Challenges from St. Petersburg-Type Paradoxes to the Normative Status of Expected Utility Theory

Ying Shi
M.Sc. Economics and Philosophy

Abstract. There is an uncomfortable cohabitation between expected utility theory (hereafter EU theory) and paradoxes that threaten its normative stature. This essay demonstrates that the paradoxical nature of the St. Petersburg game is only true in a positive sense. That is to say, the game is counterintuitive because real individuals exhibit behaviour that diverges from rational expectations. The game's descriptive inconsistencies have no bearing on norms established for a rational agent, and pose no challenge to the normative relevancy of EU theory.

In contrast, the inability of EU theory to prescribe a stable expectation for the Pasadena game exposes more enduring normative troubles. Possible resolutions include adopting finite states and bounded utilities, weak expectations, and statistical parameters, although shortcomings limit the applicability of each. The addition of the Altadena game illustrates the possibility of competing decision rules. In light of these observations, I advise a lexicographic approach that prioritizes and reconciles pluralistic rules as a necessary remedy to the normative ails of EU theory.

I. Resolving the St. Petersburg 'paradox'

The St. Petersburg paradox describes a series of fair coin flips that ends when the coin lands on tails. The prize is determined by the formula $\$2^n$, where n is the total number of flips. The number of potential acts for the game is unbounded, because there is a possibility, however small, that the coin continually yields heads. The series of expected payouts is as follows:[24]

$$\begin{aligned} EV &= \frac{1}{2}(\$2) + \frac{1}{4}(\$4) + \frac{1}{8}(\$8) + \frac{1}{16}(\$16) + \frac{1}{32}(\$32) + ... \\ &= \$1 + \$1 + \$1 + \$1 + \$1 + ... \\ &= \$\infty. \end{aligned}$$

This game has an expected value of positive infinity. A risk-neutral rational player enters the game if the entry price is below the expected value. Hence a risk-neutral rational agent willingly pays any finite sum to play the game.

The common perception of the St. Petersburg game as a 'paradox' derives from its counterintuitive conclusion of any finite payment as a legitimate price of entry. I will argue for a more nuanced look at the paradoxical stature of the game, by first giving a brief definition and then formulating two perspectives of paradox I label as positive and normative. While the St. Petersburg game is arguably a positive paradox, it does not admit of normative inconsistencies and hence does not pose a legitimate challenge to the normative status of EU theory.

A basic characterization of paradox refers to a set of acceptable premises and reasoning that leads to an absurd conclusion.[25] Within

[24] An adaptation of the St. Petersburg game formulation uses utiles over money. This variation considers the possibility of decreasing marginal utility and increases the utile payouts for each coin toss to yield the same infinitely high expectation.

[25] There is a degree of flexibility in defining 'absurd.' A set of otherwise reasonable premises can logically lead to a contradictory conclusion. Alternatively

this definition we decompose the concept of paradox into positive and normative views. The positive view examines conclusions from the perspective of actual individuals over rational agents. An otherwise logical result from inference is shown to depart significantly from human behaviour or strike most individuals as counterintuitive. Yet a seemingly absurd result can still be proven true using rigorous logic. In the mid-18th century, Leohard Euler computed the sum of the alternating infinite series $1 - 2 + 3 - 4 + ...$ as $\frac{1}{4}$. A series of integers yielding a non-integer total is paradoxical, and made even more so by the small sum relative to integers increasing in absolute value. Subsequent mathematical proofs confirmed the result. Thus what initially strikes individuals as absurd can be rational nonetheless. Such cases of counterintuitive conclusions arising from sound premises and logic exemplify a *positive* paradox.

In contrast, a *normative* paradox requires the conclusion to be absurd from the perspective of *rational* beings. What these agents ought to do must depart from predictions inferred from the original premises. Empirical inconsistencies are unnecessary and irrelevant for establishing the grounds of a normative paradox. That is to say, people's contradictory behaviour or intuitions have no bearing on what is normatively important.

The St. Petersburg game only has a legitimate claim as positive paradox. There is an observed discrepancy between the wagering behaviour of real individuals and the game's predictions of willingness to bet any finite amount. Hacking estimates that few people believe an amount greater than $25 is a reasonable price of entry. The prediction that any finite amount should be surrendered for an opportunity to play is absurd, thereby rendering the game paradoxical. This absurdity presupposes that individuals believe in the terms of the game, including the possibility of infinite coin flips and un-

the result may be 'absurd' because it runs counter to common opinion. The latter begs the questions, "Absurd to whom?" and "At what stage does the group deem the conclusion acceptable?"

bounded payouts.[26] Their choice of a conspicuously lower valuation than the game's expectation marks a departure from what rational agents will pay, and thus makes the game a positive paradox.

Critically, this type of paradox has no bearing on the normative status of EU theory. EU theory governs choice-making in the context of rational beings. In the case of the St. Petersburg game, the results of EU theory are conditional on assumptions of agents with unbounded time, wealth and capabilities. The set of these assumptions make the price of *any* finite amount a reasonable choice. This normative result is independent of the price that people *actually* pay, and hence permits instances when real individuals bet significantly lower than the game's expectation. In other words, whether a real individual wagers $10 or $15 on the game is not relevant in discussions of what people *ought* to pay under assumptions of full rationality. In order to undermine the normative status of EU theory, we must establish the St. Petersburg game as a normative paradox.

To do so, we must show that the rational agent ought to behave in a way that is inconsistent with the predictions of the St. Petersburg game. Put another way, the agent lacks reason to bet any finite amount to play the game. Consider a rational being with unbounded resources and time playing the game into eternity. The only possibilities for violating the game's normative results argue that an agent ought to: 1) pay an infinite amount to play the game, 2) refuse paying to participate or 3) choose a restricted set of finite amounts instead of any finite amount. The first option is irrational because the agent will surely pay less than the expected value to participate. So she will not bet as high as an infinite sum. The second option is equally unfeasible because a rational being seeking to maximize utility will definitely take a bet that yields an expected value of positive infinity. This leaves only the third option. If this is true

[26] I admit of a possibility that the individual judges that she is NOT actually offered the St. Petersburg game. Hence her low entry price reflects the expected value of the game she believes is being presented. As long as her price matches the game's expectation, the individual is rational and this is no longer paradoxical.

then there must be explanations for why select wagers are chosen over others. Yet there is no reason justifying the conscious choice of only a subset of entry prices. An ad hoc decision process is also not permissible. This exhausts all possibilities of establishing the game as a normative paradox. Despite its positive inconsistencies, the St. Petersburg game is normatively logical. As such it cannot pose a threat to the normative status of expected utility theory.

II. Indecision plagues the Pasadena paradox

The resolution of the St. Petersburg game leads us to another supposed paradox. The Pasadena game, as proposed by Nover and Hájek (2004), similarly requires tossing a fair coin until it lands tails for the first time. The prize alternates according to the formula $(-1)n-1*2n/n. The resulting alternating harmonic series converges to an expected value of $\$\ln 2$, as follows:[27]

$$\begin{aligned} EV &= \frac{1}{2} * \$2^1 - \frac{1}{4} * \frac{\$2^2}{2} + \frac{1}{8} * \frac{\$2^3}{3} - ... \\ &= \$1 - \$0.50 + \$0.\overline{33} - ... \\ &= \$\ln 2 \end{aligned}$$

The initial perception of a straightforward game frays when we begin reordering the payouts. Consider a different sequence of coin tosses comprised of an infinite run of positive prizes followed by small negative prizes.[28] The expected value of the game thus specified is positive infinity. The properties of infinity places the Pasadena game on equal footing with the St. Petersburg based on expected value.

[27] An adaptation of the Pasadena game formulation uses utiles over money. This variation considers the possibility of decreasing marginal utility and increases the utile payouts for each coin toss to yield the same range of expectations.

[28] Possible sequence as presented by Nover and Hájek: EV = $1 + (\frac{1}{3} + \frac{1}{5} + \frac{1}{7} ... + \frac{1}{23} - \frac{1}{2}) + (\frac{1}{25} + \frac{1}{27} + ... - \frac{1}{4}) + ...$

Yet this would be too simplistic a view, for a variant ordering of the Pasadena game yields a very different result. Suppose that instead of the infinite run of payouts described above, the ordering changes so that there is a sequence of negative prizes followed by a small positive prize. The series diverges to negative infinity. The plurality of outcomes here lead to puzzling indecision for the rational agent âĂŞ is participation in the game superior, inferior, or indifferent to the status quo of not playing?

Nover and Hájek (2004) note that expected utility theory is uncomfortably silent on the proper course of action for a rational individual deciding whether to participate in the Pasadena game. If we accept that EU theory governs choice-making for rational agents, then what logically follows from its silence is a failure on normative grounds. EU theory is unable to specify what value these agents *ought* to assign to the Pasadena game. There is no appropriate singular course of action when the game has an infinite number of possible solutions and no means of prioritizing these solutions.[29] A rational agent is left without guidance, and thus the Pasadena game becomes a normative paradox. The legitimacy of EU theory as a normative decision rule hinges on whether there is a way out of the Pasadena paradox and others with similar results.

III. Two (or more) solutions and rebuttals

Restrict decision theory to finite spaces. The paradoxical nature of the St. Petersburg and Pasadena games arises primarily from the shared characteristic of being infinite decision problems. Both games deal with unbounded values as well as an infinite ordering of payouts to generate these values. The normally sound mathematical stipulations of EU theory encounter paradoxes when they aim to value the worthiness of these games or conduct comparisons. One

[29]The game is particularly paradoxical because altering the payout ordering should not change the game itself, provided that no additional terms are added or subtracted.

suggested mode of resolution is to do away with infinite states spaces and unbounded payoffs altogether.

Nover and Hájek (2004, 2006) compellingly argue against this possibility via two strands of reasoning. The first objection is a conceptual one, in which all decision problems are limited to finite states as a matter of *conceptual necessity*. Yet restricting decision theory to a bounded theoretical space is problematic. We incorporate notions of infinity in its many applications, from the field of physics to general statistical decision theory which often evokes the concept of infinity through the law of large numbers and calculations of limits. If we acted to confine decision theory to finite state spaces, then how long before we did the same for probability theory and related mathematical axioms? The arms of infinity are long, and it is no easy task to rid decision theory of this foundational concept. Nor is it necessarily a wise move to do so.

The other key objection to the coherency of decision theory deals with *realistic considerations*. Commentators point to the necessity of finite states in decision problems because human agents are incapable of generating infinite spaces. Specifically, individuals do not live long enough, conduct experiments efficiently enough, or possess adequate resources to enable the execution of an infinite decision problem. Two arguments highlight the weakness of this position.

First, EU theory must prioritize rational norms and logically attractive decisions over facts. Consider the possibility of employing the martingale betting strategy in a casino. This class of strategies stipulate that a gambler should double his bets after every loss, so that a win recoups all previous losses in addition to gaining a profit equal to the size of the stake. The fact that few people actually use the martingale is explainable by the finite wealth of gamblers. An individual often bankrupts herself before the strategy is successfully executed. Yet the empirical uselessness of the martingale does not undermine its position as a rationally sound strategy. The strategy's normative status depends on the expected positive payouts from the martingale, but should not be conditional on limited abilities of human beings.

Secondly, it is productive to question the rationale behind bounded human utility. Who is to say that human behaviour necessarily functions within limits, that there must exist upper bounds to happiness or goodness? What construction of finite utility can be perceived as maximally good? It is possible to imagine an individual with set preferences over an infinite set of consequences. Hence, the stipulation of bounded utility seems both arbitrary and convenient, but not rationally defensible.

The method of weak expectations. Thus far we established the volatility of any single attempt to value the Pasadena gamble, because the result is conditional upon ordering. One possible approach of augmenting stability is by examining the case of repeated trials. Specifically, what happens to the distribution of expected values when the number of plays approaches infinity?

Easwaran (2008) was the first to systematically apply the Law of Large Numbers to the case of the Pasadena paradox. In doing so he distinguishes between the Strong and Weak Laws of Large Numbers. The former states that the sample mean *converges almost-surely* to the expected value while the latter describes a *convergence in probability* to the expected value. Let us define a sequence of independent Pasadena distributions as X_i. S_n represents the sum of n plays of the Pasadena game such that $S_n = X_1 + X_2 + ... + X_n$. Formally, the Strong Law says that for any ε,

$$Prob(lim_{x\to\infty}|\frac{S_n}{n - EX}| < \varepsilon) = 1$$

Intuitively, a long sequence of plays will eventually lead to the expected value of EX with probability 1. The agent's ability to decide when to continue playing or quit guarantees that she can attain any EX within the range of expected values for the Pasadena game.[30] In contrast, the Weak Law switches the limit and probability operators

[30] $EX \in (-\infty, \infty)$

to yield:

$$lim_{x\to\infty} Prob(|\frac{S_n}{n - EX}| < \varepsilon) = 1$$

The critical difference is that the number of plays is fixed in advance. The resulting probability of being arbitrarily close to EX given an infinite number of plays is 1. Easwaran states that in this scenario, the average payout per play of the Pasadena game after playing an arbitrarily high number of times is ln2. He defines this method of valuation as 'weak expectations,' relative to the 'strong expectations' method associated with the Strong Law of Large Numbers.

The choice between two approaches to assessing the Pasadena game begs the question, "Which, if any of the options, provides a rational norm for valuing a single play of the game?" Easwaran argues for weak expectations on the grounds that what serves as a fair guide to valuation in infinite repeated plays should also inform valuation for an individual trial. If the weak expectation rule is not adopted, or the number of plays not fixed in advance so that an agent elects to value based on strong expectations, the value chosen can seem ad hoc. One can make a stronger case for the weak expectation rule if the aim is to find a valuation methodology with the most normative weight.

Multiple arguments exist to undermine the applicability of weak expectations. One line of reasoning focuses on the assumption of an arbitrarily large fixed number of plays. Convergence to the weak expectation of \$$ln2$ occurs with almost certain probability for substantial repeated plays of the Pasadena game. An expected value or utility-maximizing agent would seek a payout greater than \$$ln2$ by choosing to stop the game when the payout is between \$$ln2$ and positive infinity. Such an example argues for the rationality of adopting strong expectations and non-fixed number of plays over weak expectations. Yet strong expectation is not a way out of the Pasadena paradox, because it leaves open the possibility of an infinite number of expected values. Strong expectation restores the game's normative troubles when it is unable to prescribe a unique expectation for the game. As such, weak expectation remains the better means of resolving the normative Pasadena paradox.

A more severe challenge to weak expectations arises not within the Pasadena game itself, but from other classes of paradoxes. Alexander (2010) examines at length the Cauchy distribution, whose expectation of repeated plays is exactly the same as the expected value of an individual trial. The application of weak expectation fails because the Cauchy distribution lacks a defined expected value for any single play, so calculating the distribution of average payouts becomes fruitless.

Use of statistical properties. In the absence of a fixed expected value, one can resort to the use of certain statistical attributes to assign value to a single game. For instance, an individual game can be valued at the median or mode of the distribution. This begets several concerns. The first focuses on the ad hoc nature of the selection of statistical parameters. What justifies the choice of median over mode or vice versa? What facts about the bias or variance of the distribution must be known before making such a choice? Secondly, the skewness of the probability distribution can undermine the feasibility of parameters. While the median coincides with the expected value in a Gaussian normal distribution, it can depart significantly from the mean when the distribution is highly skewed. Hence the decision to use median and mode must take into consideration more information on the distribution itself. Such information can also include whether a distribution is unimodal or multimodal, as well as the size and thickness of its tail.

IV. Scale of normative challenges and a pluralistic path forward

The full scope of EU theory's normative troubles does not end with the Pasadena paradox, Alternating St. Petersburg game, or Cauchy distributions. Another commonly invoked paradox is the Altadena game, in which every Pasadena payout increases by $1. It faces similar troubles to the Pasadena game at the onset. Rearranging the payout order yields an infinitely large range of expectations. Hence EU theory cannot define a stable single-game value for a rational agent. At this point, two solution paths emerge, beginning with the method of weak expectations. Repeating the game an arbitrarily large number of times yields a weak expectation a dollar higher than the Pasadena game. Therefore a rational agent willingly pays more to participate in the Altadena game.

An alternate solution is found in Colyvan's Relative Expected Utility (REU).[31] His decision rule assesses the expected relative advantage of choosing an act in one game over the corresponding act in another game. The expected utility of benefits or gains associated with all act-pairs determines the preferable game. This manner of construction allows REU to integrate dominance reasoning within expected utility theory. The resulting axiomatic rule is consistent with intuitions in valuing the Altadena over the Pasadena game.

These pluralistic methods nevertheless leave us in a confounding situation. On one hand, certain decision problems like the Altadena vs. Pasadena dilemma find pluralistic decision rules at their disposal. Meanwhile, other paradoxes render all existing solutions unworkable. The Cauchy distribution is one class of paradox that exposes the full magnitude of EU theory's normative challenges. There is no stable expectation, even if the number of repeated plays is fixed arbitrarily high in advance. Despite having a median of 1, the skewness of the Cauchy distribution makes the choice of this statistical property problematic. The issue remains of how to systematically address these paradoxes, given that some have numerous solutions and others have nothing that appears immediately workable.

One feasible path forward is lexicographically-based. If we admit that weak expectation resolves the normative paradox in games like Pasadena or Altadena, we should prioritize the use of the Weak Law of Large Numbers in coming up with stable solutions for paradoxical

[31] Formally, relative expected utility of act A_k over A_l is defined as: $REU(A_k, A_l) = \sum_{i=1}^{\infty} p_i(u_k i - u_l i)$ where p_i is the probability associated with the state S_i and u_{ji} represent utilities resulting from act A_j in state S_j

games.[32] This addresses certain categories of games, but leaves the determination of others unanswered. Particularly, cases of comparing two divergent series such as the St. Petersburg and Petrograd[33] games requires the use of REU or similar types of dominance reasoning to reach a choice. This modification to EU theory avoids the indecision arising from properties of infinity and prescribes a choice for the rational agent. If both weak expectation and the incorporation of dominance principles fail to yield a conclusion, a statistical parameter can guide decision-making. The choice of parameter depends critically on the probability distribution and its known properties, from variance to skewness and kurtosis. Understanding the distribution is necessary for choosing a preferable parameter as a proxy for expectation. Although a general rule for this is desirable, the wide differences in probability distributions suggest that the determination may need to happen on a case-by-case basis. Further investigation into this methodology is required.

This lexicographic method is incomplete in the sense that it can only resolve a subset of all normative paradoxes for decision-making. Nevertheless, it provides several basic formulaic guidelines to addressing normative challenges. After all, the most sorely needed tool is not a register of all existing paradoxes and applicable decision rules, but rather a unifying framework that dictates which rule is rationally preferable, and under what circumstances.

[32] It is likely that weak expectation is mainly applicable for games that are conditionally convergent and hence subject to Riemann's Rearrangement Theorem.

[33] A variation of the St. Petersburg game in which every prize is $1 higher.

References

Alexander, Jason (2010), "Decision Theory Meets the Witch of Agnesi," [working paper].

Colyvan, Mark (2006), "No Expectations," *Mind*, Vol. 115, No. 459, pp. 695-702.

Colyvan, Mark (2008), "Relative Expectation Theory," *Journal of Philosophy*, Vol. 105, pp. 37-44.

Easwaran, Kenny (2008), "Strong and Weak Expectations," *Mind*, Vol. 117, No. 467, pp. 633-641.

Fine, Terrence (2008), "Evaluating the Pasadena, Altadena, and St. Petersburg Gambles," *Mind*, Vol. 117, No. 467, pp. 613-632.

Hacking, Ian (1980), "Strange Expectations," *Philosophy of Science* 47, pp. 562-567.

Martin, Robert (2001), "The St. Petersburg Paradox," in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2001 Edition)

Nover, Harris and Alan Hájek, (2004), "Vexing Expectations," *Mind*, Vol. 113, No. 450, pp. 237-249.

Nover, Harris and Alan Hájek, (2006), "Perplexing Expectations," *Mind*, Vol. 115, No. 459, pp. 721-730.

Nover, Harris, and Alan Hájek, (2008), "Complex Expectations," *Mind*, Vol. 117, No. 467, pp. 643-664.