Seidl, Christian

Working Paper
The Petersburg Paradox: Menger revisited


Provided in Cooperation with:
Christian-Albrechts-University of Kiel, Department of Economics

Suggested Citation: Seidl, Christian (2012) : The Petersburg Paradox: Menger revisited, Economics Working Paper, No. 2012-04, Kiel University, Department of Economics, Kiel

This Version is available at:
http://hdl.handle.net/10419/56036

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.
If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
the peterburg paradox: menger revisited

by Christian Seidl
Abstract

The Petersburg Paradox and its solutions are formulated in a uniform arrangement centered around d’Alembert’s ratio test. All its aspects are captured using three mappings, a mapping from the natural numbers to the space of the winnings, a utility function defined on the space of the winnings, and a transformation of the utilities of the winnings. The main attempts at a solution of the Petersburg Paradox are labeled according to their most fervent proponents, viz. Bernoulli and Cramer, Buffon, and Menger. This paper also investigates the role of the probabilities for the Petersburg Paradox: they may well be used to “solve” a Petersburg Paradox, or to re-gain it by means of appropriate transformations. Thus, the probabilities are also instrumental for the Petersburg Paradox. The Petersburg Paradox can only be avoided for bounded utility functions. Its various solution proposals are but disguised attempts of filling in the missing behavioral justification for the boundedness of utility functions. This paper also corrects several misconceptions which have crept in the respective literature.

1 The Petersburg Paradox Outlined

“Ever since mathematicians first began to study the measurement of risk there has been general agreement on the following proposition: Expected values are computed by multiplying each possible gain by the number of ways in which it can occur, and then dividing

*I am indebted to Peter J. Hammond, Stanford University, U.S.A., and University of Warwick, U.K., and Kirill Pogorelskiy, CalTech, Pasadena, U.S.A., for most helpful comments. Remaining errors are mine.
the sum of these products by the total number of possible cases where, in this theory, the consideration of cases which are all of the same probability is insisted upon.” (D. Bernoulli (1954, p. 23))

This was the state of the art of evaluating the fair value of a gamble before Nicolas Bernoulli (1713-1732) and Daniel Bernoulli (1738) discovered an amazing paradigm for which this rule failed. If a fair coin is flipped and comes up head, the gambler receives one ducat. If it comes up head only at the second toss, the gambler receives 2 ducats. If it comes up head first at the third toss, the gambler receives 4 ducats. Generally, if it comes up head first at the $i$-th toss, the gambler receives $2^{i-1}$ ducats. Although the expected value of this gamble is infinity, no reasonable person will stake more than a comparatively modest amount for participating in this game.

The challenge to “solve” this paradox has captivated the attention of many prominent scholars, starting with Gabriel Cramer and Daniel Bernoulli. The final keystone of this edifice was set by Menger (1934), who established that a necessary and sufficient condition of securing a finite expected value of this gamble is the boundedness of the winnings. Since Menger’s article is rather verbose and idiosyncratic, the present paper endeavors to revisit Menger’s seminal contribution. It also takes the occasion to concisely restate the various ways to resolve this paradox.

The main sources are the correspondence of Nicolas Bernoulli (1713-1732) with contemporary mathematicians and Daniel Bernoulli’s (1738) famous essay. Since this essay was published in the publication series of the Petersburg Academy of Sciences, this paradox has since become known as the Petersburg Paradox. The original essay was written in Latin, and was twice translated. First into German by the mathematician Alfred Pringsheim (Thomas Mann’s father in law) [D. Bernoulli (1896)], and second into English by Louise Sommer for Econometrica [D. Bernoulli (1954)]. Whereas the Sommer translation reproduces the plain Latin original in English, the Pringsheim translation in addition abounds in sagacious and helpful comments by Pringsheim. Moreover, it is accompanied by a thoughtful introduction by Ludwig Fick. Hence, I used the Pringsheim translation for this paper. I have also named the various solutions according to their most fervent proponents.

The mathematics is presented in Section 1 and the discussion in Section 2.

---

1Around 1654, the nobleman gambler, Chevalier de Mérè, brought the famous mathematicians Fermat and Pascal to elaborate for him guidance for optimum gambling. Their suggestions centered around the expected value of the winnings. Cf. Samuelson (1977, pp. 37-38).
2 Mathematics

THEOREM 1: Let \( \mathbb{N} \) denote the set of natural numbers and let \( x : \mathbb{N} \rightarrow X \subset \mathbb{R}_+ \) denote a strictly increasing mapping, i.e., \( x_j > x_i \ \forall j > i, \ i,j \in \mathbb{N} \). Let \( u : X \rightarrow \mathbb{R}_+ \) denote a nondecreasing function such that \( u(x_i) < \infty \ \forall i < \infty \), and \( p : X \rightarrow [0, 1] \), a nonincreasing function such that \( \sum_{i=1}^{\infty} p(x_i) = 1 \), i.e., a probability distribution. We consider four cases:

Case 1: \( \sum_{i=1}^{\infty} u(x_i)p(x_i) < \infty \) if \( \exists i^* < \infty \) such that \( p(x_i) = 0 \ \forall i \geq i^* \).

Case 2: \( \sum_{i=1}^{\infty} u(x_i)p(x_i) < \infty \) if \( \exists i^* < \infty \) such that \( \sup_{i \geq i^*} \frac{u(x_{i+1})p(x_{i+1})}{u(x_i)p(x_i)} < 1 \).

Case 3: \( \sum_{i=1}^{\infty} u(x_i)p(x_i) = \infty \) if \( \exists i^* < \infty \) such that \( \sup_{i \geq i^*} \frac{u(x_{i+1})p(x_{i+1})}{u(x_i)p(x_i)} > 1 \).

Case 4: \( \sum_{i=1}^{\infty} u(x_i)p(x_i) \) may converge or diverge if \( \exists i^* < \infty \) such that \( \sup_{i \geq i^*} \frac{u(x_{i+1})p(x_{i+1})}{u(x_i)p(x_i)} = 1 \).

PROOF: Case 1 is immediate. Cases 2, 3, and 4 follow from d’Alembert’s ratio test (see, e.g., Stephenson (1973, pp. 75-76)).

COROLLARY 2 (Petersburg Case): For each probability distribution \( p(x_i) > 0 \), \( \sum_{i=1}^{\infty} p(x_i) = 1 \), which is strictly decreasing for all \( i \geq i^* \), \( i^* < \infty \), there exist functions \( u(x_i) \) which are strictly increasing for all \( i \geq i^* \), \( i^* < \infty \), such that \( \sum_{i=1}^{\infty} u(x_i)p(x_i) = \infty \).

PROOF: Applying Case 3 of Theorem 1, we consider functions \( u(x_i) \) such that

\[
\frac{u(x_{i+1})}{u(x_i)} \geq \frac{p(x_i)}{p(x_{i+1})} > 1 \quad \forall i \geq i^*,
\]

that is, the growth of the \( u(\cdot) \)-function does not fall below the shrinkage of the \( p(\cdot) \)-function for infinitely many items. If the equality sign holds instead of the weak inequality sign in the above relation, then Case 4 of Theorem 1 holds.

REMARK 3: The “classical” Petersburg paradox assumes \( x_i = i \), \( u(x_i) = 2^{i-1} \), \( p(x_i) = 2^{-i} \), which implies \( \sum_{i=1}^{\infty} u(x_i)p(x_i) = 2^{-1} \times \infty = \infty \). Note that \( \frac{u(x_{i+1})}{u(x_i)} = 2 \), \( \frac{p(x_i)}{p(x_{i+1})} = 2 \), and \( \frac{u(x_{i+1})p(x_{i+1})}{u(x_i)p(x_i)} = 1 \) for all \( i \in \mathbb{N} \). That is, the “classical” Petersburg Paradox comes up to Case 4 of Theorem 1 for which \( \sum_{i=1}^{\infty} u(x_i)p(x_i) = \infty \). Hence, the Petersburg Case may hold for some instances of Case 4 and holds quite generally for Case 3 of Theorem 1.
Proof: By assumption $\frac{p(x_i)}{p(x_{i+1})} > 1 \forall i < \infty$. Select $\psi(\cdot)$ such that $\frac{\psi[u(x_{i+1})]}{\psi[u(x_i)\cdot]} < \frac{p(x_i)}{p(x_{i+1})} \forall i > i^*, i^* < \infty$. This implies $\frac{\psi[u(x_{i+1})]p(x_{i+1})}{\psi[u(x_i)]p(x_i)} < 1 \forall i > i^*$. Hence, by Case 2 of Theorem 1 we have $\sum_{i=i^*}^\infty \psi[u(x_i)]p(x_i) < \infty$. This implies $\sum_{i=1}^\infty \psi[u(x_i)]p(x_i) < \infty$.

Remark 5: Note that Bernoulli proposed $\psi[u(x_i)] = \ln[u(x_i)] = \ln 2^{i-1}$, and Cramer proposed $\psi[u(x_i)] = \sqrt{u(x_i)} = \sqrt{2^{i-1}}$. Since these functions were proposed for $p(x_i) = 2^{-i}$, we have

for the Bernoulli proposal: $\frac{\psi[u(x_{i+1})]p(x_{i+1})}{\psi[u(x_i)]p(x_i)} = \frac{i}{2(i-1)} < 1$ for $i > 2$.

for the Cramer proposal: $\frac{\psi[u(x_{i+1})]p(x_{i+1})}{\psi[u(x_i)]p(x_i)} = \frac{\sqrt{2}}{2} < 1$ for $i \in \mathbb{N}$.

Remark 6: Note that a Petersburg Paradox is brought about by the winnings and the probability distribution. Suppose we replace the probabilities of the Cramer solution by $\tilde{p}(x_i) = \frac{1}{\pi \sqrt{2}^i}$, where $\pi = \sum_{i=1}^\infty 2^{-\frac{i}{2}} = \frac{\sqrt{2}}{\sqrt{2} - 1} = \frac{2^{\frac{3}{2}}}{2^{\frac{3}{2}} - 1}$. Obviously,

$$\sum_{i=1}^\infty \tilde{p}(x_i) = \sum_{i=1}^\infty \frac{1}{\pi \sqrt{2}^i} = \sum_{i=1}^\infty \frac{\frac{1}{2^\frac{3}{2}}}{\frac{1}{2^\frac{3}{2}} - \frac{1}{2^\frac{3}{2}}} = (1 - 2^{\frac{-\frac{3}{2}}{2}}) \sum_{i=1}^\infty 2^{-\frac{i}{2}} = 1,$$

and

$$\sum_{i=1}^\infty \psi[u(x_i)]\tilde{p}(x_i) = \sum_{i=1}^\infty \frac{\sqrt{2^{i-1}}}{\sqrt{\frac{1}{2} - \frac{1}{2}}} = \sum_{i=1}^\infty \left( \frac{1}{\sqrt{2}} - \frac{1}{2} \right) = \infty,$$

which re-establishes the Petersburg Paradox. Hence, a concave transformation of the winnings.

---

2It seems that Tintner (1941, p. 301) was the first to suggest the concept of a preference functional, which included beyond utility also properties of the probability distribution.

3In addition to the gambler's winnings, Bernoulli took also the gambler's wealth into consideration. We will deal with this extension in the next section.

4The equivalent roles of the winnings or their utilities on the one hand and the probability distribution on the other for expected values has ever been daily routine in statistics; see, e.g., DeGroot (1970, pp. 89-91). In the literature on the Petersburg Paradox, the role of the probabilities seems to have been largely neglected.

5It is easily checked that $\frac{\psi[u(x_{i+1})]}{\psi[u(x_i)]} = \frac{\sqrt{2}}{\ln 2^i} = \sqrt{2}$ and $\frac{\tilde{p}(x_{i+1})}{\tilde{p}(x_i)} = \frac{1}{\pi \sqrt{2}^{i+1} \sqrt{2}^{i}} = \frac{1}{\sqrt{2}}$. Hence,
nings does not necessarily “solve” the Petersburg Paradox. It is regained if the shrinkage of the probabilities is adequately slowed down.

However, an adequate probability adaption does not work for the Bernoulli solution. In order to regain \(\sum_{i=1}^{\infty} \ln 2^{i-1} \hat{p}(x_i) = \infty\), we would at least need \(\hat{p}(x_i) = \frac{1}{\lambda \ln 2^i}\), where \(\lambda > 0\) denotes a calibration constant. Yet \(\sum_{i=1}^{\infty} \frac{1}{\ln 2^i} = \frac{1}{\ln 2} \sum_{i=1}^{\infty} \frac{1}{i} = \infty\) because the sum of a harmonic series is infinite. This means that for the Bernoulli solution probabilities which shrink slow enough to compensate the growth of the logarithm of the winnings do not exist.

**Remark 7:** Alternatively, the Petersburg Paradox can be solved by manipulating the probability distribution rather than applying a transformation to the winnings. Suppose we have a Petersburg game with winnings \(2^{i-1}\) at the \(i\)-th toss, but the game is somewhat re-defined. Two coins are flipped at the same time. The gambler gets the winning \(2^{i-1}\) if some other result than two heads comes up for the first time at the \(i\)-th toss. The respective probabilities are \((\frac{1}{4})^{i-1} (\frac{3}{4})\) and \((\frac{3}{4}) \sum_{i=1}^{\infty} \left(\frac{1}{4}\right)^{i-1} = 1\). This implies \((\frac{3}{4}) \sum_{i=1}^{\infty} \left(\frac{1}{4}\right)^{i-1} 2^{i-1} = \frac{3}{2}\). This example demonstrates that the Petersburg Paradox cannot exclusively be “solved” by a transformations of the winnings. Probabilities which shrink relatively faster to zero than the winnings increase relatively to infinity perform the same job. This other source of solving the Petersburg Paradox was largely ignored by the literature.

**Corollary 8 (Buffon Case):** For all nondecreasing functions \(u(x_i) < \infty, i < \infty\), we have \(\sum_{i=1}^{\infty} u(x_i)p(x_i) < \infty\) for all probability distributions with \(p(x_i) = 0 \forall i > i^*, i^* < \infty\).

**Proof:** Case 1 of Theorem 1.

**Corollary 9 (Menger Case):** For all nondecreasing functions \(u(x_i) < \infty, \forall i < \infty\), and all probability distributions \(p(x_i) \geq 0, \sum_{i=1}^{\infty} p(x_i) = 1, \sum_{i=1}^{\infty} u(x_i)p(x_i) < \infty\) holds if and only if \(\exists B < \infty\) such that \(u(x_i) \leq B \forall i \in \mathbb{N}\).

**Proof:** Obviously \(u(x_i) \leq B < \infty \forall i \in \mathbb{N}\) implies \(\sum_{i=1}^{\infty} u(x_i)p(x_i) \leq \sum_{i=1}^{\infty} Bp(x_i) = B < \infty\).

Conversely, assume \(u(x_i) \rightarrow \infty\) as \(x_i \rightarrow \infty\). Consider a strictly decreasing probability distribution \(p(x_i) > 0 \forall i < \infty\). By assumption \(\frac{p(x_i)}{p(x_{i+1})} > 1 \forall i < \infty\). Then there exist
strictly increasing transformations \( \psi(\cdot) \) such that

\[
\frac{\psi[u(x_{i+1})]}{\psi[u(x_i)]} > \frac{p(x_i)}{p(x_{i+1})} \iff \frac{\psi[u(x_{i+1})]p(x_{i+1})}{\psi[u(x_i)]p(x_i)} > 1 \quad \forall \, i \geq i^*, \, i^* < \infty.
\]

Then we have Case 3 of Theorem 1 for \( i \geq i^* \). Hence, boundedness of \( u(\cdot) \) is a necessary and sufficient condition for \( \sum_{i=1}^{\infty} u(x_i)p(x_i) < \infty \).

### 3 Discussion

#### 3.1 Petersburg Case

On 9th September 1713, Nicolas Bernoulli\(^6\) wrote a letter to Pierre Rémond de Montmort, asking him five problems. The fourth problem was: “A promises to give a coin to B, if with an ordinary die he achieves 6 points on the first throw, two coins if he achieves 6 on the second throw, 3 coins if he achieves this point on the third throw, 4 coins if he achieves it on the fourth and thus it follows; one asks what is the expectation of B?” The fifth problem was more tricky. Here, Bernoulli replaced 1,2,3,4, etc., by 1,2,4,8, etc., or 1,3,9,27 etc., or 1,4,9,16,25 etc., or 1,8,27,64. We see immediately that the expected value of the fourth example is \( \sum_{i=1}^{\infty} \frac{i}{6} \left(\frac{5}{6}\right)^{i-1} \). The ratio test shows \( 1 + \frac{1}{6} \left(\frac{5}{6}\right) < 1 \) for \( i > 5 \), which implies Case 2 of Theorem 1 indicating a finite expected value of the fourth example. The expected value of the fifth example is for the series 1,2,4,8, etc.:

\[
\sum_{i=1}^{\infty} \frac{i}{6} \left(\frac{10}{6}\right)^{i-1}.
\]

Since the ratio test shows \( \frac{10}{6} > 1 \), the expected value of the fifth example for this (and all other) series amounts to infinity.

In his response of 15th November 1713, Montmort expressed opinion that these examples have easy solutions along the lines of geometric progressions. However, meanwhile Montmort (1713, p. 402) had entered Bernoulli’s problems into the second edition of his *Essay d’analyse sur les jeux de hazard*. It was only in his letter of 20th February 1714, that Nicolas Bernoulli demonstrated the momentousness of his discovery to Montmort, who expressed skepticism, but had to admit that he was unable to solve this problem.

Gabriel Cramer, a professor of mathematics at the University of Geneva, read Montmort’s book and pondered on Bernoulli’s fifth problem. In a letter of 21st May 1728 from London, he addressed Bernoulli proposing first a simplification of the gamble by replacing throwing a die by flipping a coin (whereby he reformulated the problem as it had become known since). In addition to that he proposed two solutions of the paradox,

---

\(^6\)All citations are to be found in N. Bernoulli (1713-1732).
viz. by bounding the (gambler’s perception of) winnings from above, or by applying an increasing transformation to the winnings. These proposals will be taken up below.

In a letter of 27th October 1728, Nicolas Bernoulli communicated the fifth problem in Cramer’s simplified version to his cousin Daniel Bernoulli who was at that time a professor of mathematics at the University of St. Petersburg. Subsequently, Daniel had become interested in this paradox and had sent Nicolas a first draft of his later publication. In his response of 5th April 1732, Nicolas thanked Daniel for a copy of his essay and remarked: “I have read it with pleasure, and I have found your theory most ingenious, but permit me to say to you that it does not solve the knot of the problem in question.” Together with this response, Nicolas sent Daniel a copy of Cramer’s letter, which Daniel appended to his manuscript for its final publication (D. Bernoulli (1896, pp. 56-60)). Here we find the same version of the Petersburg Paradox (D. Bernoulli (1896, p. 56)), but in terms of Thaler’s rather than in ducats.

This illustrates the origin of the Petersburg Paradox. Several other prominent scholars have also dealt with it.

An important consequence of the Petersburg Paradox is that it obscures obvious superiority among gambles. For instance, replace in a Petersburg gamble the winnings $2^i$ by $3^i$. Then a partial sum test shows the following:

$$\sum_{i=1}^{j} \frac{2^{i-1}}{2^i} < \sum_{i=1}^{j} \frac{3^{i-1}}{2^i} \forall j > 1 \text{ and } j < \infty, \text{ but } \lim_{j \to \infty} \sum_{i=1}^{j} \frac{2^{i-1}}{2^i} = \lim_{j \to \infty} \sum_{i=1}^{j} \frac{3^{i-1}}{2^i} = \infty.$$

Thus, whereas a partial sum test reveals clear superiority of the second gamble, this superiority breaks down as the expected values approach infinity. For this case, the ordering among all respective gambles becomes paralyzed (see also Arrow (1970, pp. 64-65)).

---

7Pringsheim translates uncle (D. Bernoulli (1896, pp. 46 and 56)), but Samuelson (1977, p. 37) has convincingly shown that the said Nicolas was Daniel’s cousin. The confusion might be explained because Nicolas’ father had also the first name Nicolas and was the brother of Daniel’s father Johann. Perhaps Daniel Bernoulli was also led to express esteem for the older Nicolas Bernoulli.

8Note that D. Bernoulli was ignorant of the content of Cramer’s letter to N. Bernoulli before that time. Instead he had concluded the main part of his manuscript before N. Bernoulli had sent him a copy of Cramer’s letter. Savage (1972, p. 92) erroneously argued that D. Bernoulli had developed his ideas under the influence of Cramer’s letter. Moreover, Savage (1972, pp. 94-95) purported that Cramer, in his aforementioned letter, had pointed to the disadvantages of the logarithm. But Cramer’s letter dates from 21st May 1728, way before D. Bernoulli had proposed the logarithmic function as a solution of the Petersburg paradox. Rather D. Bernoulli had sent the first draft of his essay on 4th July 1731 to N. Bernoulli; see N. Bernoulli (1713-1732).

9For an exhaustive account of prominent scholars cf. Samuelson (1977, pp. 36-53). Also Keynes (1921, pp. 362-368) dealt at length with the Petersburg Paradox, however, without having taken a position of his own of how to “solve” it.
Note that by an appropriate assumption a partial sum test can be rendered serviceable for ordering gambles. If there exists a \( j^* \) < \( \infty \) such that \( \sum_{i=1}^{j} u(x_i)p(x_i) \geq \sum_{i=1}^{j} \tilde{u}(\tilde{x}_i)\tilde{p}(x_i) \) for all \( j \geq j^* \), then the first lottery is considered as superior to the second lottery. This is basically Assumption \( U_6 \) of DeGroot (1970, p. 112).

Moreover, although the expected value of a Petersburg game is infinity, no reasonable person would pay more than a comparatively modest amount to participate in such a game. The following sections deal with attempts at a resolution of the Petersburg Paradox.

### 3.2 Bernoulli-Cramer Solutions

Daniel Bernoulli and Gabriel Cramer suggested appropriate transformations of the winnings, which we have come to denominate as utility,\(^{10}\) such that the growth rate of the utilities fell short of the shrinkage rate of the probabilities. To come to grips with the classical Petersburg Paradox, D. Bernoulli (1896, pp. 34-38) suggested a transformation by the natural logarithm, while Cramer (cf. D. Bernoulli (1896, pp. 59-60)) suggested a transformation by taking the square root. Buffon (1777/2010, p. 44-45) proposed to replace \( 2^i \) in the Petersburg Paradox by \( (\frac{9}{5})^i \) to capture the shrinking “morale value” of the winnings,\(^{11}\) a proposal which was later rekindled in a similar fashion by Fry (1928, pp. 196-197).

Menger (1934, pp. 467 and 481) belittled the ad-hoc nature of this approach, which, while applying to Cramer’s proposal, appears ill-founded for Bernoulli’s logarithmic transformation. D. Bernoulli (1896, pp. 27-28 and 34-38) adopted the assumption that the utility of a winning is inversely proportional to the gambler’s wealth. Let \( w \) denote a gambler’s wealth, then his or her utility is \( u(x + w) = b\ln\left(\frac{w+x}{w}\right) \),\(^{12}\) which is a rather modern behavioral assumption. This relationship shows, first, that the marginal utility of a winning is the smaller the larger the gambler’s wealth is, second, that additional constant winnings have decreasing marginal utility, and, third, that the utility of a gain falls

---

\(^{10}\)With respect to the transformed winnings, the translations differ: whereas we find in D. Bernoulli (1896, p. 26) the term \textit{Vorteil} (advantage), we find in D. Bernoulli (1954, p. 24) the term \textit{utility}. In the German translation, Cramer is attributed the terms \textit{moralischer Wert} (moral value) and \textit{moralische Hoffnung} (moral hope) (D. Bernoulli (1896, pp. 58 and 60)), whereas he is attributed the terms \textit{moral expectation} and \textit{psychic expectation} in the English translation (D. Bernoulli (1954, p. 34)). From a modern point of view we would, however, not fail to conceive of this device as \textit{expected utility}.

\(^{11}\)Buffon (1777/2010, p. 36) reported that he had visited Cramer in Geneva in 1730 and was well acquainted with the Petersburg Paradox. He expressed agreement with the Cramer and D. Bernoulli solutions.

\(^{12}\)D. Bernoulli (1896, pp. 34-37, and, in particular, cf. Pringsheim’s revealing Footnote 4). Savage (1972, p. 94) remarked: “To this day, no other function has been suggested as a better prototype for Everyman’s utility function.”
short of the disutility of an equivalent amount of money lost. These assumptions are in line with modern economics. Moreover, Bernoulli’s hypothesis was later confirmed by Weber (1834) and Fechner’s (1860) experimental investigations on psychophysics (cf. Fick in D. Bernoulli (1896, pp. 11 and 18) and Stigler (1950b, pp. 375-377)).

D. Bernoulli made good use of the dependence of the marginal utility of winnings and losses on a person’s wealth. In particular, he was able to explain why some hazard should be insured by less wealthy persons while richer persons should better carry the risk of the same hazard by themselves (D. Bernoulli (1896, pp. 42-44)). He also presented a quite modern theory of portfolio selection for risk spreading (D. Bernoulli (1896, pp. 44-45)). Dependence of utility on the gambler’s wealth provokes a lot of consequences, e.g. that a gambler’s willingness to pay and to accept may diverge, and that the selling price of a gamble will exceed a gambler’s willingness to pay. Samuelson (1977, pp. 27-30) has painstakingly dealt with such extensions. However, this gives rise to the question why only the gambler’s wealth should matter and not also his or her utility function (cf. also Menger (1934, pp. 480 and 484). But this would lead to a plethora of possible cases.

Samuelson (1977, pp. 30-31) dealt also with the bankruptcy problem. While this problem is legitimate for the bookmaker of the gamble, it sounds far-fetched with respect to the gambler. Indeed, a gambler’s stake may exceed his wealth, but then the gamble does not came about when the stake has to be paid in cash. Moreover, in a modern economy a gambler can take a credit jointly with taking insurance against too early (and thus, insufficient) winnings, which should be available in face of the infinite expected value of the gamble. After all, an insurer can insure plenty of (stochastically independent) Petersburg gambles (this can be inferred from Menger (1934, pp. 463-464); see in analogy also Knight (1921, Section I.II.36)). Moreover, explaining bankruptcy may require the assumption of a convex utility function (Samuelson (1977, p. 31)).

But such considerations are beyond the mere solution of the Petersburg Paradox. Hence, we will proceed along the lines of D. Bernoulli (1896, pp. 46-55). The expected value of a Petersburg game in monetary terms, which was donated to a gambler, is $\exp\{\sum_{i=1}^{\infty} 2^{-i} \ln[w + 2^{i-1}]\} - w$. We may consider it likewise as the selling value of this Petersburg game. If the gambler had paid the stake $s$ for this gamble, its expected value

---

13Hence, Daniel Bernoulli can be considered a precursor of Kahneman and Tversky’s (1979) prospect theory. The same can be said of Buffon (1777/2010, pp. 31-33). Blavatsky (2005) observed that the overweighing of small probabilities in Tversky and Kahneman’s (1992) cumulative prospect theory may well restore a Petersburg Paradox.

14This problem was also recognized by Menger (1934, p. 473).

15Note that D. Bernoulli (1896, pp. 32-33, 41-42, and 52-53) ruled out that a gambler’s stake would exceed his or her wealth. See also Menger (1934, p. 461 and 475-478).
in monetary terms is \( \exp\{\sum_{i=1}^{\infty} 2^{-i} \ln[w - s + 2^{i-1}]\} - w. \) Since \( s > 0, \) the latter value is smaller than the former. Bernoulli argues that the latter expression approximates the former one if \( w \) is comparatively big in relation to \( s. \)

Menger (1934, p. 467-468) objected to the Bernoulli-Cramer solution that a transformation of the winnings which is "sufficiently concave" is only a sufficient, but not a necessary condition for "solving" the Petersburg Paradox. To demonstrate that he proposed to replace \( (x_i) = 2^i \) by \( x_i = \exp 2^i. \) Applying now \( \psi(\cdot) = \ln(\cdot) \) to \( x_i = \exp 2^i \) regains the Petersburg Paradox. [Replacing \( (x_i) = 2^i \) by \( (x_i) = (2^i)^2 \) regains the Petersburg Paradox in Cramer's version for \( \psi(\cdot) = \sqrt{\cdot}. \)] Samuelson (1977, p. 32) called this Menger's Super-Petersburg Paradox. More generally, for each increasing and unbounded utility function can an increasing transformation be defined such that the transformed winnings converge relatively faster to infinity than the probabilities converge to zero.

### 3.3 Buffon Solutions

Among several other solutions, Buffon (1777/2010, pp. 23, 38, and 46) proposed for "solving" the Petersburg Paradox, to set all probabilities smaller than \( \frac{1}{10000} \) equal to zero, although he (p. 24) had some difficulty of neglecting the probability of sudden death. Related to setting small probabilities equal to zero is Buffon’s (1777/2010, pp. 39-40) argument that a Petersburg gamble to end at the 30th round at least because the needed payoff would be as great as the whole fortune in the Kingdom of France. Since no bookmaker would dispose of so much money, it can be inferred that the probability of the effectuation of payoffs associated with tossing a head after the 30th or later toss of a tail comes up to be zero. Moreover, the time required to come to an end with a Petersburg gamble might become excessively long (Buffon (1777/2010, pp. 43-44); Pringsheim, see D. Bernoulli (1896, pp. 47-48)). Buffon (1777/2010, pp. 41-44) reports also on a coin tossing experiment, which he had conducted, and compares its results with the theoretical payoffs. Both the experiment and the theoretical payoffs for the same number of tosses yield an expected value of some 5 ecus (a French currency unit used by Buffon).

In these reasonings, Buffon was followed by Fontaine, Poisson, d’Alembert, and Condorcet (Samuelson (1977, pp. 39-40)). Also Pringsheim argued along the line that huge payoffs can never be effectuated (D. Bernoulli (1896, pp. 46-52 and 57)).

---

16D. Bernoulli (1896, pp. 49-55). For the Cramer proposal, the logarithm should be replaced by the square root, and the exponential function by taking the square. For the Buffon proposal, \( 2^i \) should be replaced by \( \left(\frac{5}{8}\right)^i \).
Interestingly enough, even Menger (1934, pp. 473 and 476) believed that small probabilities are undervalued by subjects.¹⁷ Whereas Menger did not wonder about the success of commercial lotteries with enormous prizes which could be won with very small probabilities, the success of commercial lotteries had troubled Buffon (1777/2010, p. 48). He explained the success of commercial lotteries sarcastically:

> The display of hope is the business of all money swindlers. The big art of the lottery maker is to present large sums with very small probabilities, soon swollen by the spring of greed. These swindlers still enlarge this ideal product, dividing it and giving it for very little money, which everyone can loose, a hope that, though much smaller, seems to be a part of the grandeur of the total sum.

Hence, Buffon explained the success of commercial lotteries by fraudulent concealment of the tininess of the probabilities of large winnings.

### 3.4 Menger Solutions

The chronologically last solution proposal of the Petersburg Paradox was proposed by Menger (1934). Menger indeed proved Corollary 9, but in a somewhat more circumstantial way. In his article, he proved sufficiency in Section 6 (pp. 469-471), and necessity with the help of his Super-Petersburg game in Section 5 (pp. 468-469). In the necessity part of his proof he argued that for each unbounded function \( u(x_i) \) there exists a related Petersburg game whose expected transformed winnings amount to infinity.

As concerns sufficiency, Menger’s proposal was not novel. In his letter to Nicolas Bernoulli of 21st May 1728, Cramer suggested bounding the winnings at \( 2^{24} \) coins, because larger sums do not make for more pleasure of the gambler or because the gambler is never able to receive more than this sum (N. Bernoulli (1713-1732); see also D. Bernoulli (1896, pp. 58-59)). Following this assumption, Cramer ascertains a value of 13 coins for this gamble. In one of his proposals to solve the Petersburg Paradox, Buffon (1777/2010, p. 39) had argued in analogy to Cramer. Several other authors justified boundedness of the winnings by referring to the finiteness of the bookmaker’s wealth. A bouquet of such authors includes Buffon (1777/2010, pp. 39-40), Pringsheim (D. Bernoulli (1896, 1896, 

¹⁷Menger (1934) was unable to adduce empirical evidence for this assumption. Backed by ample empirical research, modern theories, e.g., Kahneman and Tversky (1979) and Tversky and Kahneman (1992), assume overvaluation of small probabilities. Prospect theory became famous for its accurate forecasts of human behavior.
Fry (1928, pp. 198-199), Shapley (1977a, p. 440) and, although considerably more cautious, Shapley (1977b, pp. 448-449). A similar argument was put forward by Brito (1975) and Cowen and High (1988), viz. that the utility of winnings is finite because a gambler has not enough time to consume that much money involved by extremely high winnings, or because the marginal utility of money becomes zero as the gambler disposes of extremely high sums of money.

A modern variety of truncation to come to grips with unbounded winnings or utility functions was developed by Wakker (1993). Central to this—very complicated—theory is the concept of truncation continuity, which works in effect by way of truncation of the upper (and/or lower) tail of expected utility. By this assumption unbounded utility functions become manageable.

The principal contribution by Menger (1934) was his necessary condition, i.e., to have shown that a Petersburg gamble has a finite solution only if the utility of winnings is bounded.\textsuperscript{18} Samuelson, who had a lot to quarrel with Menger (Samuelson (1977, pp. 32-33 and pp. 47-49)), frankly acknowledged that Menger’s breakthrough was a quantum jump in the analysis of the Petersburg Paradox (Samuelson (1977, pp. 32 and 37)). Menger’s achievement was also acclaimed by other Nobel laureates such as Arrow (1970, pp. 23 and 64-69) and Aumann (1977). In a painstaking analysis, Arrow (1970, pp. 64-69) has shown that a utility function for risky prospects has to be bounded from above and from below when instances of the Petersburg Paradox should be avoided.\textsuperscript{19} This implies that a utility function for risky prospects must have at least one Friedman-Savage inflexion point (Samuelson (1977, p. 50, Footnote 14)) which forms also the central constituent of Kahneman and Tversky’s prospect theory.

In view of two centuries’ discussions on such a provocative subject like the Petersburg Paradox, it takes wonder why such simple mathematics like Corollary 9 had to wait until Menger’s (1934) discovery.\textsuperscript{20} This seems to be due to the comprehension of utility as it developed in the course of several centuries. After early contemplations by Thomas Hobbes, Francis Hutcheson, and David Hume, it was, in particular, Bentham (1781/2000)  

\textsuperscript{18}\textsuperscript{18}Before its 1967 English translation in the Morgenstern-Festschrift, Menger’s result was widely ignored in the Anglo-Saxon literature, except by Arrow in 1951 and 1965 (reprinted as Essays 1 and 2 in Arrow (1970)). Perhaps misled by an ambiguous footnote in Arrow’s 1951 \textit{Econometrica} article (Arrow (1970, p. 23, Footnote 18)), Savage (1972, p. 95) erroneously attributed this insight to Cramer: “Cramer therefore concluded, and I think rightly, that the utility of cash must be bounded, at least from above.” Interestingly enough, Stigler (1950b, p. 375, Footnote 129) drew on this result, however, without any reference to Menger.

\textsuperscript{19}\textsuperscript{19}Fishburn (1970, pp. 206-207) was the first to prove that the utility function of Savage’s theory is bounded.

\textsuperscript{20}\textsuperscript{20}Bassett (1987, p. 522), who had raised this question eventually, had to resign to the fact that “there is really no explanation”.
who elevated utility to the central driving power of human actions. In Chapter I, Section III, he remarked:

By utility is meant that property in any object, whereby it tends to produce benefit, advantage, pleasure, good, or happiness, (all this in the present case comes to the same thing) or (what comes again to the same thing) to prevent the happening of mischief, pain, evil, or unhappiness to the party whose interest is considered...

Bentham’s influential disciple John Stuart Mill engrossed Bentham’s ideas and propagated them further. Other early advocates of utility were Auguste Walras and Juvénal Dupuit.

These beginnings of utility theory spawned later on the edifice of marginal utility propagated by Heinrich Hermann Gossen, William Stanley Jevons (mainly influenced by Bentham), Léon Walras, Carl Menger, Friedrich von Wieser, Eugen von Böhm-Bawerk, and John Bates Clark.\(^{21}\) For all those scholars utility was something palpable and immutable. It was only Pareto who was the first to aver interpersonal noncomparability of utility. It was not until the Hicks and Allen (1934) ordinal revolution that the profession had understood that uniqueness of the utility function up to an ordinal transformation is sufficient to analyze and rationalize economic decision-making. It was not until von Neumann and Morgenstern’s (1947, Appendix) second edition of their *Theory of Games and Economic Behavior* that expected utility was axiomatized and systematically explored. Their utility function of expected utility is cardinal and unique up to a linear transformation (although Baumol (1951) had raised doubt about the cardinality of von Neumann-Morgenstern utility). This means that, if we consider D. Bernoulli’s and Cramer’s solutions of the Petersburg Paradox as paradigms of expected utility, then we come upon finite expected values under all feasible (linear) transformations. Utility functions which are unique only up to ordinal transformations seem to be alien to expected utility.

When considering Corollary 9, we can look at the reasons for Menger’s Super-Petersburg Paradox and its avoidance. These may be embodied in \(x_i, u(\cdot), \psi(\cdot)\).

We can immediately rule out \(x_i < B < \infty\) for all \(i = 1, 2, \ldots, \infty\), because this will pull the rug out from under the emergence of a Petersburg Paradox: the expected value of such a design is trivially finite under all nondecreasing functions \(u(\cdot)\) and \(\psi(\cdot)\) which are

\(^{21}\)For comprehensive surveys on the development of utility theory cf. Viner (1925a,b) and Stigler (1950a,b). Whereas Viner did not mention Bernoulli at all, Stigler (1950b, pp. 373-7) devoted slightly over four pages to Bernoulli’s utility function. However, amidst only theories of utility under certainty, this part looks rather alien in the Stigler survey.
barred from infinity for finite domain values. Hence, we have to stick to the assumption that $x_i \to \infty$ as $i \to \infty$.

Let us, therefore, focus on $u(\cdot)$ and assume a Bernoulli-Cramer case, i.e., $u(x_i) \to \infty$ as $x_i \to \infty$, and $\sum_{i=1}^{\infty} u(x_i)p(x_i) < \infty$. Recall from Corollary 4 that finiteness of the expected utility results because $u(x_i)$ increases to infinity relatively slower than the $p(x_i)$ decreases to zero. How can then finite expected utility of a Bernoulli-Cramer case become infinite? Two instances can provide for that. First, $u(\cdot)$ stays put, but the $x_i$'s are replaced by other winnings, say, by $x_i^*$, such that the $u(x_i^*)$'s now increase relatively faster to infinity than the $p(x_i)$'s (which are assumed to be the same) decrease to zero. This is obviously the case of Menger’s Super-Petersburg Paradox, for which $u(x_i) = 2^{i-1}$ is replaced by $u(x_i^*) = \exp 2^{i-1}$. This means that the very same persons with the same utility functions fall victim to the Petersburg Paradox when the winnings are adequately increased. Such instances are ruled out once and for all if $u(\cdot)$ is bounded. No winnings, however enticing they are, can elevate the respective gambler’s utility to approach infinity. This is basically the gist of Menger’s theorem.

Alternatively, we can, as we did above, consider strictly increasing transformations $\psi(\cdot)$ of a utility function. Suppose the expected value of the $u(x_i^*)$'s is finite, but the $\psi[u(x_i^*)]$’s now increase relatively faster to infinity than the $p(x_i)$’s decrease to zero. In this view, $\psi(\cdot)$ is interpreted as the utility functions of some other gamblers who exhibit more greed such that the finite expected utility for some gamblers becomes infinite for others. This view takes account of different utility functions for different gambler personalities. Note again that boundedness of $u(\cdot)$ rules out the occurrence of a Petersburg Paradox in general.

Such increasing transformations may also work the other way round. Suppose the expected value of the $u(x_i^*)$’s is infinite. Consider a strictly increasing transformation such that the $\psi[u(x_i^*)]$’s increase relatively slower to infinity than the $p(x_i)$’s decrease to zero. This illustrates that a gamble with infinite expected value can be transformed into a gamble with finite expected value provided the gamblers are modest enough.

Let us now focus on $\psi(\cdot)$. It may be that, although $u(\cdot)$ is unbounded, $\psi(\cdot)$ is bounded. This may be consequential of some social conventions prevailing in a society, or may result from another contingency. It seems that Machina (1982, p. 310) alluded to this case as a solution of the Petersburg Paradox. This means that boundedness is shifted to a higher level.

22 An alternative formulation would be that the winnings “increase faster than the marginal utility of money diminishes.” (Cowen and High (1988, p. 220)).

23 In view of Arrow’s (1974) theorem discussed below, $\psi(\cdot)$ has to be convex to achieve that.
Recall from Remarks 6 and 7 that a Petersburg game can be “solved” or re-established by manipulating the probabilities rather than the winnings. When we use probabilities which shrink relatively faster to zero than the winnings or their utilities grow to infinity, then a Petersburg game may assume a finite expected value. On the other hand, using probabilities which shrink relatively slower to zero than the winnings or their utility grow to infinity, then a finite expected value of a Petersburg game can be transformed in a Petersburg game with infinite expected value. However, this works only if we can find probabilities which sum up to one, which cannot be always secured. Remark 6 showed us that this does not work for the Bernoulli solution of the Petersburg Paradox.

The Petersburg Paradox holds, of course, independently of the monetary unit of the winnings as long as it is greater than zero. Interestingly enough, Samuelson (1960) has demonstrated that, when the monetary unit approaches zero, while the number of tosses approaches infinity, and the utility function is twice continuously differentiable, then the relative stake of a Petersburg game approaches infinity, irrespective of whether the utility function is bounded or not. Let \( s = S(n, y; w) \) denote the fair stake, \( n \) the number of tosses, and \( y \) the value of the monetary unit, then \( s = S(\cdot) \) is implicitly defined by

\[
u(w) = \sum_{i=1}^{n} \frac{u(w-sy+2^{i-1})}{2^i} + \frac{u(w)}{2^n}.
\]

Samuelson (1960) proved that taking the double limit yields

\[
\lim_{n \to \infty} \lim_{y \to 0} \frac{S(n, y; w)}{y} = \lim_{y \to 0} \lim_{n \to \infty} \frac{S(n, y; w)}{y} = \infty.
\]

This example demonstrates that, when dealing with infinity, a different quality of solutions emerges. This is trivially illustrated by a single-toss lottery: the expected value of a lottery yielding 0 with probability of \( \frac{1}{2} \) and \( x \) with probability of \( \frac{1}{2} \) is obviously \( \frac{x}{2} \). But \( \lim_{x \to \infty} \frac{x}{2} = \infty \).

Finally, there is also a revealing theorem by Arrow (1974). It says that if \( u(x) \) is concave and monotone increasing with \( u(0) \) finite and \( E(x) \) finite, then expected utility is finite. Using Jensen’s inequality, the proof can be simplified: \( E[u(x)] \leq u[E(x)] \) holds for concave utility functions. Hence, if \( E(x) \) and \( u[E(x)] \) are finite, so is \( E[u(x)] \). Note that this theorem readily extends to concave and monotone increasing transformations \( \psi(\cdot) \) if \( E[u(x)] \) and \( \psi(E[u(x)]) \) are finite. Recall that finiteness of \( E(x) \) and \( u[E(x)] \) depends, inter alia, on the prevailing probability distribution. Hence, as shown in Remark 6, changing the probability distribution to engender an infinite value of \( Eu(x_i) \) or \( E\psi[u(x_i)] \) implies that Arrow’s theorem is not applicable. Remark 7 shows that it becomes applicable if the probability distribution is changed to yield a finite value of \( Eu(x_i) \) or \( E\psi[u(x_i)] \).

Whereas Menger’s theorem is impeccable from a mathematical point of view, the profession has largely failed to provide a behavioral rationale for the boundedness of utility. It seems that the many proposals brought forward to “solve” the Petersburg Paradox are nothing but disguised attempts of filling in the missing behavioral justification for the boundedness of utility functions. But rather than pegging them to the realm of behavioral
psychology, their proponents sought to offer them as stringent mathematical solutions. Hence, with this purpose in mind, they were doomed to failure.

4 Conclusion

The Petersburg Paradox is a special example of expected value. As such it is linear, either in winnings or in utilities. From a mathematical point of view it is indistinguishable whether the game is played with $u(x_i)$ as prizes and linearity in utility, or with $x_i^* := u(x_i)$ as winnings (possibly in a different monetary unit) and linearity in money. Switching from monetary winnings to utility is nothing but resorting to a behavioral assumption for explaining why subjects do not behave according to expected monetary value. But, from a mathematical point of view it is equivalent to replacing $u(x_i)$ by $x_i^*$. In both cases the “solution” consists of defining a different game with prizes of $u(x_i)$ or $x_i^*$ instead of $x_i$. The rationale for relying on the utilities of the winnings is the conviction that subjects perceive the winnings in a different way.

The second component of expected value is the probabilities. In many cases a Petersburg Paradox can be “solved” by re-defining the probabilities rather than the winnings. As the winnings or the utilities of the winnings go to infinity, we can define other probabilities which shrink relatively faster to zero than the winnings or their utilities grow to infinity. This instrument can also be used in the opposite direction. When the winnings or their utilities grow to infinity, but have a finite expected value, then the probabilities can be replaced by other probabilities which shrink relatively slower to zero so that a Petersburg Paradox emerges or re-emerges. Note that, as Remark 6 showed us, this does not work in all cases. It may be that, although ratio tests would suggest such a solution, the resulting “probabilities” lack their existential property, viz. that their sum be equal to one. A rationale for using adapted probabilities may, for instance, be that subjects perceive them in a different way.

Common to all approaches at a solution of the Petersburg Paradox is Menger’s quantum-jump insight that a necessary and sufficient condition of avoiding a Petersburg Paradox is boundedness of the winnings or of the admissible utility functions in their capacity as transformed winnings.
References


