Complex Expectations

Alan Hájek and Harris Nover

In our (2004), we introduced two games in the spirit of the St. Petersburg game, the Pasadena and Altadena games. As these latter games lack an expectation, we argued that they pose a paradox for decision theory. Terrence Fine has shown that any finite valuations for the Pasadena, Altadena, and St. Petersburg games are consistent with the standard decision-theoretic axioms. In particular, one can value the Pasadena game above the other two, a result that conflicts with both our intuitions and dominance reasoning. We argue that this result, far from resolving the Pasadena paradox, should serve as a *reductio* of the standard theory, and we consequently make a plea for new axioms for a revised theory. We also discuss a proposal by Kenny Easwaran that a gamble should be valued according to its 'weak expectation', a generalization of the usual notion of expectation.

I can't imagine going on when there are no more expectations.
- Dame Edith Evans

1. Pasadena *redux*

Pasadena may be an agreeable town of orderly houses and manicured lawns, but the *Pasadena game* is the terror of Colorado Boulevard.[1] It resembles the St. Petersburg game in two important respects:

- a fair coin is tossed until it lands Heads for the first time;
- the payoffs grow in magnitude without bound.

But the Pasadena game also differs from the St. Petersburg game in two important respects:

- the former alternates rewards with punishments according to whether *n*, the

[1] With apologies to Jan and Dean and The Beach Boys.

number of tosses required, is odd or even, whereas the latter offers only rewards;

- the former's payoffs grow in absolute value as dollars—regarded as utiles—whereas the latter's payoffs grow as dollars.

So with probability the Pasadena game pays $\$(-1)^{n-1}2^n/n$, while the St. Petersburg game pays \$. As a result, there is an important similarity and an important difference in expected utility theory's evaluations of the games. Important similarity: the expectation series of each game fails to converge absolutely—replacing each term in the series with its absolute value yields a series that fails to converge. Important difference: the expectation series of the St. Petersburg game,

$$1 + 1 + 1 + \ldots$$

diverges, while that of the Pasadena game

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots$$

conditionally converges (converges, but not absolutely).[2] Consequently, expected utility theory judges both games to be problematic, but in two different ways. Famously, the St. Petersburg game's expectation is infinite, so it seems that it should be valued infinitely, and it is thus regarded as preferable to any finite reward, however large. Less famously, the Pasadena game's expectation is undefined, so it seems that it cannot be valued at all, and it is thus regarded as incomparable to any finite reward.

But things are not always as they seem. As we rehearse in section 2, Terrence Fine (REF) shows that consistent with the preference axioms of standard decision theory, both the Pasadena game and the St. Petersburg game can be valued at any real number whatsoever. In particular, the Pasadena game can be valued *above* the St. Petersburg

[2] See Nover and Hájek (2004), and Hájek and Nover (2006) for fuller presentations of the payoff and probability tables of both the Pasadena and St. Petersburg games, and further discussion of their expectations.

game. Offhand, this is startling—after all, the St. Petersburg game's payoffs are all at least as great as the Pasadena game's corresponding payoffs, the former are *strictly* greater apart from the payoff for the coin landing Heads immediately ($2 dollars), and —icing on the cake—the former escalate so that they quickly become *much* greater. If you play the Pasadena game and we simultaneously play the St. Petersburg game with the same coin experiment dictating both our outcomes, in every possible state of the world we do at least as well as you, and in some of them we do strictly better than you. Whose shoes would you rather be in? In decision-theoretic parlance, the St. Petersburg game *weakly dominates* the Pasadena game. And offhand, you should prefer the game that does the weak dominating to the one that is weakly dominated; in any case (allowing for indifference or incommensurability between the games) your preference certainly should not go the other way. This vividly brings out a conflict between dominance reasoning and standard decision theory.

Altadena is another agreeable town, a little higher in altitude than Pasadena. The *Altadena game* is just like the Pasadena game, except that each payoff is increased by $1. In every possible state of the world, the Altadena game's payoff is strictly greater than its Pasadena game's counterpart. In decision-theoretic parlance, the Altadena game *strictly dominates* the Pasadena game. Standard decision theory's conflict with dominance reasoning is brought out just as strikingly by another, closely related result of Fine's: consistent with the preference axioms of decision theory, the Altadena game can be valued at any real number whatsoever, *independent of the valuation of the Pasadena game*. In particular, you can value the former *below* the latter, and by any (finite) amount that you like. Offhand, this too is startling.

We think it is too startling. If the preference axioms conflict with dominance reasoning, then too bad for those axioms, we say. Indeed, we may localize the conflict: as Fine shows, it is specifically the *Archimedean* axiom that conflicts with dominance reasoning. Too bad for the Archimedean axiom, we say—and argue in section 3.

This debate could easily stalemate. Enter Kenny Easwaran (MS). His major contribution, which we discuss in section 4, is to generalize the notion of a gamble's expectation to its 'weak expectation', an homage to its role in the weak law of large numbers. He shows that the Pasadena game's weak expectation is well defined, and he argues that it may well be a good guide to choiceworthiness. This provides some vindication for our dominance-dominated intuitions, since the Altadena game's weak expectation is exactly a dollar greater than that of the Pasadena game, and the St. Petersburg game's weak expectation is apparently infinite and thus greater than both. Moreover, there is no freedom in where these games are valued relative to gambles with defined utilities, an issue that we take up in section 5.

We conclude in section 6 with some suggestions of avenues for future research. In particular, we call for a new set of preference axioms for decision theory that will yield both dominance reasoning and a representation theorem for weak expectations.

2. Fine

Fine gives an exemplary primer on linear utility theory; we recap the highlights. Begin with a set of objects of value, which we will call *prizes*—monetary payoffs, goods of various kinds, or what have you. Most of our discussion will concern dollar amounts. Note that some of these things may be undesirable to you—*booby* prizes, monetary *losses*, *bads* of various kinds. Now consider various 'discrete' gambles over these prizes, with known probabilities for each, and with only countably many possible outcomes. Fine initially considers the set $\mathcal{G}$ of all gambles with only finitely many possible (that is, having positive probability) outcomes, the 'simple' gambles. We assume that you have rational preferences over these gambles: let $G_1 \succ G_2$ denote your preferring to , and let G_1 G_2 [NOTE TO READER: THE MIDDLE SYMBOL SHOULD LOOK LIKE A CURLY '≥' AS IN FINE'S, BUT DOESN'T DISPLAY CORRECTLY ON SOME COMPUTERS.] denote your regarding as at least as desirable as (i.e. you either prefer to , or you are indifferent between them). Our goal is to 'represent' these preferences by assigning to each gamble in Ģ a real number $u(G)$, the *utility* of , such that G_1 G_2 if and only if $u(G_1)$ $u(G_2)$. So far this is analogous to beginning with a set of possible weather conditions and a comparative notion of 'hotter than' defined on them, which can then be represented with a real-valued temperature function.

But we want more from our preferences: we would also like to constrain them in such a way that the utility function for various complex gambles supervenes on the values of the function for simpler gambles. (This has no analogue in the temperature case.) To this end, we impose some axioms on preferences that are also meant to be rationally compelling in their own right. Two are structural 'richness' axioms: the *axiom of totally ordered preference* and the *axiom of mixture sets of gambles* guarantee that preferences are suitably defined over all mixtures (convex combinations) of more basic gambles. Since gambles can themselves be prizes of more complex gambles, it is natural to consider what happens when we replace one such prize with another that is at least as desirable. The *independence* axiom says roughly that if is at least as desirable as , then a complex gamble with as one of its prizes is at least as desirable as the complex gamble that replaces with but that is otherwise identical. The *Archimedean axiom* says that if ħ ħ , then there is a non-trivial mixture of and that is at least as desirable as , and another non-trivial mixture of   and   that is no more desirable than   . This axiom will prove to be important in our later discussion.

With these axioms in place, Fine then cites a *representation theorem:* any set of preferences that satisfy the axioms can be represented by a finitely additive linear utility function *u*. This tells us that the utility of a mixture of two gambles is the corresponding mixture of their utilities taken separately. More precisely, for any pair of gambles   and   , and for any real number 0 < λ < 1, we may form the mixture gamble $\lambda G_1 + (1-\lambda)G_2$ whose utility equals $\lambda u(G_1) + (1-\lambda)u(G_2)$.

Given a utility function   defined on □, following Fine we may define a *value function*   on the set of basic prizes, namely the restriction of   to the sure-thing gambles (whose prizes are guaranteed). □'s finite additivity assures us that the value it assigns to any simple gamble   that yields prizes $< c_1, c_2, \quad , c_n >$ with probability distribution $< G(c_1), G(c_2), \quad , G(c_n) >$ is the *expectation* of the   -value of these prizes with respect to this distribution:

$$u_1(G) = \sum_{i=1}^{n} v(c_i)G(c_i) = E_G v$$

In the case of simple gambles, the detour through □ is unnecessary—this is just a longwinded way of saying that the □-value of a G gamble is the expectation of the □-values of its various prizes. But when we come to more complicated gambles with infinitely many possible prizes, we will not want to take this identification for granted. In particular, when the gambles are badly behaved, as the Pasadena, Altadena, and St. Petersburg games are, we will need to keep the distinction between and □ clear.

So let us turn to more complicated gambles. Fine proceeds in two stages: first, he extends the preference relation to well-behaved discrete gambles (which may have infinitely many possible prizes); then, he extends it again to all mixtures of these gambles, the Pasadena game, the Altadena game, and the St. Petersburg game. The test for well-behavedness of a gamble is whether the expectation of the □-value of its prizes according to the gamble's probability distribution converges absolutely. Fine's $\mathcal{G}_2$ is the mixture set of discrete gambles on prizes from $\mathcal{G}_1$ passing this test. Intuitively, $\mathcal{G}_2$ is the set of all gambles (finite or infinite) for which the expectation is a sum that allows no monkey business: it converges, and its value does not depend on the order of summation (the way that conditionally convergent series do). The St. Petersburg game involves one kind of monkey business: its expectation series diverges. The Pasadena and Altadena games involve another kind: their expectation series are sensitive to their order of summation (being conditionally convergent).

Fine's final extension is to $\mathcal{G}_3$, the smallest mixture set containing the Pasadena game, the Altadena game, the St. Petersburg game, and all gambles in $\mathcal{G}_2$. This is where Fine makes his most significant contribution. He shows that consistent with the preference axioms, there exist linear utility functions on $\mathcal{G}_3$, and so in particular one can assign mutually consistent valuations to games in $\mathcal{G}_2$ and to the Pasadena, Altadena, and St. Petersburg games. That's the good news. *Paradox lost?* No, for there is also bad news, or so we insist. The valuations of the three games are entirely *arbitrary* and *independent* of each other. Firstly, the arbitrariness: the Pasadena game can be placed wherever you like in comparison with those gambles whose linear utilities are well-defined as expected values (the well-behaved gambles); so too the

Altadena game; and so too the St. Petersburg game. Secondly, the independence: fixing the valuation of one of these games puts no constraint whatsoever on the valuations of the other two, although valuations of their mixtures are determined linearly. In particular, as far as the preference axioms are concerned, you can consistently value the Pasadena game *above* the Altadena game, and this in turn *above* the St. Petersburg game. Offhand, this seems exactly back-to-front.

The back-to-frontness intuition stems largely from dominance reasoning, which Fine addresses in his penultimate section. As we have seen, two applications of dominance reasoning yield that both the St. Petersburg game and the Altadena game are preferable to the Pasadena game. It is plausible, though not entailed by dominance reasoning, that the St. Petersburg game is also preferable to the Altadena game—not entailed, because there is a state of the world (i.e. the coin lands Heads immediately) in which the former's payoff ($2) is less than the latter's ($3). But even if this last judgment is negotiable, we maintain that it is *not* negotiable that the Pasadena game should be ranked strictly *below* the other two games, as dominance reasoning demands.

How, then, can expected utility theory allow valuations of these games that dominance reasoning forbids? The reason is that dominance reasoning is not implied by the standard axioms of the theory. In fact, one of the axioms is *inconsistent* with dominance reasoning. As Fine points out, the Archimedean axiom implies that all utilities are real-valued. But as we will rehearse in the next section, dominance reasoning requires us to assign infinite utility—*non*-real-valued utility—to the St. Petersburg game. Since Fine holds on to the Archimedean axiom at the expense of dominance reasoning, he can accept the rankings of the three anomalous games that we have called back-to-front. Since we hold on to dominance reasoning at the expense of the Archimedean axiom, we cannot. We turn to a defence of our position.

3. Dominance vs. the Archimedean axiom

In the blue corner, we have dominance reasoning; in the red corner, utility theory's Archimedean axiom. Fine shows that there is a conflict between these two putative rationality constraints (assuming the other preference axioms). We gave part of the argument in our 2006: By dominance, the St. Petersburg game is preferable to each of its finite truncations. For example, it is preferable to the game that is called off if Heads is not reached by the tenth toss; to the game that is called off if Heads is not reached by the eleventh toss; and so on. So it is preferable to something with expectation 10; it is preferable to something with expectation 11; and so on. So its expectation is infinite. But as is well known, and as Fine points out, decision theory's Archimedean axiom is violated if there is a gamble of infinite expectation. For then letting be this gamble, be a sure payoff of $1, and the status quo, we have

ℏ ℏ ; but any non-trivial mixture of and is more desirable than (indeed, infinitely more so).

Either dominance reasoning or the Archimedean axiom has to give—which is it to be? Colyvan endorses dominance reasoning as compelling even when expected utility theory runs aground, and we agree. We also agree with Easwaran that the fact that the axioms forbid valuing the St. Petersburg game infinitely is a reason to question them—in particular, the Archimedean axiom. But before making our final judgment, let us look briefly at some arguments for and against both it and dominance reasoning.

Archimedes famously calculated how many grains of sand would be required to fill the known universe. His answer was finite (8×10^{63} in modern notation). If he were right, the sizes of the universe and a grain of sand would be *comparable* in the sense of having a finite, real-valued ratio. The real numbers display a similar Archimedean comparability: if *a* and *b* are two positive real numbers, then there is a positive integer n such that $n.a > b$. And a similar thought undergirds the Archimedan axiom; its main purpose is to guarantee that utilities are real-valued. As such, it renders utility theory susceptible to the methods of real analysis—a rich reward.

But let us not mistake an *idealization* in our *theory of* rationality for an *ideal* of rationality. It is one thing to appeal to the Archimedean axiom in our theorizing *about* rational agents, quite another to demand obedience to that axiom *of* such agents. Compare: it might help our theorizing about gases to treat gas molecules as point-sized, but let us not project that idealization back on to the gases themselves as if it were an insight into their nature. The Archimedean axiom makes our theory of rationality tractable; that does not make it an insight into the nature of rationality itself.

Perhaps the axiom is nevertheless compelling in its own right? We think not. Its status is surely not the same as the independence axiom, any violation of which seems to involve a kind of doublethink, and is thus irrational. At least some violations of the Archimedean axiom are not obviously irrational—for example, preferring $10 to $1, and $1 to excruciating death, without regarding any gamble between $10 and excruciating death as worth at least $1. Or again, recall the Beatles' song 'Can't Buy Me Love'. This can be interpreted as a plea for a violation of the Archimedean axiom (although we doubt that Lennon and McCartney explicitly intended this interpretation![3]). Arguably, love is just not the sort of good that can be equalled or surpassed by piling on a sufficient amount of money. Furthermore, it might be the sort of good that is worth any risk to obtain, without any offence to rationality. So it seems that a rational agent could prefer a life of love to $1, and $1 to $0, while preferring any chance at love to $1. We are not claiming that this preference structure is rationally required; it may not even actually belong to any human (Lennon and

[3] The lyrics of the song do give some useful information about the relevant preference ordering: 'I don't care too much for money—Money can't buy me love'.

McCartney included). We merely insist that it is rationally *permissible*—contra the Archimedean axiom.

Indeed, it may even be *rationally required* to violate the Archimedean axiom. A Kantian might say that it is rationally required to recognize certain categorical imperatives—for example, a duty not to commit murder. But then the preference structure:

$\$10 \succ \$1 \succ$ murder

apparently violates the Archimedean axiom: any gamble between \$10 and murder is so contaminated by the latter that it is dragged below \$1. More generally, rationality might require our preferences to be lexically ordered, defying an Archimedean representation.

Finally, the conflict between the Archimedean axiom and dominance is itself an argument against that axiom. We have seen how the St. Petersburg game provides another counterexample to the Archimedean axiom, provided we allow ourselves infinitely many appeals to dominance reasoning. Interestingly, it is a counterexample in which goods *of the same kind*—lotteries with cash prizes—are compared with each other, unlike the other cases we have considered.

So the Archimedean axiom is not sacrosanct. At best, it earns its keep as part of a mostly successful theory, and it basks in a reflected glory from that success. When that theory breaks down, as we believe it does for the Pasadena game and its comparison with the Altadena game, that glory is mitigated.

How about dominance reasoning? It is important to stress that the usual concern that one might have with dominance reasoning, when states are probabilistically dependent on acts, does not apply here. Obviously, whether or not you take the Pasadena game or the Altadena game has no impact on how the coin lands. Then there are alleged arguments against dominance from certain multi-player games (e.g. the prisoner's dilemma), or perhaps certain sequential choice situations. But these introduce extra complications that are not germane to the Pasadena/Altadena game comparison. Closer to home, but not close enough, is the two envelope problem, which really involves a violation of so-called *conglomerability* rather than dominance. (See Arntzenius and McCarthy 1997, Arntzenius, Elga and Hawthorne 2004.) We do not know of any counterexample to synchronic, one-person, act/state-independent choice obeying dominance—yet that is what is at issue in the non-negotiable comparisons that we have adduced.

Indeed, it is hard to imagine what a counterexample to dominance would look like. Of course, it is possible for you to gain no extra utility from the increase of some good—that just means you have reached a saturation point for that good. But it is *not* possible for you to gain no extra utility by the addition of *further utility.* That would be tantamount to you not obeying your own utility representation!

To be sure, valuing individual payoffs is one thing, valuing gambles among them is another. Perhaps somehow unanimously strict preferences state-by-state could fail to 'agglomerate' to an overall preference for a strictly dominant act? But consider what this would amount to for the Altadena game vs. Pasadena game comparison. You play the Pasadena game, and end up in some state of the world. We then offer you the choice to upgrade, at no charge, to an outcome that is one utile better. By definition of a utile, you prefer the upgrade, and so are rationally required to choose it, thus ending up where you would have been if you had played the Altadena instead of Pasadena game. But this is true for every possible outcome. So given the choice up front, it seems that you are rationally required to choose the Altadena game. Put another way, if you are offered the choice between the Altadena and Pasadena games and choose the Pasadena game, then you will wish afterwards that you had chosen the Altadena game, no matter what the outcome; whereas if you choose the Altadena game, then whatever the outcome, you will be happy with your choice. So you are presumably required to choose the Altadena game. A similar argument supports weak dominance reasoning in a choice between the St. Petersburg and Pasadena games—not quite as decisively, perhaps, but decisively enough. If you choose the Pasadena game, then you may be sorry that you did, and whatever the outcome, you will not be genuinely glad that you did; whereas if you choose the St. Petersburg game, then whatever the outcome, you will not be sorry that you did, and you may be genuinely glad that you did (this with probability ½).

And so it goes with dominance reasoning more generally. In sum, even if there is some intuition in favour of the Archimedean axiom, we insist that the intuition in favour of dominance is stronger. (And even if the intuitions were of equal strength, that still may not justify tilting exclusively for decision theory in its current form. It may only justify a pluralism of such theories, one privileging the Archimedean axiom, another privileging dominance—see Colyvan 2006 and the final section of Hájek and Nover 2006 for further discussion of pluralism.)

We have appealed to dominance reasoning in support of our preference of the St. Petersburg game over the Pasadena game and of the Pasadena game over the Altadena game, and we could rest our case there. But to lodge our protest against standard decision theory, it suffices to appeal to something weaker still—and again, we draw our inspiration from Fine, although presumably not in a way that he would have intended.

Now consider the *Negative St. Petersburg game*, in which *punishments* grow exponentially, just as the St. Petersburg game's rewards grow—that is, we switch the signs of the payoffs in the St. Petersburg game. Which game would you rather play? Intuition tells us, yells at us, indeed bellows at us, that the St. Petersburg game is preferable. Here we may appeal not merely to dominance, but to a kind of *superdominance* (cf. McClennen 1994, Hájek 2004). Say that gamble G_1 *strictly superdominates* G_2 just in case each payoff of G_1 is strictly greater than every payoff of G_2. In particular, the *worst* possible result of G_1 is strictly better than the *best* possible result of G_1, if both of these are defined. Then in the strongest sense you can't go wrong by choosing G_1 over G_2. Clearly, the St. Petersburg game strictly superdominates the Negative St. Petersburg game: the smallest possible payoff of the former ($2) is greater than the largest possible payoff of the latter ($–2). So we insist that the former is preferable to the latter. If we have not convinced you above of the cogency of dominance reasoning, so be it; but we dare you to question with a straight face the cogency of strict-superdominance reasoning, such as we have just deployed.

And yet applying Fine's methods one more time, we may show that consistent with the preference axioms of standard decision theory, the Negative St. Petersburg game can be valued at any real number whatsoever. Putting this together with his result about the St. Petersburg game, we have the disastrous consequence that by the lights of standard decision theory, you are free to value the Negative St. Petersburg game *above* the St. Petersburg game[4]. That is surely a reductio of orthodox decision theory.

Thus, a rethinking of that theory is in order.

4. Easwaran

Gambles can be regarded as random variables. For example, the Pasadena game can be regarded as a random variable whose value is $(-1)^{n-1}2^n/n$ with probability .

Gambles can also be repeated, generating a sequence of random variables. Let X_i (i = 1, 2, …) be a sequence of independent, identically distributed random variables.

[4] We thank Aidan Lyon for pointing this out to us.

Assume for the moment that the expectation EX_i is defined; you will not be surprised to learn that soon we will suspend this assumption. Let $S_n = X_1 + X_2 + \cdots + X_n$ S_n =

X_1 + X_2 + ... + X_n.. The random variable $\frac{S_n}{n}$ will prove to be important in our discussion; think of it as the averaged payoffpayoff of n plays of the same gamble. If

n is large, then we call $\frac{S_n}{n}$ a *long run average*. It plays centre stage in the celebrated *laws of large numbers*, formal expressions of the long run stabilization of certain random processes that the folk call the 'law of averages'. Easwaran gives precise statements of the so-called *strong* and *weak* laws of large numbers. In slogan form: the strong law of large numbers is a happy claim about the *probability 1 of convergence* of the long run average to the expectation, while the weak law is a happy claim about the *convergence of probability to 1* of the long run average's discrepancy from the expectation being small. The former is called *almost sure convergence,* the latter *convergence in probability*.

For those gambles that have an expectation, there are three roles that the expectation can play:

1. It is a measure of choiceworthiness of the gamble.

2. It is the quantity to which $\frac{S_n}{n}$ converges almost surely, as per the strong law of large numbers.

3. It is the quantity to which $\frac{S_n}{n}$ converges in probability, as per the weak law of large numbers.

For typical, well-behaved gambles, roles 1, 2 and 3 coincide. But in the Pasadena game, nothing can play role 2. Yet it is still to be hoped that something plays role 1.

Easwaran argues that this hope might be satisfied—since something *does* play role 3 in the game. Specifically, he proves that $\frac{S_n}{n}$ converges in probability to *ln* 2—he calls this the *weak expectation* of the game.

You may recognize this number: it is the sum of the Pasadena game's expectation series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots$$

taken in that order. It appeared in our initial presentation of the Pasadena game as the first candidate for the value of the game; the trouble was that soon afterwards it had to compete with other candidates. In numerous discussions of the Pasadena game with us, various critics have insisted that its expectation is *ln* 2. For after all, they repeatedly say, the coin toss outcomes come in a particular order, favouring the corresponding order of summation of the expectation series.

We have remained unconvinced. Firstly, in insisting that *ln* 2 is the expectation of the Pasadena game, these critics are apparently not just taking us on, but also the orthodoxy among probabilists and decision theorists, according to which the expectation of a game is undefined if it is not absolutely convergent. Secondly, anticipating exactly this insistence, in our original paper we imagined the mechanism that determines the payoffs being placed inside a black box. You are ignorant of its workings; all you know is the utility and probability profile that it produces. Now there really seems to be no privileged ordering. Compare: when deciding whether to bring an umbrella to work tomorrow, there is no privileged ordering of the states {rain, not rain}. Permuting the columns in a decision matrix should not change the decision problem; yet this seems to correspond to permuting the order of terms in the expectation series. Finally, if these critics are right, then the Pasadena game is far more important than we ever thought. It would apparently overthrow nearly a century of decision-theoretic orthodoxy. All these years we have been told that the utility and probability profiles suffice for settling issues of choiceworthiness; now, apparently, we learn that we need knowledge of a third thing, the mechanism that produces the outcomes, so that we can privilege one ordering! (We wonder what the special relativistic effects might be on that ordering as seen from various frames of reference —for example, if there is one according to which all the tosses of the coin are simultaneous—but let that pass.) We would be delighted if the game taught such a profound, revisionary lesson, and we would thank these critics for pointing this out to us. But we are happy to settle for less.

Now Easwaran has given a far more subtle argument for valuing the Pasadena game at *ln* 2. In our 2006 paper we lamented the fact that current expected utility theory lacks the resources to put a value on the game, and we lodged a plea for better technology that would do so. One way of understanding his important contribution is that such technology was under our noses all along in the familiar notion of convergence in probability, made explicit in his notion of 'weak expectation'.

So let us consider the case that he makes for weak expectation being a guide to choiceworthiness. His main argument is this:

> consider the case of an agent playing a game a fixed very large number of times. If she plays repeatedly at a price that is slightly higher than the weak expectation, then she has a very high probability of ending up behind. If she plays repeatedly at a price that is less than the weak expectation then she has a very high probability of ending up ahead. Playing at a price exactly equal to the weak expectation for each play doesn't guarantee anything about the probability of eventually being likely to be ahead or behind. Thus, it seems that an agent should not be willing to pay more than the weak expectation, and should be willing to pay any amount less than it. (Easwaran MS, 3)

To be sure, it would be nice to find a price for the Pasadena game such that, with probability 1, after infinitely many trials of paying that price you would exactly break even—that would be a good candidate for the value of the game. Alas, there is no such price. Failing that, it seems prima facie reasonable to look for a price that is neither too high nor too low in the sense of eventually almost guaranteeing an advantage to either the agent or the house. The weak expectation is that price, and it is thus another candidate for the value.

How good a candidate is it? One cause for concern—or is it a cause for wonder?—is that Easwaran's argument, which arrives at the 'naïve' *ln* 2 valuation, goes through just as well even for the 'black box' presentation of the Pasadena game. This seems almost too good to be true. Remember that the utilities and probabilities could be given in any order, and in the absence of any knowledge of the mechanism that generates them, no order seems privileged. To be sure, this much can be said in favour of the ordering of the expectation series given above: it uniquely corresponds to the increasing order of absolute values of the game's utilities, and the decreasing order of its probabilities, and the decreasing order of its absolute values of utilities multiplied by probabilities (which is just to say that terms in the series successively get closer to 0). These facts are invariant across any mechanism that might generate the utility and probability profile. Still, we wonder what bearing these facts might have on choiceworthiness.

All that said, Easwaran has given us an ingenious proposal that may genuinely push forward our understanding of such anomalous games. Now, if weak expectations are good guides to choiceworthiness, then there is at least one respect in which the Pasadena game is *less* anomalous than the St. Petersburg game. After all, if we can make sense of a weak expectation at all for the latter, it is infinite, only confirming the original 'paradox' that you should gladly pay any finite amount to play it. But *ln* 2 is a perfectly anodyne value to place on a game; you could pay for it with your lunch money, and still get plenty of change. And the Altadena game's value is placed at *ln 2 + 1*, vindicating our intuition that it should be exactly one utile (dollar) above the Pasadena game, and that it should be below the St. Petersburg game.

5. 5 positions on ladder positions

Fine and Easwaran endorse positions at opposite ends of a spectrum. Fine writes: 'There is no meaningful finite price for the Pasadena and Altadena games when this price is to be assessed from an unlimited number of repetitions of such a game' (7). His argument is that the long run average winnings in these games will with probability 1 exceed any pre-assigned finite number infinitely often, and with probability 1 fall below any pre-assigned finite number infinitely often. Easwaran proves this and notes that this fact thwarts the possibility of the games having (strong) expectations, for almost sure convergence must fail. But he goes on to argue in favour of valuing these games uniquely at their weak expectations.

Fine interprets his own result about the unlimited freedom in valuing the Pasadena game as concerning 'a single presentation' of it (REF 7). Nonetheless, one might take this result to speak to repeated plays of it as well: by the lights of expected utility theory, *any* finite price for the game is meaningful, even in repeated plays. Conversely, Easwaran's argument for valuing the Pasadena game at its weak expectations is premised on its being played a 'very large number of times' (REF 3). Nonetheless, one might take the argument to speak to single presentations of it as well: by symmetry, the same valuation will be given of each presentation in a series of repeated plays, but *each* presentation is intrinsically the same as a *single* presentation.

As we have seen, Fine's and Easwaran's arguments apply equally to the Altadena game. Fine allows us to value it however we like, independently of the Pasadena game; Easwaran would have us value it uniquely at *ln* 2 + 1, exactly one utile (dollar) above the Pasadena game. But there are also various other positions, as we will now see.

Picture a fireman's ladder that comes in two parts. Start with a fixed ladder; then overlay it with a sliding second ladder that can be locked into any position against the fixed ladder. Now think of an infinite 'ladder' of gambles of known utilities; for concreteness, let it be the following:

$2 for sure

$1 for sure

0 for sure

–$1 for sure

–$2 for sure

The Pasadena game generates its own infinite sequence of gambles, which we will call *the Pasadena sequence*. It is tempting also to think of it as a ladder, and soon we will argue that it is:

Pasadena game + $2

Pasadena game + $1 = Altadena game

Pasadena game

Pasadena game – $1

Pasadena game – $2

Now consider the problem of superimposing this sequence on the first ladder. We can characterize five positions as to how this might be done:

5.1 The Pasadena sequence cannot be superimposed on the ladder of defined utilities

The first position insists that it cannot be done—any putative superimposition is mistaken. (Compare: any value given to 1/0 is mistaken.) There are two ways this position might be spelled out:

(i) The Pasadena game itself is incoherent

We argued in our 2004 and 2006 papers for the coherence of the Pasadena game. Now we can also approvingly cite Fine's assessment of the Pasadena and Altadena games: 'mathematically they are finite-valued, discrete random variables and, as such, are as well-defined as any other such object' (REF 3). The same is true of all their kin in the Pasadena sequence.

(ii) The Pasadena game is coherent, but the choice between it and any gamble of finite known utility is ill-defined

This is Colyvan's (2006) position, which we rebutted at some length in our 2006 paper. We contend that the untenability of (i)—on which Colyvan and we agree—implies the untenability of (ii). If one can in principle play the Pasadena game, then one can in principle be made to choose between playing and not playing it (= the status quo). And at that point, a rational agent must do something, the decision of which gives some guidance as to where to place the Pasadena game on the ladder of defined utilities. Moreover, Colyvan grants us that *some* choices involving the Pasadena game are well defined—for example, by dominance reasoning it should be dispreferred to the Altadena game. But by Fine's axiom of totally ordered preference, or by any similar axiom of 'completeness' or 'connectedness' of preferences, if there is *any* other gamble to which a gamble can be compared, then that gamble can be compared to *every* other gamble. [HARRIS: I'M HAPPY WITH THIS AS IT STANDS, BUT IF YOU WANT TO ADD A BRIEF ARGUMENT, PLEASE GO AHEAD

Colyvan's reason for denying that the Pasadena game can be placed on the ladder of known utilities is that its expected utility is undefined. But as Fine writes, 'it suffices for rationality that preferences between gambles be represented by a linear utility function, but it is not necessary that this linear utility function be of the form of an expected utility' (REF 9). And as we have seen, Fine shows that preferences involving the Pasadena game and its kin *can* be represented by a linear utility function. To be sure, we do not take ourselves to be beholden to linear utility theory—in fact, we have used Fine's results to give what we take to be a reductio of it. But giving up that theory only further undercuts Colyvan's argument, which was premised on it.

5.2 The gambles in the Pasadena sequence can be superimposed anywhere on the ladder of defined utilities, independently of each other

On Fine's view, the Pasadena sequence does not form a ladder at all, but rather a loose set of rungs that can be dropped anywhere you like on the fixed ladder of defined utilities. (Well, almost anywhere—not at positive infinity or negative infinity.) You are free to value the Pasadena game wherever you like, and independently of that, the Altadena game wherever you like, and so on.

We reply: to be sure, this is the verdict of standard utility theory. But we *tollens* where Fine *ponenses*—we find this verdict unacceptable, flouting as it does our intuitions about the seemingly rigid, non-negotiable relationships between the various Pasadena-like games. We might try to retrain our intuitions accordingly, thus finding a reflective equilibrium between them and the theory. Or we might hold those intuitions as sacrosanct for as long as we can, and try to revise decision theory accordingly. We plump for the latter option. We thus want to drive Fine's argument in reverse, appropriating it to cast doubt on standard utility theory. In particular, repeated appeals to dominance reasoning force the ordering of the Pasadena sequence according to the ladder above.

5.3 The Pasadena sequence is ordered, but internal differences in utilities among its members are not fixed

On this view, the Pasadena sequence forms a ladder of sorts, but it is a rubbery, indefinitely stretchable ladder. For example, the Altadena game is better than the Pasadena game, but there is no saying how *much* better. This would be the position we would be left with if we had recourse to dominance reasoning in valuing these games and their relatives, but nothing else. For dominance reasoning forces us to make certain comparative preference judgments, but anything more it leaves underdetermined.

We reply: It is not just that the Altadena game is better than the Pasadena game; it is *a dollar (utile)* better. Consider: you have a choice of either playing the Pasadena game, or playing it and then receiving a further dollar when it is over. Clearly the latter is better, but more than that: it is clearly *one dollar* better. This leads us to the next view.

5.4 The Pasadena sequence forms a rigid ladder, but it can be superimposed anywhere on the ladder of defined utilities

Granting that we have no freedom in the *relative* placement of the members of the Pasadena sequence, that still leaves open whether we have freedom in their *absolute* placement. This view claims that we do—indeed, we have unlimited freedom. This is where the fireman's ladder analogy is most apt. We may 'lock in' the Pasadena game wherever we like on the ladder of defined utilities, but having done so, the values of all the other members of the Pasadena sequence are determined. Or to use another analogy a little closer to home, recall the familiar view that the conditional probability of *A* given *B* is unconstrained by *P*(*A* & *B*) and *P*(*B*) when *P*(*B*) = 0. The equation

$P(A \mid B).P(B) = P(A \,\&\, B)$

becomes

$P(A \mid B).0 = 0,$

so we are free to assign *P*(*A* | *B*) any value we like. But having assigned it, other values are then determined—for example, *P*(¬*A* | *B*).

We reply: This view posits an interesting failure of the supervenience of the values of certain compound gambles on the values and probabilities of their constituents. Absolutely convergent gambles (that live in Fine's G_2) do seem to supervene in this way. If that supervenience breaks down, there is an arbitrariness in how we extend expected utilities to G_3 gambles that makes us uneasy. But we admit that an argument from uneasiness is not much of an argument.

While it seems to us unlikely that *every* placement of the Pasadena sequence, thought of as a rigid ladder, is rationally permissible, we are not opposed to the idea that there may be many rational placements. For example, without the Archimedean axiom, one might employ non-standard extensions of the real numbers, such as the hyperreals, for the values of our utility functions. That may leave the valuation of the Pasadena game partially but not fully constrained. It might, for example, be constrained to be some infinite hyperreal number, without further commitment. Less dramatically, it might be constrained to lie in the monad of a given real number (i.e. lie infinitesimally close to it), but no more.

However, we note that if this supervenience fails, it is unlikely to do so at the boundary between G_2 and G_3. As Fine mentions in passing (REF 25), the standard axioms do not force us to value games in G_2 at their expected value. These axioms constrain only the valuations of simple gambles, and there is no difference as far as the axioms are concerned between games in G_2 and games in G_3. When we value games in G_2 at their expected values, we may be guided by our intuitions, but we are not guided by the standard axioms. In fact, Fine's proof, *mutatis mutandis*, goes through just as well to show that a given gamble in G_2 can be valued arbitrarily compared to gambles in G_1. But for G_2, unlike G_3, our intuition to use the expected value is not only strong but also self-consistent. However, additional axioms are needed to bring rigor to this intuition, and these axioms are likely to say something about G_3 games as well. Put another way, drawing a line between gambles in and out of G_2 is unprincipled. The current axioms care only about membership in G_1, and rectifying that must bear on G_3.

5.5 The Pasadena sequence forms a rigid ladder, and there is a unique place where it can be superimposed on the ladder of defined utilities

The final view is the most committal regarding the valuation of the Pasadena game and its kin. Rationality requires unique valuations of all of them. (Presumably the Pasadena game's position on the ladder of defined utilities is locked in, and the positions of the rest are determined in the obvious way.)

We reply: this gives us a welcome supervenience of the value of G_3 gambles on the values and probabilities of their constituents, and it removes all arbitrariness in valuing them. But we admit that an argument from welcomeness is not much of an argument either.

More cause for optimism here is provided by Easwaran. All the games in the Pasadena sequence have unique weak expectations. To be sure, Easwaran himself seems to be a little ambivalent about the normative force of weak expectations, and his points are well taken. Perhaps they are just a step towards a more complete decision theory. We will pursue this thought further in the final section, but for now leave this discussion on a note of ambivalence of our own.

To summarize this section: we are opposed to all the views countenanced except 5.4 and 5.5. Ideally, 5.5 will eventually hold sway, but we are not yet committed to it.

6. Fine-tuning decision theory

Fine writes: ‘Linear utility theory is about ‘rationally’ reducing the complexity of choices between gambles. By making certain simpler choices we can then delegate to the mathematics to say what our choices between more complex gambles should be’ (REF 15). Complex gambles yield complex expectations. When the gambles become sufficiently complex in certain ways, the usual methods for evaluating choices among them break down. As theorists observing this situation, *we* face a decision problem of our own: we can choose to not worry and be happy, or we can try to develop new methods to handle such gambles. The situation is not unlike that which confronted mathematicians in the 19th century when anomalous functions that were not Riemann integrable were discovered. They could have rested content with their theory of integration. Instead, Lebesgue, Stieltjes and others developed new theories that generalized Riemann integration, yielding its results while also successfully handling the anomalous functions.

The Pasadena and Altadena games are complex gambles that frustrate our expectations, as calculated by our usual decision theory. We start with intuitive rational preference axioms. These lead to a representation theorem, which tells us that expected utility is a sure-fire guide to rational preference for simple gambles. In the Pasadena and Altadena games, expected utility theory breaks down. But Fine and Easwaran show that they are amenable to valuation nonetheless, albeit in very different ways. Fine's approach is more conservative (in a good sense of that word): finding a way, indeed infinitely many ways, of accommodating these gambles within existing linear utility theory. Easwaran deploys the familiar mathematics of the weak law of large numbers, but gives it a novel twist, characterizing a new notion of choiceworthiness, 'weak expectation', that applies even to these gambles.

We suggest two broad directions for further research. One approach

begins by offering a new decision rule, one that may make some use of existing utility functions and other representations. Easwaran suggests maximizing weak expectation, and Fine states, without endorsing, a generalization of dominance he calls 'stochastic dominance'. The more compelling such a rule is, the better. The next step is to look for a new set of preference axioms (each plausible, one hopes), which yield a *new* representation theorem: that preference goes (at least in part) by the new decision rule. This second step is necessary, as otherwise our new rule for choiceworthiness is merely a suggestion based on intuition, rather than a consequence of a new, refined definition of rationality.

To be sure, there may still be recalcitrant games even for such an enriched theory.

For example, consider the *alternating St. Petersburg game,* which alternates rewards with punishments, changing the sign of those St. Petersburg payoffs associated with ending on an even number of tosses the n^{th} payoff is $(-1)^n 2^n$. Its expectation series is 1 - 1 + 1 - 1 + ..., so its expectation is undefined. It is unclear if it even has a weak expectation, so Easwaran's technique may break down. In any event, there are surely games that do lack a weak expectation. And yet presumably we should still prefer their 'Altadena' counterparts that pay an additional dollar in every state. This means that even a new-look decision theory, with axioms that provide a representation by weak expectations, will still be incomplete, much as the Pasadena game shows the current theory to be incomplete.

Dominance reasoning suggests proceeding in the reverse direction—and here *we* suggest a second direction for future research. Begin by listing plausible new axioms couched purely in terms of preferences among gambles, unconcerned at this stage by any representation theorems. (Dominance reasoning might be a good candidate to be on that list.) Assuming this list is consistent, then see what sort of new representation theorems, and associated decision rules, might emerge. Ideally, the resulting decision theory will deliver intuitive results on the various anomalous games that we have considered, and it will guide intuition on other anomalous games where previously it was lacking.

So much for the future; where do things stand today? The streets of Pasadena may be safer than they were—but for the time being, there is still some danger out there.[5]

[5] We thank especially John Broome, David Chalmers, Mark Colyvan, Kenny Easwaran, Jason Grossman, Aidan Lyon, Mike Smithson, and Wlodek Rabinowicz for helpful comments.

Philosophy Program ALAN HÁJEK
Research School of the Social Sciences
Australian National University
Canberra, ACT 0200
Australia
alan.hajek@anu.edu.au

Department of Mathematics HARRIS NOVER
University of Wisconsin - Madison
480 Lincoln Dr
Madison WI 53706-1388
USA
nover@math.wisc.edu

References

Arntzenius, Frank and David McCarthy 1997: 'The Two Envelope Paradox and Infinite Expectations'. *Analysis*, 57(1), pp. 28-34.

Arntzenius, Frank, Adam Elga, and John Hawthorne 2004: 'Bayesianism, Infinite Decisions, and Binding'. *Mind*, 113, pp. 251-83.

Colyvan, Mark 2006: '2006. No Eexpectations'. *Mind*, 115 (July), pp. 695–702.

Easwaran, Kenny 2007: 'Strong and Weak Expectations'. MS.

Fine, Terrence 2007: 'Evaluating the Pasadena, Altadena, and St. Petersburg Gambles'.

Hájek, Alan 2004: "Pascal's Wager". *The Stanford Encyclopedia of Philosophy* (Spring 2004 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2004/entries/pascal-wager/>.

Hájek, Alan and Harris Nover 2006: 'Perplexing Expectations'. *Mind*, 115 (July), pp. 703-20.

McClennen, Edward 1994: 'Finite Decision Theory'. In Jeff Jordan (ed.), *Gambling on God: Essays on Pascal's Wager*, Rowman & Littlefield.

Nover, Harris and Alan Hájek 2004: ‘Vexing Expectations’. *Mind*, 113 (April), pp. 237-49.