Submitted to Econometrica

COMMENT ON D. BERNOULLI (1738)

Ole Peters

Daniel Bernoulli's study of 1738 Bernoulli [1738] is considered the beginning of expected utility theory. Here I point out that in spite of this, it is formally inconsistent with today's standard form of expected utility theory. Bernoulli's criterion for participating in a lottery, as written in Bernoulli [1738], is *not* the expected change in utility.

1. INTRODUCTION

I show that the decision theory developed by Bernoulli [1738] is not equivalent to the decision theory axiomatized by von Neumann and Morgenstern [1944].

Bernoulli's 1738 study was translated from Latin into English and republished in Econometrica in 1954, by Louise Sommer, helped by Karl Menger. It is regarded as the origin of expected utility theory. The original Latin paper is freely available online, and the problem I discuss here is not one of mistranslation. Expected utility theory is a cornerstone of formal economics. It is part of the standard university curriculum and features in many textbooks, often in its axiomatized form put forward by von Neumann and Morgenstern [1944]. While its descriptive shortcomings and its somewhat circular logical structure have been criticized, leading to the development of so-called "non-expected utility theories," it has at least the status of a very influential null model.

Both because of its influential status and because of the controversies surrounding it, it is important to clarify that Bernoulli's original study is inconsistent with what is today regarded as the standard form of expected utility theory, *i.e.* with what von Neumann and Morgenstern axiomatized, and indeed with most modern re-tellings of Bernoulli's work. This constitutes the main result of the present comment. Over the centuries, including some 200 years with little access to the original text, Bernoulli's study was digested by the scientific community, and what was perceived as errors or simply overlooked was changed in popular retellings, including those by Laplace [1814] and Todhunter [1865].

It is prudent to keep speculation about the consequences of this situation to a minimum. According to the logical principle "ex falso quodlibet" it is possible to prove mathematically any statement (quodlibet) – true or false – from the wrong assumption (falso) that [Bernoulli, 1738] and *e.g.* [von Neumann and Morgenstern, 1944] are representations of the same mathematical structure. As an illustration, I use this wrong assumption in Appendix A to prove that Bertrand Russell is the Pope.

It is certainly unusual for a mathematical inconsistency, or error, to remain in an important place of the scientific literature without detection for several

London Mathematical Laboratory, 8 Margravine Gardens, London W6 8RH, UK
Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, 87501 NM, USA
E-mail: o.peters@lml.org.uk

centuries. It is also unusual for an eminent scientist like Daniel Bernoulli to commit what later researchers implicitly interpreted as a mathematical error in the first place. The former anomaly is partially explained by the original Latin work being quickly replaced by greatly shortened – corrected – re-tellings in living languages.

To understand the latter, it is helpful to recall the science-historical context of Bernoulli's work. In the second half of the 17th century it was generally believed that humans, when faced with the choice between two gambles of equal duration[1] will choose the one that possesses the larger expected wealth change. This earliest decision theory, circa from the 1650s, is well summarized by C. Huygens's statement "if any one should put 3 shillings in one hand without telling me which, and 7 in the other, and give me choice of either of them; I say, it is the same thing as if he should give me 5 shillings..." Huygens [1657]. We will refer to this as "Huygens's decision theory" (HDT).

In the early 18th century doubts were raised about the realism of this mathematical model of human behavior. Humans do *not* generally choose the gamble with the largest expected wealth change. In particular, this was demonstrated by Nicholas Bernoulli's St. Petersburg gamble [Montmort, 1713, p. 402], whose expected wealth change is divergent, *i.e.* does not exist. How should such a gamble be evaluated? The divergence leads HDT to predict that people would be willing to pay any finite price to participate in the gamble, but that was not observed.

Expected utility theory emerged out of ideas put forward by Cramer (cited in Bernoulli's 1738 paper) and Daniel Bernoulli, in response to these doubts about HDT. It is a means of accounting for the fact that people do not optimize expected wealth changes when evaluating gambles. We will distinguish between Bernoulli's decision theory (BDT) and what we mean today by "expected utility theory" (EUT). Although Bernoulli [1738] is considered the beginning of EUT, it does not actually contain EUT as it is understood today.

An alternative modern accounting for the failure of HDT, using concepts that had not been developed by Bernoulli's time, is presented in [Peters and Gell-Mann, 2016]. This modern perspective, based on ergodic theory, allows a deep understanding of Bernoulli's work and brought to light the inconsistency between BDT and EUT that I am about to discuss.

Readers familiar with the context may jump straight to Sec. 4, which contains the core result: an exposition of the inconsistency between today's expected utility theory and Bernoulli's work. Section 2 makes explicit what I mean by a random variable and an expectation value. In Sec. 3, both the earliest decision theory (expected-wealth maximization) and the currently dominant form of decision theory (expected utility theory) are defined for reference. In Sec. 5 possible reasons for and consequences of Bernoulli's inconsistency are discussed.

[1] I will disregard matters of time and discounting here, as they were not considered in Bernoulli's early work, which is the focus of the present study. My perspective on these problems is elaborated in Peters and Gell-Mann [2016].

2. RANDOM VARIABLE AND EXPECTATION VALUE

For the discussion, we will need two concepts. Different authors favor slightly different nomenclatures, so for the avoidance of doubt I will define explicitly the terms used in this comment. The required concepts are: (i) a discrete random variable, and (ii) the expectation value.

Discrete random variable

A discrete random variable is an object Y, defined by

- a set of possible values, $\{y_i\}$, labelled with a number i, often called an "event."
- a set of probabilities, $\{P_Y(y_i) = p_i\}$, associated with these values.

The probabilities are non-negative, $p_i \geq 0$ and add up to one,

$$\sum_i p_i = 1. \tag{1}$$

A function $z(y)$ defines another random variable, Z, whose set of possible values are the values of the function $z_i = z(y_i)$ at the possible values y_i of the original random variable, Y. Its probabilities are also those of the corresponding events, $P_Z(z_i) = p_i$.

Expectation value

The expectation value of a random variable, $\langle y \rangle_{P_Y}$, is the sum over the products of all possible values and their probabilities,

$$\langle y \rangle_{P_Y} = \sum_i y_i P_Y(y_i). \tag{2}$$

The operator $\langle \cdot \rangle$ is a simple (linear) sum. As a consequence, if $z(y)$ is a *linear* function of y, then the expectation values $\langle z \rangle_{P_Z}$ and $\langle y \rangle_{P_Y}$ are trivially related, namely $\langle z \rangle_{P_Z} = z\left(\langle y \rangle_{P_Y}\right)$. This is easily proved, see Appendix B. If Huygens's statement had been restricted to linear $z(y)$, it would be closer to the truth: in that case, it often really is "the same thing" to consider the full distribution of y as it is to work with its expectation value. Crucially, however without linearity, (Eq. 8), the trivial relationship does not hold. In other words, $\langle z \rangle_{P_Z} = z\left(\langle y \rangle_{P_Y}\right)$ implies linearity of $z(y)$. This is also easily proved, see Appendix C.

Therefore, Huygens's statement is both false and misleading: even if we're only interested in expectation values, working with the random variable Y is not "the same thing" as working with its expectation value $\langle Y \rangle$. However, because many functions $z(y)$ are linearizable for small changes in their arguments, researchers in the early days of probability theory were not fully aware of this complication.

3. EARLY AND RECENT DECISION THEORIES

We model wealth as a dollar amount, x, and a gamble as a random variable, ΔX, whose possible values, $\{\Delta x_i\}$, represent the possible changes in wealth resulting from participating in the gamble.

For example, consider a coin toss: we have to pay a \$1 fee to participate, if the coin lands tails we receive nothing, if it lands heads we receive \$3. This would be modelled as a random variable with two discrete possible values, $\Delta x_1 = -\$1$ (net change in wealth if the coin shows tails, we receive nothing but have paid the fee) and $\Delta x_2 = +\$2$ (net change in wealth if the coin shows heads). The probabilities would be $P_{\Delta X}(\Delta x_1) = P_{\Delta X}(\Delta x_2) = 0.5$.

3.1. *Huygens's decision theory*

Imagine now deciding whether to participate in a gamble. This amounts to a comparison between no change in wealth (non-participation, which is a trivial gamble, sometimes called the null-gamble) and some evaluation of the random variable ΔX. The evaluation of ΔX in HDT was $\langle \Delta x \rangle$: the gamble is evaluated by comparing $\langle \Delta x \rangle$ to 0. If $\langle \Delta x \rangle > 0$ (like in the example, where $\Delta X = +\$0.50$), this decision theory predicts that people will accept the gamble, if $\langle \Delta x \rangle < 0$ the prediction is non-participation.

Following discussions in the early 18th century [Montmort, 1713, p. 402], it became clear that this decision theory is not a realistic model of human behavior, and expected utility theory was developed as a better fit to the observations.

3.2. *Expected utility theory*

We will come to Bernoulli's theory in the next section. First, I present today's standard EUT, almost exclusively in use at least since Laplace's 1814 book on then-nascent probability theory [Laplace, 1814] and famously axiomatized by von Neumann and Morgenstern [1944].

Instead of maximizing $\langle \Delta x \rangle$ in the choice of gambles, EUT posits that humans in their decisions appear to maximize the expected change of a *different* random variable. Wealth is mapped into utility of wealth $u(x)$, and human decisions are now modeled as maximizing the expected change in utility,

$$\langle \Delta u(x) \rangle = \langle u(x + \Delta x) \rangle - u(x). \tag{3}$$

This is a different mathematical model of human behavior. Crucially, EUT is sensitive to a reference level: wealth before the gamble, which we call x_0, unless the utility function is linear (which is not an interesting case because it predicts the same human behavior as maximizing $\langle \Delta x \rangle$), proved in Appendix D.

EUT thus introduces into decision theory the intuitively plausible notion that our ability and willingness to risk losing resources depends on the resources we have. A millionaire may not notice the loss of \$1,000, whereas for someone else this may mean starvation.

<ins>Summary:</ins> EUT, as presented in countless textbooks, including [Laplace, 1814, von Neumann and Morgenstern, 1944], models human decision-making by mapping wealth to utility $u(x)$. It then posits that decisions are made by selecting the gamble with the maximum expected change in utility.

Object	Bernoulli's notation	Our notation
Wealth before the gamble	AB	x_0
Utility before the gamble	0 (zero)	$u(x_0)$
Maximum ticket fee in Bernoulli's theory	pB	$f_{\mathrm{m}}^{\mathrm{B}}$
Maximum ticket fee in EUT	no symbol	$f_{\mathrm{m}}^{\mathrm{U}}$
Minimum prize received	BC	π_1
2nd-smallest prize	BD	π_2
3rd-smallest prize	BE	π_3
4th-smallest prize	BF	π_4
Utility change due to smallest prize at zero fee	CG	Δu_1^+
Utility change due to 2nd-smallest prize at zero fee	DH	Δu_2^+
Utility change due to 3rd-smallest prize at zero fee	EL	Δu_3^+
Utility change due to 4th-smallest prize at zero fee	FM	Δu_4^+
Expected utility change at zero fee	PO=AN	$\langle \Delta u^+ \rangle$
Utility change due to loss of maximum fee	po=An	Δu^-
Wealth corresponding to utility $u(x_0) + \langle \Delta u^+ \rangle$	AP	x^+
Probability of π_1	$\frac{m}{m+n+p+q+\cdots}$	p_1
Probability of π_2	$\frac{n}{m+n+p+q+\cdots}$	p_2
Probability of π_3	$\frac{p}{m+n+p+q+\cdots}$	p_3
Probability of π_4	$\frac{q}{m+n+p+q+\cdots}$	p_4
Net wealth change from i^{th}-smallest prize	no symbol	Δx_i
Net utility change from i^{th}-smallest prize	no symbol	Δu_i
Expected net wealth change	no symbol	$\langle \Delta x \rangle$
Expected net utility change	no symbol	$\langle \Delta u \rangle$

TABLE I

Translation of Bernoulli's 1738 notation into modern notation. Key concepts, especially the expected net change in utility do not appear in Bernoulli's work.

4. BERNOULLI'S DECISION THEORY IS DIFFERENT

Laplace, in 1814, and many authors after him, present EUT as I have done above. Laplace ascribes this fully to Bernoulli and does not mention that Bernoulli actually wrote something else.

Bernoulli considers a gamble where a fee is paid, and prizes are received with different probabilities. He uses a cumbersome geometric notation which we replace with a more modern one, see Table I and Fig. 1.

Modern EUT (equivalent to Laplace) predicts that the maximum ticket fee to be paid for such a gamble is the amount of money at which the expected utility gain is zero. Substituting in (Eq. 3) for Bernoulli's gamble,

$$\langle \Delta u \rangle = 0 = \sum_i p_i u(x_0 + \pi_i - f_{\mathrm{m}}^{\mathrm{U}}) - u(x_0). \tag{4}$$

However, Bernoulli [1738, p. 26–27] writes: "If we wish, further, to know how large a stake the individual should be willing to venture on this risky proposition, our curve must be extended in the opposite direction in such a way that the abscissa Bp now represents a loss and the ordinate po represents the corresponding decline in utility. *Since in a fair game the disutility to be suffered by* *losing must be equal to the utility to be derived by winning,* we must assume that An = AN, or po = PO. Thus Bp will indicate the stake more than which persons who consider their own pecuniary status should not venture." [emphasis mine]

The language here suggests that Bernoulli is *not* computing the expected net change in utility, and the equations he writes confirm this. Bernoulli proceeds as follows:

- he computes the expected change in utility, *assuming that the ticket fee is zero.*
- next, he computes the loss in utility from paying the fee *but receiving none of the prizes.*
- he then finds the fee where these two quantities are equal.

Bernoulli's decision theory – in contrast to EUT – predicts that the maximum fee to be paid is $f_{\mathrm{m}}^{\mathrm{B}}$, defined by

$$0 = \left\langle \Delta u^{+} \right\rangle - \Delta u^{-} \tag{5}$$

$$= \sum_i [p_i u(x_0 + \pi_i) - u(x_0)] - [u(x_0) - u(x_0 - f_{\mathrm{m}}^{\mathrm{B}})]. \tag{6}$$

Generally, this expression cannot be brought into the form of (Eq. 4), wherefore **Bernoulli's decision theory is not expected-utility theory.** The difference is summarized in Table II.

PROPOSITION 4.1 *Bernoulli's 1738 decision theory is not expected utility theory.*

The proposition is proved in Appendix A: assuming that BDT and EUT are equivalent implies the contradiction 1=0.

Only a linear utility function guarantees the equality $f_{\mathrm{m}}^{\mathrm{B}} = f_{\mathrm{m}}^{\mathrm{U}}$. In other words, Bernoulli's decision theory, which is often wrongly presented as equivalent to EUT, is really only equivalent to EUT under the assumption of linear utility. But that is equivalent to Huygens's decision theory that we encountered in Sec. 1, and that was deemed an unrealistic model of human behavior.

Bernoulli's equation on his p. 26 is accompanied by a figure, reproduced in Fig. 1. The figure is not part of the original manuscript but was produced by Sommer and Menger. It is a clear illustration that Bernoulli's decision theory is not only inconsistent with EUT but indeed questionable in itself. Here is why: the maximum fee to be paid is represented in the figure by the length pB (f_{m}). This length is clearly less than the length BC (π_1) that represents the minimum prize to be received from playing the lottery[2]. The minimum net change in wealth, *i.e.* the worst possible result from playing the lottery, is thus $\pi_1 - f_{\mathrm{m}} > 0$, which is positive. This lottery represents a guaranteed win of at least $\pi_1 - f_{\mathrm{m}}$ and possibly much more. There is no good reason to reject it, and a theory that predicts people

[2]The types of lotteries considered by Bernoulli always result in receiving some prize, *i.e.* it is not possible to receive nothing in return for the fee. For example, in the famous St. Petersburg lottery (the focus of Bernoulli's paper), at least $1 is paid out (or one ducat in Bernoulli's currency).

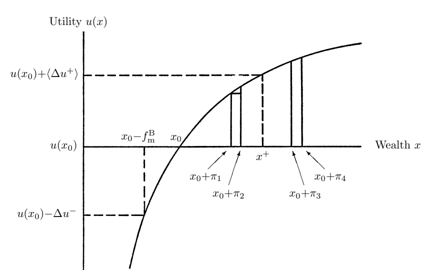
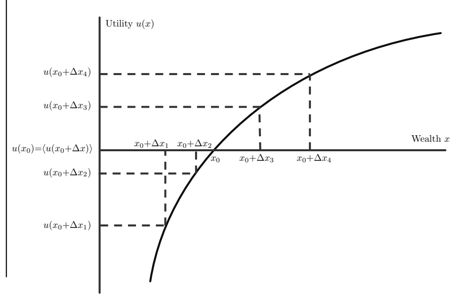
	Bernoulli's 1738 decision theory	**Expected-utility theory**
Algorithm	1. Find expected utility gain, $\langle \Delta u^+ \rangle$, assuming zero fee. 2. Find the fee, f_m^B, that would lead to a utility loss of equal size, Δu^-, if no prize were won.	Find the fee, f_m^U, at which the expected net change in utility is zero.
Equation	Choose f_m^B such that $\sum_i [p_i u(x_0 + \pi_i) - u(x_0)]$ $- [u(x_0) - u(x_0 - f_\mathrm{m}^\mathrm{B})]$ $= 0$.	Choose f_m^U such that $\sum_i p_i u(x_0 + \underbrace{\pi_i - f_\mathrm{m}^\mathrm{U}}_{\Delta x_i}) - u(x_0) = 0$
Figure		

TABLE II

Maximum fee to be paid for a lottery ticket, where prize π_i is won with probability p_i, wealth x, initial wealth x_0, and utility function $u(x)$. **The decision theory put forward by Bernoulli in 1738 is not expected utility theory (EUT).** Bernoulli did not compute the expected change in utility, whereas this is the object that determines the value of a lottery under EUT. Bernoulli's geometric construction does not correspond to EUT, and consequently the maximum fee computed by Bernoulli, f_m^B, differs from the maximum fee according to EUT, f_m^U. EUT works with net changes in wealth, $\Delta x_i = \pi_i - f$, whereas Bernoulli works with prizes, π_i, treating the fee, f, separately. The two theories constitute different mathematical models of human decision-making.

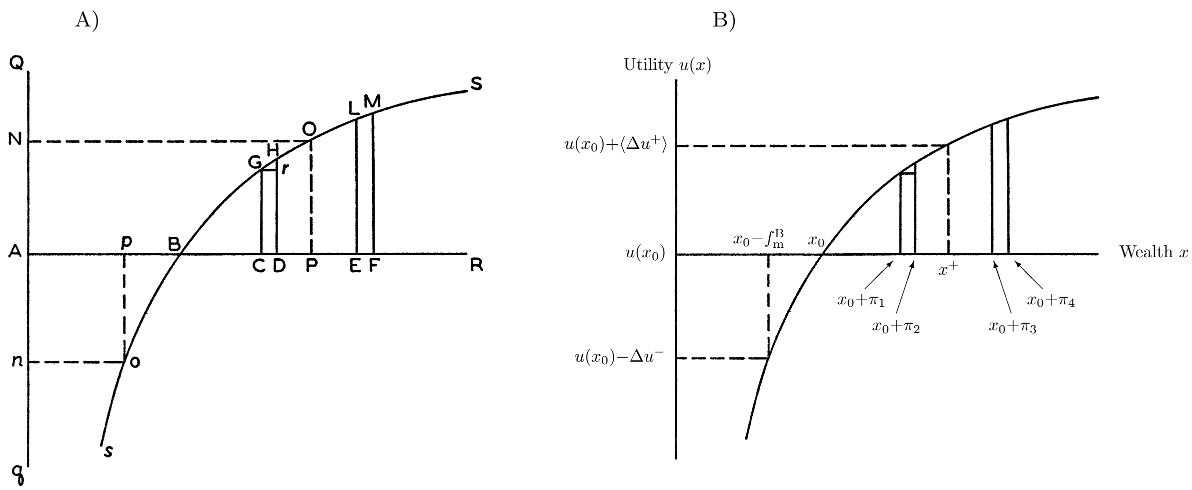


FIGURE 1.— A) Bernoulli's 1738 computation of the maximum fee to be paid for a lottery. B) the same figure translated into modern notation. In Bernoulli's decision theory, unlike in modern EUT, the maximum fee to be paid for a gamble is found as follows: find 1) the expected gain in utility, $\left\langle\Delta u^+\right\rangle$, that would arise from playing without paying the fee. Find 2) the fee such that the loss in utility that would arise from paying it without playing, Δu^- is the same as $\left\langle\Delta u^+\right\rangle$. This contradicts modern utility theory where the maximum fee is the fee that renders the expectation value of the change in utility due to net changes in wealth zero, $\langle\Delta u\rangle = 0$.

would reject it will soon be falsified by observation. Ask the person next to you if he'd like to play the following gamble: toss a coin, and if heads shows the person gets \$20, if it's tails he gets \$10.

5. DISCUSSION

It is impossible to reconcile BDT with EUT. We are, however, free to speculate about the reasons for Bernoulli's inelegant theory. It is possible that Bernoulli developed a different decision theory on purpose. He proposed a different model of human psychology, and perhaps he meant to do just that. Over the centuries, as we learned more about stochastic systems, we diverged from his original vision.

We may also suggest that Bernoulli made a sloppy error; in effect a typo. This is not a likely scenario, due to Bernoulli's consistent use of (Eq. 6) throughout his paper. A sloppy error would most likely occur in just one place.

A third interpretation is that the difference between BDT and EUT is a more serious confusion on Bernoulli's part about the early mathematical concept that is the expectation value. In this reading, Bernoulli actually wanted to compute $\langle\Delta u\rangle$ but didn't notice that his computation did not achieve this. This interpretation is supported by the following passage [Bernoulli, 1738, p. 24]: "Meanwhile, let us use this as a fundamental rule: If the utility of each possible profit expectation is multiplied by the number of ways in which it can occur, and we then divide the sum of these products by the total number of possible cases, a mean utility will be obtained, and the profit which corresponds to this utility will equal the value of the risk in question." This strongly suggests that Bernoulli wanted to compute $\langle \Delta u \rangle$, and indeed this is how Laplace interpreted it when he claimed that Bernoulli had proposed to compute $\langle \Delta u \rangle$. Laplace omitted the information that his claim implies that Bernoulli made an error. This scenario is not unlikely, given the state of probability theory at the time of Bernoulli's writing. But the absence of an equation at [Bernoulli, 1738, p. 24] leaves the statement open to interpretation. The following interpretations would make this statement equivalent to modern EUT and contradict Bernoulli's BDT:

- Bernoulli meant "net profit" when he wrote "profit," including losses as negative profits.
- He meant to restrict the statement to risks that require no fee to participate.

On the other hand, another plausible interpretation would keep Bernoulli's statement here consistent with the rest of his paper but inconsistent with modern EUT:

- By "profit" Bernoulli means "prize," disregarding the fee.

Laplace presents modern EUT that computes $\langle \Delta u \rangle$ as a criterion for gamble participation and ascribes it to Bernoulli, omitting the fact that Bernoulli did not actually compute this object. If the inconsistency between Bernoulli and the far more natural modern form of EUT seems astonishing, we should remind ourselves of how early in the development of probability theory, mathematics, and of course economics Bernoulli's study sits. Probability theory was in its infancy (the axioms of Kolmogorov [1933] are often considered the foundation of modern probability theory). The key concept Bernoulli stumbles over was in the process of being developed in the 18th century. This concept is the following: a function of a random value, such as $u(x)$, defines a new random variable (see Sec. 2) whose expectation value is not trivially related to the expectation value of X, namely $\langle u(x) \rangle \neq u(\langle x \rangle)$, and $\langle \Delta u(x) \rangle \neq u(\langle \Delta x \rangle)$ *etc.* This was certainly not universally known in Bernoulli's days. In fact, this complication is behind Bertrand's circle problem [Bertrand, 1889], put forward as late as 1889 in response to Laplace's flawed "principle of insufficient reason" of 1812, which it exposes as invalid, see the discussion in [van Kampen, 2007, p. 20].

It is my impression that the inconsistency between Bernoulli's decision theory and modern EUT has caused a great deal of confusion. For instance, Karl Menger's famous result that utility functions must be bounded [Menger, 1934] makes use of Bernoulli's original work and falls apart when one tries to derive it from modern EUT [Peters, 2011, Peters and Gell-Mann, 2016]. Campbell [2017, p. 5] notes that this result is now "routinely ignored", yet it persists in the literature without a retraction of, or comment on, the original paper. Similarly, discussions of prospect theory and cumulative prospect theory often confusingly claim that EUT does not contain a reference value. Indeed the first two equations in [Kahneman and Tversky, 1979, pp. 263–264] suggest that the authors misrepresent EUT as an evaluation of (Eq. 3) without knowledge of initial wealth x_0. The reference value in EUT is, of course, present wealth, x_0. The confusion may be related to the problem under investigation: if we insist that Bernoulli's and modern EUT are equivalent, we are forced to assume linear utility, and in that special case the reference value, x_0, cancels out from the computation of the decision criterion $\langle \Delta u \rangle$ in (Eq. 3).

A clear understanding of Bernoulli's work and its deviations from modern EUT prevents this sort of confusion. Moreover, it leads to an interpretation of EUT in terms of ergodicity transformations that has proved very useful recently, shedding new light on long-standing problems in economic theory, including the St. Petersburg paradox, the cooperation conundrum, and the equity premium puzzle.

REFERENCES

Daniel Bernoulli. Specimen Theoriae Novae de Mensura Sortis. Translation "Exposition of a new theory on the measurement of risk" by L. Sommer (1954). *Econometrica*, 22(1):23–36, 1738. . URL http://www.jstor.org/stable/1909829.

J. Bertrand. *Calcul des probabilités*. Gauthier-Villars, Paris, 1889.

John Y. Campbell. *Financial decisions and markets, A course in asset pricing*. Princeton University Press, 2017.

Christiaan Huygens. *De ratiociniis in ludo aleae. (On reckoning at Games of Chance)*. T. Woodward, London, 1657.

Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–91, 1979.

Andrey Nikolaevich Kolmogorov. *Grundbegriffe der Wahrscheinlichkeitsrechnung. Translation "Foundations of the Theory of Probability" by N. Morrison, New York: Chelsea, 1956 (2nd ed.)*. Springer, Berlin, 1933.

Pierre Simon Laplace. *Théorie analytique des probabilités*. Paris, Ve. Courcier, 2 edition, 1814.

Karl Menger. Das Unsicherheitsmoment in der Wertlehre. *J. Econ.*, 5(4):459–485, 1934. . URL http://dx.doi.org/10.1007/BF01311578.

Pierre Rémond Montmort. *Essay d'analyse sur les jeux de hazard*. Jacque Quillau, Paris. Reprinted by the American Mathematical Society, 2006, 2 edition, 1713.

Ole Peters. Menger 1934 revisited. *http://arxiv.org/abs/1110.1578*, 2011. URL http://arxiv.org/abs/1110.1578.

Ole Peters and Murray Gell-Mann. Evaluating gambles using dynamics. *Chaos*, 26:23103, February 2016. . URL http://dx.doi.org/10.1063/1.4940236.

I. Todhunter. *A history of the mathematical theory of probability*. Macmillan & Co., 1865.

Nicholas G. van Kampen. *Stochastic Processes in Physics and Chemistry*. Elsevier (Amsterdam), 3rd edition, 2007.

John von Neumann and Oskar Morgenstern. *Theory of games and economic behavior*. Princeton University Press, 1944.

Appendix A: Ex falso quodlibet

It is fascinating to speculate what past, present, and future difficulties in economic theory may be caused by the inconsistency between BDT and EUT. It is, however, far more important to point out that any statement, true or false, may be proved by wrongly assuming consistency. It is impossible to place a bound on the potential consequences of not clarifying this inconsistency. To illustrate, I prove that Bertrand Russell is the Pope.

Bertrand Russell, it is said, once lectured on the logical principle "ex falso quodlibet" – assuming the truth of any false proposition, any statement, irrespective of its truth or falsity, may be proved. A student raised his hand and challenged: "in that case, prove that you are the Pope based on the assumption that $1 = 0$."

THEOREM .1 $1 = 0$ *implies that Bertrand Russel is the Pope.*

PROOF: The set containing the Pope and Bertrand Russell has 2 members. From the false assumption $1 = 0$, it follows that $2 = 1$ (by adding 1). Therefore, the set only has 1 member, wherefore Bertrand Russell is the Pope. *Q.E.D.*

By extension, any statement that implies $1 = 0$ also implies that Bertrand Russell is the Pope. To show that the equivalence of BDT [Bernoulli, 1738] and EUT [von Neumann and Morgenstern, 1944] implies that Bertrand Russell is the Pope, it is therefore sufficient to prove that it implies $1 = 0$.

LEMMA .2 *The assumption that BDT and EUT are equivalent implies that Bertrand Russell is the Pope.*

PROOF: The assumption that BDT and EUT are equivalent formally means that the corresponding decision criteria yield identical values for any given gamble,

$$\underbrace{\left\langle \Delta u^+ \right\rangle - \Delta u^-}_{\text{BDT}} = \underbrace{\langle \Delta u \rangle}_{\text{EUT}} \tag{7}$$

With parameters is $F = \$1$, $G_1 = \$\frac{e+1+\sqrt{e^2+2e-3}}{2}$, $p_1 = 1$, and $x_0 = \frac{G-f}{e-1}$, the left-hand side side of (Eq. 7) is 0 and the right-hand side 1. By Theorem 1, it follows that Bertrand Russell is the Pope. *Q.E.D.*

Appendix B: linear functions imply trivial expectation value

PROOF:

$$\text{Assume linearity,} \quad z(y) = ay + b. \tag{8}$$

$$\text{Then} \quad \langle z \rangle_{P_Z} = \sum_i p_i(ay_i + b) \tag{9}$$

$$= a\left(\sum_i p_i y_i\right) + b \tag{10}$$

$$= z\left(\langle y \rangle_{P_Y}\right) \tag{11}$$

Q.E.D.

Appendix C: Trivial expectation value implies linear function

PROOF: Assume the trivial relationship

$$\langle z \rangle_{P_Z} = z\left(\langle y \rangle_{P_Y}\right) \tag{12}$$

$$\sum_i p_i z(y_i) = z\left(\sum_i p_i(y_i)\right) \tag{13}$$

Since the equation has to hold for general y_i, it has to hold term by term, and we have

$$p_i z(y_i) = z(p_i y_i) \tag{14}$$

Differentiating with respect to y_i yields

$$p_i z'(y_i) = p_i z'(p_i y_i) \tag{15}$$

$$z'(y_i) = z'(p_i y_i) \tag{16}$$

Again, since the equation has to hold for any y_i it implies a constant first derivative of $z(y)$, which means $z(y)$ is linear. *Q.E.D.*

Appendix D: insensitivity to x_0 implies linear utility

PROPOSITION .3 *If utility changes are a function of wealth changes only, then utility is a linear function of wealth.*

PROOF: To see this, require that Δu be a function of Δx only, $u(x + \Delta x) - u(x) = f(\Delta x)$. Now imagine fixing Δx and varying x: the equation then implies that changes in the function in response to a change in the argument are constant, $\frac{du}{dx} =$ const., whose solution is a linear function $u(x) = ax + b$. *Q.E.D.*