Studies in Epistemology*, edited by Tamar Szabó Gendler and John Hawthorne, 172–94. Oxford University Press.

Nover, Harris, and Alan Hájek. 2004. "Vexing Expectations." *Mind* 113 (450). Oxford University Press: 237–49.

Parfit, Derek. 1984. *Reasons and Persons*. Oxford University Press.

Pettigrew, Richard. 2013. "Accuracy and Evidence." *Dialectica* 67 (4): 579–96. https://doi.org/10.1111/1746-8361.12043.

———. 2019. "Epistemic Utility Arguments for Probabilism." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2019. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2019/entries/epistemic-utility/.

Raiffa, Howard. 1997. *Decision Analysis: Introductory Lectures on Choices Under Uncertainty*. New York: McGraw-Hill College.

Russell, Jeffrey Sanford. 2020. "Non-Archimedean Preferences over Countable Lotteries." *Journal of Mathematical Economics* 88 (May 2020): 180–86.

Savage, Leonard J. 1954. *The Foundations of Statistics*. Wiley Publications in Statistics.

Schlee, Edward E. 1997. "The Sure Thing Principle and the Value of Information." *Theory and Decision* 42 (1). Springer Netherlands: 21–36.

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. Research Showcase @ Cmu: 329–40.

Sprenger, J., and R. Heesen. 2011. "The Bounded Strength of Weak Expectations." *Mind* 120 (479): 819–32.

Västfjäll, Daniel, Paul Slovic, Marcus Mayorga, and Ellen Peters. 2014. "Compassion Fade: Affect and Charity Are Greatest for a Single Child in Need." Edited by Claus Lamm. *PLoS ONE* 9 (6): e100115. https://doi.org/10.1371/journal.pone.0100115.

von Neumann, John, and Oskar Morgenstern. 1944. *Theory of Games and Economic Behavior*. Princeton University Press.

Williamson, Timothy. 2002. *Knowledge and Its Limits*. Oxford University Press. http://dx.doi.org/10.1093/019925656X.001.0001.