E

Expected Utility Theory

Joshua C. Teitelbaum
Law Center and Department of Economics,
Georgetown University, Washington, DC, USA

Synonyms

EUT; Expected utility hypothesis; Expected utility model

Definition

Expected utility theory is the dominant model of decision-making under uncertainty in law and economics. It posits that people choose among risky prospects, or lotteries, modeled as probability distributions over a set of possible outcomes, as if they assign a utility value to each outcome x according to a function $u(x)$ and select the lottery that maximizes the expected value of $u(x)$. This entry surveys expected utility theory and its development. It covers the St. Petersburg Paradox and Bernoulli's seminal contribution; the axiomatic formulations of expected utility theory by von Neumann and Morgenstern, Savage, and Anscombe and Aumann; and the modeling of risk and risk aversion in expected utility theory. It then discusses the challenges to expected utility theory posed by the Allais and Ellsberg paradoxes and the Rabin critique and points to non-expected utility theories that respond to these challenges.

The St. Petersburg Paradox

In the early eighteenth century, Daniel Bernoulli (1954 [1738]) introduced the twin ideas of expected utility and diminishing marginal utility to resolve a problem submitted by his cousin, Nicolas Bernoulli, to the mathematician Pierre Rémond de Montmort in 1713. The problem, known as the St. Petersburg Paradox, runs as follows: Peter tosses a fair coin until it lands on heads. He offers to give Paul two dollars if the coin lands on heads in the first toss, four dollars if it lands on heads in the second toss, eight dollars if it lands on heads in the third toss, 16 dollars if it lands on heads in the fourth toss, and so on, so that with each subsequent toss the payoff doubles. In short, Paul wins 2^n dollars if the coin lands on heads in the n th toss. What is the value to Paul of playing the game?

The "paradox" is that the value to Paul is finite, whereas the expected value of the game is infinite. In other words, if Paul valued risky prospects at their expected value, he would be willing to pay any price to play the game. But Paul is not willing to pay any price. Indeed, people generally are willing to pay a relatively small amount to play the game. As Ian Hacking surmised, "few of us would pay even $25 to enter such a game" (Hacking 1980).

A. Marciano, G. B. Ramello (eds.), *Encyclopedia of Law and Economics*,
https://doi.org/10.1007/978-1-4614-7883-6_805-1

To resolve the paradox, Bernoulli suggested that people do not value risky prospects at their expected value, but rather at their expected utility. That is, he suggested that, given (finite) initial wealth $w < \infty$ and entry price $p < w$, the value of the game to Paul was not $EV = \sum_{n=1}^{\infty} \left(\frac{1}{2^n}\right)(w + 2^n - p)$, but rather $EU = \sum_{n=1}^{\infty} \left(\frac{1}{2^n}\right) u(w + 2^n - p)$, where u is Paul's utility function, which captures how Paul values money. Moreover, Bernoulli suggested that Paul's utility function is concave – specifically, $u(x) = \log(x)$ – which captures the idea of diminishing marginal utility: "a gain of one thousand ducats is more significant to a pauper than to a rich man" (Bernoulli 1954 [1738]). Together, these suggestions "unravel the knot" (ibid.): although the expected value of the game is $EV = \left[\frac{w+2-p}{2} + \frac{w+4-p}{4} + \frac{w+8-p}{8} + \cdots\right]$, which is infinite, the expected utility of the game is $EU = \left[\frac{\log(w+2-p)}{2} + \frac{\log(w+4-p)}{4} + \frac{\log(w+8-p)}{8} + \cdots\right]$, which is finite. It follows that the maximum entry price $\overline{p}$ that Paul is willing to pay is given implicitly by $\sum_{n=1}^{\infty} \left(\frac{1}{2^n}\right) \log(w + 2^n - \overline{p}) = \log(w)$. For example, if Paul is a millionaire ($w = 1{,}000{,}000$), then he is willing to pay 20 dollars and change ($\overline{p} \cong 20.88$).

Although Bernoulli's theory provided a resolution to the St. Petersburg Paradox, it did not answer the "why" question: Why would Paul value risky prospects at their expected utility? That is, why would people choose among risky prospects so as to maximize expected utility?

Objective Expected Utility Theory

Two centuries after Bernoulli, mathematician John von Neumann and economist Oskar Morgenstern offered an answer to this question (von Neumann and Morgenstern 1953 [1947]). They proposed a set of axioms that, if satisfied by an individual's preferences over risky prospects, are logically equivalent to the individual valuing risky prospects according to their expected utility. Stated another way, they proposed a set of axioms on preferences over risky prospects that guarantee the existence of a utility function such that one risky prospect is preferred to another risky prospect if and only if the expected utility of the former risky prospect is greater than the expected utility of the latter risky prospect.

The original formulation by von Neumann and Morgenstern (vNM) has been described as "terse and somewhat enigmatic" (Fishburn 1988). For this reason, textbook treatments of the vNM theorem frequently rely on subsequent reformulations, including those by Hernstein and Milnor (1953), Luce and Raiffa (1957), Jensen (1967), and Fishburn (1970). I follow suit.

Let X denote an arbitrary set of outcomes and let $\Delta(X)$ denote the set of simple probability distributions on X. Elements of $\Delta(X)$ are known as lotteries. For any two lotteries $p, q \in \Delta(X)$, and any $\alpha \in [0, 1]$, define the compound or mixed lottery $\alpha p + (1 - \alpha)q \in P(X)$ by $(\alpha p + (1 - \alpha)q)(x) = \alpha p(x) + (1 - \alpha)q(x)$ for every outcome $x \in X$. People have preferences and make choices over lotteries. Let $\succcurlyeq$ denote an individual's preference relation on $\Delta(X)$. That is, for any $p, q \in \Delta(X)$, $p \succcurlyeq q$ means that the individual prefers lottery p to lottery q, in the sense that the individual either strictly prefers lottery p to lottery q (i.e., $p \succ q$) or is indifferent between lottery p and lottery q (i.e., $p \sim q$).

The vNM theorem relies on three axioms.

A1. Weak Order: $\succcurlyeq$ is complete and transitive.

A2. Continuity: For every $p, q, r \in \Delta(X)$ such that $p \succ q \succ r$, there exist $\alpha, \beta \in (0, 1)$ such that $\alpha p + (1 - \alpha)r \succ q \succ \beta p + (1 - \beta)r$.

A3. Independence: For every $p, q, r \in \Delta(X)$ and every $\alpha \in (0, 1)$, $p \succcurlyeq q$ if and only if $\alpha p + (1 - \alpha)r \succcurlyeq \alpha q + (1 - \alpha)r$.

The first axiom (weak order) requires that the individual's preference relation $\succcurlyeq$ is complete and transitive. The individual's preference relation $\succcurlyeq$ is complete if, for every $p, q \in \Delta(X)$, $p \succcurlyeq q$ or $q \succcurlyeq p$ or both. That is, completeness requires that, for every pair of lotteries, the individual either strictly prefers one to the other or is indifferent between them. The individual's preference relation $\succcurlyeq$ is transitive if, for every $p, q, r \in \Delta(X)$, $p \succcurlyeq q$ and $q \succcurlyeq r$ implies $p \succcurlyeq r$. That is, transitivity requires that if the individual prefers one lottery to a second lottery and prefers the second lottery to a third lottery, then the individual must prefer the first lottery to the third lottery. Transitivity ensures that the individual's preferences are acyclic.

The second axiom (continuity) requires that small changes in probabilities do not change the individual's preferences. Stated another way, continuity requires that no lottery is infinitely more or less desirable than another lottery. Given $p \succ q \succ r$, it rules out the possibility that the individual would prefer the lottery q for sure to a mixed lottery involving a near one chance of p and a near zero chance of r (which would imply that the lottery r is infinitely less desirable than the lottery q) or that the individual would prefer a mixed lottery involving a near zero chance of p and a near one chance of r to the lottery q for sure (which would imply that the lottery p is infinitely more desirable than the lottery q).

The third axiom (independence) requires that if two lotteries are each mixed with a third lottery, then the individual's preference between the resulting mixed lotteries is independent of the third lottery and depends only on the individual's preference between the original two lotteries. Stated another way, independence requires that the individual's preferences over mixed lotteries satisfy a form of separability: the individual compares mixed lotteries based on their distinct elements, disregarding their common elements.

The normative appeal of the independence axiom stems in part from the fact that it equates to a form of dynamic consistency. We can interpret $\alpha p + (1 - \alpha)r$ and $\alpha q + (1 - \alpha)r$ as two-stage lotteries in which nature first determines whether the individual receives the choice between p and q (with probability α) or alternatively receives r (with probability $1 - \alpha$). Provided that the individual's preferences obey consequentialism (the individual's preferences over lotteries depend only on future consequences and not on past history or counterfactuals) and reduction of compound lotteries (the individual is indifferent between a compound lottery and the corresponding reduced lottery), independence is violated if the individual plans to choose p over q in the event that she receives the choice (which implies $\alpha p + (1 - \alpha)r \succcurlyeq \alpha q + (1 - \alpha)r$) but then chooses q over p if and when she receives the choice (which implies $q \succ p$). See, e.g., Karni and Schmeidler (1991), Volij (1994), and Nebout (2014).

The vNM theorem states that an individual's preferences over lotteries satisfy A1, A2, and A3 if and only if there exists a utility function $u : X \to \mathbb{R}$ such that for every $p, q \in \Delta(X)$,

$$p \succcurlyeq q \text{ if and only if } \sum_{x \in X} p(x)u(x) \geq \sum_{x \in X} q(x)u(x).$$

That is, the vNM theorem provides that a person chooses among risky prospects so as to maximize expected utility if and only if his preferences over lotteries obey the vNM axioms.

While vNM provided an answer to the question of why (or, more precisely, when) a person would value risky prospects according to their expected utility, vNM's theory of expected utility presupposed the existence of objective probabilities. About a decade before vNM, Frank Ramsey (1931) and Bruno de Finetti (1964 [1937]) independently formulated theories of expected utility in which probabilities are subjective, i.e., in which probabilities represent degrees of belief in the likelihood of uncertain events. However, both Ramsey and de Finetti effectively presupposed the existence of utilities. Moreover, Ramsey only sketched a proof of the existence of subjective probabilities, and de Finetti only considered monetary bets (and thus formulated a theory of expected value, not expected utility, with subjective probabilities).

Subjective Expected Utility Theory

About a decade after vNM, the American mathematician Leonard Savage (1972 [1954]) provided a formulation of subjective expected utility theory, regarding as the "crowning glory" of decision theory (Kreps 1988), that synthesized the ideas of vNM, Ramsey, and de Finetti. Fishburn (1970) has called it the "most brilliant axiomatic theory of utility ever developed."

The primitives of Savage's formulation are a set of outcomes, denoted by X, and a set of states of the world, denoted by S. People have preferences and make choices over acts, which are functions from states to outcomes, $f : S \to X$. Let F denote the set of all acts and let $\succcurlyeq$ denote an individual's preference relation on F.

Savage proposed a set of seven axioms that, if satisfied by an individual's preferences over acts, guarantee the existence of a utility function $u : X \to \mathbb{R}$ and a subjective probability measure P on S such that for every $f, g \in F$,

$$f \succcurlyeq g \quad \text{if and only if} \quad \int_S u(f(s))dP(s) \geq \int_S u(g(s))dP(s).$$

In other words, Savage's axioms provided an answer to the question of why/when a person would choose among risky prospects so as to maximize his subjective expected utility.

At the core of Savage's theory of subjective expected utility is the sure-thing principle, which plays a role analogous to that of the independence axiom in vNM's theory of objective expected utility. Essentially, the sure-thing principle requires that an individual has conditional preferences, e.g., he conditionally prefers f to g given an event $A \subset S$, and that if he conditionally prefers f to g both given A and given $A^C = S \backslash A$, then he unconditionally prefers f to g. Like the independence axiom, the sure-thing principle imposes a form of separability on preferences: the sure-thing principle implies that if f and g disagree (yield different outcomes) on A but agree (yield the same outcomes) on A^C, then the individual's conditional preference between f and g given A determines his unconditional preference, and therefore his choice, between f and g. Stated another way, the individual compares acts by focusing on events on which they disagree, and disregarding events on which they agree. And like the independence axiom, the appeal of the sure-thing principle stems in part from the fact that it equates to a form of dynamic consistency (provided that preferences satisfy a form of consequentialism). See Ghirardato (2002); see also Arrow (1971, ch. 2), Myerson (1979), and Hammond (1988).

About a decade after Savage, Frank Anscombe and Robert Aumann (1963) provided an alternative formulation of subjective expected utility theory in which subjective probabilities are defined in terms of objective probabilities. Like Savage, Anscombe and Aumann (AA) began with a set of outcomes X and a set of states S. However, they further posited the existence of objective randomization devices (e.g., roulette wheels), "with stochastic independence between successive observations, and *with stated values* for the chances of simple outcomes" (Anscombe and Aumann 1963). These give rise to objective, simple lotteries over X, as in vNM. People have preferences and make choices over acts, which AA redefined as functions from states to lotteries over outcomes, $h : S \to \Delta(X)$. Let H denote the set of all acts and let $\succcurlyeq$ denote an individual's preference relation on H. Once the individual chooses an act $h \in H$, nature determines the state $s \in S$ (resolving subjective uncertainty), yielding the lottery $p = h(s) \in \Delta(X)$, the outcome $x \in X$ of which is determined by a randomization device calibrated to p (resolving objective uncertainty).

AA's formulation entails five axioms – the three vNM axioms plus two additional axioms, each applied to preferences over AA acts, which are de facto compound lotteries. The key additional axiom is monotonicity, which requires that, for every $h, h' \in H$, if $h(s) \succcurlyeq h'(s)$ for all $s \in S$ then $h \succcurlyeq h'$. Monotonicity's appeal stems in part from the fact that it equates to a form of state independence. Denote by h_s^p the act that yields lottery p in state s and yields lottery r in all other states of the world. Monotonicity is violated if $h_s^p \succcurlyeq h_s^q$ but $h_{s'}^q \succ h_{s'}^p$ for some lottery $q \neq p$ and state $s' \neq s$. In other words, monotonicity is violated if the individual's preference between two acts that yield different lotteries in one state of the world (and only one state of the world) depend on the state of the world in which they yield different lotteries.

The AA axioms guarantee the existence of a utility function $u : X \to \mathbb{R}$ and a subjective probability measure P on S such that for every h, $h' \in H$,

$$h \succcurlyeq h' \quad \text{if and only if} \quad \int_S \sum_{x \in X} h(s)(x)u(x)dP(s) \geq \int_S \sum_{x \in X} h'(s)(x)u(x)dP(s).$$

That is, the individual chooses h so as to maximize his subjective expectation (with respect to P) of the objective expectation (with respect to $p = h(s)$) of $u(x)$.

Risk and Risk Aversion in Expected Utility Theory

Returning to the St. Petersburg Paradox, recall that Bernoulli's explanation for Paul's aversion to risk – i.e., Paul's unwillingness to pay an actuarially fair price to play Peter's game – invokes not only the idea that Paul values risky prospects at their expected utility, but also the idea that Paul's utility function is concave. More than 200 years after Bernoulli, economist Kenneth Arrow (1971) and statistician John Pratt (1964) independently developed the theory of risk aversion within the expected utility framework.

Within the expected utility framework, an individual is risk averse – in the sense that, for any risky prospect, he prefers a sure amount equal to the expected value of the risky prospect to the risky prospect itself – if and only if his utility function is concave. Formally, let u denote the individual's utility function, let w denote his initial wealth, and let $\tilde{z}$, a random variable, denote his risky prospect. The individual is risk averse if and only if

$$u(w + E[\tilde{z}]) \geq E[u(w + \tilde{z})],$$

where E is the expectation operator. This inequality is known as Jensen's inequality, and it is a characteristic property of a concave function. Assuming u is increasing and twice differentiable, a necessary and sufficient condition for this inequality to hold is that $u''(w) \leq 0$ for all w, or equivalently that $u'(w)$ is decreasing as w increases, i.e., u exhibits diminishing marginal utility.

Two useful measures of an individual's degree of risk aversion are the Arrow-Pratt coefficients of absolute and relative risk aversion:

$$r_A(w) = -\frac{u''(w)}{u'(w)} \quad \text{and} \quad r_R(w) = -\frac{wu''(w)}{u'(w)}.$$

Conceptually, both $r_A(w)$ and $r_R(w)$ are measures of the concavity of u. That is, both measure the rate at which the individual's marginal utility diminishes. The former measures the rate at which marginal utility decreases when wealth increases by one unit, whereas the latter measures the rate at which marginal utility decreases when wealth increases by 1%.

The appeal of the Arrow-Pratt coefficients of risk aversion is due not only to their interpretation as measures of the rate of decay for marginal utility, but also to their kinship with two intuitive, behavioral measures of risk aversion.

The first measure is Pratt's risk premium $\pi(w, \tilde{z})$, defined implicitly by the equation

$$u(w + E[\tilde{z}] - \pi(w, \tilde{z})) = E[u(w + \tilde{z})].$$

The risk premium is the maximum amount, beyond the expected (or actuarial) value of the risk, $E[\tilde{z}]$, that an individual with wealth w would be willing to pay to insure against the risk $\tilde{z}$. The amount $c(w, \tilde{z}) = E[\tilde{z}] - \pi(w, \tilde{z})$ is known as the certainty (or cash) equivalent of the risk; it is the sure change in wealth that has the same effect on utility as bearing the risk $\tilde{z}$. Thus, the risk premium is the amount by which the risk's expected value exceeds its cash equivalent.

Pratt (1964) showed that, under suitable regularity conditions on u and for "small" risks, the risk premium associated with risk $\tilde{z}$ is approximately equal to one-half the variance of $\tilde{z}$ times the individual's degree of absolute risk aversion evaluated at expected wealth:

$$\pi(w,\tilde{z}) \approx \frac{1}{2} V(\tilde{z}) r_A(w + E[\tilde{z}]).$$

Similarly, he showed that if we measure the risk and the risk premium not in absolute terms but as proportions of initial wealth – let $\tilde{z}^* = \tilde{z}/w$ and $\pi^* = \pi/w$ – then the risk premium is approximately equal to one-half the variance of the risk times the individual's degree of relative risk aversion evaluated at expected wealth:

$$\pi^*(w,\tilde{z}^*) \approx \frac{1}{2} V(\tilde{z}^*) r_R(w + wE[\tilde{z}^*]).$$

Roughly, the Arrow-Pratt degree of risk aversion is twice the risk premium per unit of variance, at least for small risks. See Menezes and Hanson (1970).

The second measure is Arrow's probability premium $p(w,\tilde{z})$, defined implicitly by

$$u(w) = E[u(w+\tilde{z})] = \frac{1}{2}[1 + p(w,\tilde{z})]u(w+h) + \frac{1}{2}[1 - p(w,\tilde{z})]u(w-h)$$

for the special case $\tilde{z} = \pm h$, $h > 0$. In other words, the probability premium is the increment, above fair odds, that would make an individual with wealth w indifferent between the status quo and an equiprobable bet of $\pm h$.

Arrow (1971) showed that, under suitable regularity conditions on u and for "small" stakes h, the probability premium associated with risk $\tilde{z}$ is approximately equal to one-half the stakes times the individual's degree of absolute risk aversion evaluated at initial wealth:

$$p(w,\tilde{z}) \approx \frac{1}{2} h r_A(w).$$

Similarly, he showed that if we measure the bet not in absolute terms but in proportion to wealth – let $\tilde{z}^* = \pm h^* w$, $0 < h^* < 1$ – then the risk premium is approximately equal to one-half the stakes times the individual's degree of relative risk aversion evaluated at initial wealth:

$$p(w,\tilde{z}^*) \approx \frac{1}{2} h^* r_R(w).$$

Roughly, the Arrow-Pratt degree of risk aversion is twice the probability premium the individual requires per unit risked, at least for small risks.

Note that Arrow (1971) derived not the probability premium p but rather the probability $p^* = \frac{1}{2}[1 + p]$. The above derivation of p follows the derivation in Pratt (1964).

A few years after Arrow and Pratt independently developed the theory of risk aversion within the expected utility framework, Hadar and Russell (1969), Hanoch and Levy (1969), and (most famously) Rothschild and Stiglitz (1970) independently developed the theory of increasing risk within the expected utility framework. The key concepts of this theory are first- and second-order stochastic dominance.

Let $\tilde{z}_1$ and $\tilde{z}_2$ denote two risky prospects, with cumulative distribution functions F_1 and F_2, respectively. We say that $\tilde{z}_1$ first-order stochastically dominates $\tilde{z}_2$ if $F_1(z) \le F_2(z)$ for all z. In other words, $\tilde{z}_1$ first-order stochastically dominates $\tilde{z}_2$ if the latter is obtained from the former by a transfer of probability mass from higher payoffs to lower payoffs. Equivalently, $\tilde{z}_1$ first-order stochastically dominates $\tilde{z}_2$ if the latter is obtained by adding negative noise to the former; i.e., $\tilde{z}_2$ has the same distribution as $\tilde{z}_1 + \tilde{\varepsilon}$, where $\tilde{\varepsilon} \le 0$ in all events. The importance of first-order stochastic dominance within an expected utility framework is that if $\tilde{z}_1$ first-order stochastically dominates $\tilde{z}_2$, then every expected utility maximizer with an increasing utility function (i.e., who prefers more wealth to less) necessarily prefers $\tilde{z}_1$ to $\tilde{z}_2$.

Assume now that $\tilde{z}_1$ and $\tilde{z}_2$ have the same mean, i.e., $E[\tilde{z}_1] = E[\tilde{z}_1]$. We say that $\tilde{z}_1$ second-order stochastically dominates $\tilde{z}_2$ if $\int_{-\infty}^{z} F_1(t)dt \ge \int_{-\infty}^{z} F_2(t)dt$ for all z. Rothschild and Stiglitz (1970) showed that this integral condition, and therefore second-order stochastic dominance, is equivalent to two intuitive notions of increasing risk. First, they showed that it is equivalent to the notion that $\tilde{z}_2$ is a mean-preserving spread of $\tilde{z}_1$ (or, more precisely, that $\tilde{z}_2$ is obtained from $\tilde{z}_1$ by a sequence of mean-preserving spreads). Intuitively, a mean-preserving spread is a transfer of probability mass from the center to the tails that leaves the mean unchanged. Note that if $\tilde{z}_2$ is a mean-preserving spread of $\tilde{z}_1$, then $\tilde{z}_2$ has a higher variance than $\tilde{z}_1$. Second, they showed that it is equivalent to the idea that $\tilde{z}_2$ is obtained by adding white noise to $\tilde{z}_1$; i.e., $\tilde{z}_2$ has the same distribution as $\tilde{z}_1 + \tilde{\varepsilon}$, where $E(\tilde{\varepsilon}|\ \tilde{z}_1 = z) = 0$ for all z. Most importantly, Rothschild and Stiglitz showed that if $\tilde{z}_1$ second-order stochastically dominates $\tilde{z}_2$, then every expected utility maximizer with an increasing and concave utility function necessarily prefers $\tilde{z}_1$ to $\tilde{z}_2$. In other words, they showed that within the expected utility framework, if $\tilde{z}_2$ is riskier than $\tilde{z}_1$ in the sense of second-order stochastic dominance, then every risk averter (who prefers more wealth to less) prefers $\tilde{z}_1$ to $\tilde{z}_2$.

Challenges to Expected Utility Theory

Because of its mathematical elegance and normative appeal, as well as the explanatory power of many of its predictions, expected utility theory is the dominant model of individual decision-making under risk and uncertainty in economics and hence in law and economics. Nevertheless, there have been many challenges to expected utility theory as a descriptive theory. Two of the most famous and important challenges are the Allais and Ellsberg paradoxes.

The Allais paradox is due to French economist Maurice Allais (1979 [1953]). It comprises two thought experiments, both involving two hypothetical choice problems. In the first experiment, the two choice problems are

Problem 1: Choose between
Prospect A: 100% chance of winning 100 million
Prospect B: 10% chance of winning 500 million
89% change of winning 100 million
1% chance of winning nothing

Problem 2: Choose between
Prospect C: 11% chance of winning 100 million
89% chance of winning nothing
Prospect D: 10% chance of winning 500 million
90% chance of winning nothing

Allais postulated that most individuals would prefer A to B and D to C. However, this would contradict vNM's objective expected utility theory. Under that theory, $A \succ B$ if and only if

$$u(w + 100) > 0.10u(w + 500) + 0.89u(w + 100) + 0.01u(w),$$

and $D \succ C$ if and only if

$$0.10u(w + 500) + 0.90u(w) > 0.11u(w + 100) + 0.89u(w).$$

The second inequality, however, holds if and only if

$$0.10u(w + 500) + 0.89u(w + 100) + 0.01u(w) > u(w + 100),$$

which contradicts the first inequality. One can show that it is the independence axiom that is violated by the Allais paradox. Intuitively, this is because the equivalence of the second and third inequalities relies on the fact that expected utility is linear in the probabilities, and the independence axiom is responsible for this property of objective expected utility theory.

In the second thought experiment, the two choice problems are

Problem 1: Choose between
Prospect A: 100% chance of winning 100 million
Prospect B: 98% chance of winning 500 million
2% chance of winning nothing

Problem 2: Choose between

Prospect C: 1% chance of winning 100 million
99% chance of winning nothing

Prospect D: 0.98% chance of winning 500 million
99.02% chance of winning nothing

Again, Allais postulated that most individuals would prefer A to B and D to C. Again, however, this would contradict objective expected utility theory by violating the independence axiom. The set of outcomes is $\{500, 100, 0\}$. Let a, b, c, and d denote the lotteries associated with A, B, C, and D, respectively, and let $e = (0, 0, 1)$. Observe that $c = .01a + .99e$ and $d = .01b + .99e$. Thus, the combination $a \succ b$ and $d \succ c$ evidently contravenes the independence axiom.

The Ellsberg paradox is due to the American economist Daniel Ellsberg (1961). (Ellsberg is famous outside of economics for releasing the Pentagon Papers to the New York Times in 1971.) Like the Allais paradox, the Ellsberg paradox comprises two thought experiments, both involving two hypothetical choice problems. In the first experiment, there are two urns. Urn I contains 100 red and black balls in an unknown ratio. Urn II is known to contain 50 red balls and 50 black balls. The two choice problems are

Problem 1: Choose between

Gamble A: Draw a ball from urn II and win \$100 if a red ball is drawn.

Gamble B: Draw a ball from urn I and win \$100 if a red ball is drawn.

Problem 2: Choose between

Gamble C: Draw a ball from urn I and win \$100 if a black ball is drawn.

Gamble D: Draw a ball from urn II and win \$100 if a black ball is drawn.

Ellsberg postulated that most individuals would prefer A to B and D to C. However, this would contradict Savage's subjective expected utility theory, for it would imply that the subjective probability of drawing a red ball from urn I is less than one-half in problem 1 and greater than one-half in problem 2. As Ellsberg observed, this is "inconsistent with the essential properties of probability relationships," and hence we "must conclude that [these] choices are not revealing judgments of 'probability' at all."

In the second experiment, there is one urn. It is known to contain 30 red balls plus 60 black and yellow balls in an unknown ratio. One ball is to be drawn at random from the urn. The two choice problems are

Problem 1: Choose between

Gamble A: Bet on "red" and win \$100 if a red ball is drawn.

Gamble B: Bet on "black" and win \$100 if a black ball is drawn.

Problem 2: Choose between

Gamble C: Bet on "red or yellow" and win \$100 if a red or yellow ball is drawn.

Gamble D: Bet on "black or yellow" and win \$100 if a black or yellow ball is drawn.

Again, Ellsberg postulated that most individuals would prefer A to B and D to C. Again, however, this would contravene subjective expected utility theory – specifically, the sure-thing principle. Let a, b, c, and d denote the acts associated with A, B, C, and D, respectively. The set of outcomes is $\{100, 0\}$ and the set of states is {red, black, yellow}. Let E denote the event "red or black" (i.e., "not yellow") and observe that c and d disagree on E but agree on E^C (i.e., "yellow"). Observe further that c versus d given E is tantamount to a versus b – both amount to "bet on red" versus "bet on black." By the sure-thing principle, therefore, $d \succ c$ requires $b \succ a$.

A third important challenge to expected utility theory is the so-called Rabin critique (Rabin 2000; see also Rabin and Thaler 2001). Rabin argued that expected utility theory cannot provide a plausible account of risk aversion over both small-stakes and large-stakes gambles. His starting point is the observation that under expected utility theory, risk aversion derives from a concave utility function over wealth, which captures diminishing marginal utility for money. He then demonstrated that if a person exhibits plausible risk aversion over small-stakes gambles, the implied rate of diminishing marginal utility (i.e., the implied degree of concavity) yields implausible risk aversion over large-stakes gambles.

Consider the following example. Suppose a person with initial wealth of \$200 and CARA utility function $u(x) = 1 - e^{-rx}$ rejects a gamble to lose \$20 with probability 48% or win \$20 with probability 52%, which seems plausible. (CARA stands for constant absolute risk aversion.) Under expected utility theory, this implies

$$0.48\left(1 - e^{-(200-20)r}\right) + 0.52\left(1 - e^{-(200+20)r}\right) \leq 1 - e^{-200r},$$

which in turn implies $r \geq \frac{1}{20} \log \frac{13}{12}$. Now consider a gamble to lose \$190 with probability 48% or win an infinite amount with probability 52%. Because $r \geq \frac{1}{20} \log \frac{13}{12}$, the expected utility of this gamble is less than the expected utility of the status quo. To see this, observe that

$$0.48\left(1 - e^{-(200-190)r}\right) + 0.52\left(1 - e^{-(200+\infty)r}\right) < 1 - e^{-200r}$$

if and only if $r > \frac{1}{190} \log \frac{25}{12}$, which is true because $r \geq \frac{1}{20} \log \frac{13}{12} > \frac{1}{190} \log \frac{25}{12}$. Thus, expected utility theory dictates that the person must also reject this gamble, which seems implausible.

Rabin (2000) shows that the problem is more general. His "calibration" theorem implies that for any concave utility function, if an expected utility maximizer rejects small gambles for a range of initial wealth, then he must also reject large gambles for wealth levels in that range – so large as to imply implausible degrees of risk aversion. For instance, if an expected utility maximizer rejects a 50–50 gamble to lose \$100 or win \$110 for all levels of initial wealth up to \$300,000, then he will also reject the following 50–50 gambles given an initial wealth of \$290,000: lose \$1000 or win \$718,190; lose \$2000 or win \$12 million; lose \$6000 or win \$180 million; lose \$10,000 or win \$1.3 billion. If the win amount in the initial gamble is \$125, then he will reject the following 50–50 gambles: lose \$1000 or win \$160 billion; lose \$6000 or win \$89 trillion; lose \$10,000 or win \$7.7 quadrillion. And if he rejects the initial gamble for all levels of initial wealth, then he will reject a 50–50 gamble to lose \$1000 or win an infinite amount.

The Rabin critique at first received push back (for a review, see Wakker 2010, pp. 244–245), but it is now widely accepted as a genuine critique of expected utility theory (see, e.g., Bleichrodt et al. 2019). Note that Samuelson (1963) made a related argument almost 40 years earlier. He showed that expected utility theory implies that if a person rejects a small-stakes gamble with positive expected value (e.g., a 50–50 bet to lose \$100 or win \$200), then he must also reject a large-stakes gamble comprising 100 repetitions of the gamble even though the probability of an aggregate loss is extremely low (less than 0.1% in the foregoing example).

Although the Allais and Ellsberg paradoxes and the Rabin critique were originally presented as thought experiments, they have been replicated many times in the laboratory (Camerer 1995; Cox et al. 2013; Bleichrodt et al. 2019).

Non-expected Utility Theories

There are numerous alternatives to and generalizations of expected utility theory that respond to the challenges posed by the Allais and Ellsberg paradoxes and the Rabin critique. Perhaps the most famous is prospect theory, which was developed by psychologists Daniel Kahneman and Amos Tversky (1979) and Tversky and Kahneman (1992). Other well-known non-expected utility theories include rank-dependent expected utility theory (Quiggin 1982), regret theory (Bell 1982; Fishburn 1982; Loomes and Sugden 1982, 1987), disappointment theory (Bell 1985; Loomes and Sugden 1986; Gul 1991), Yaari's (1987) dual theory, Choquet expected utility theory (Schmeidler 1989; Gilboa 1987), maxmin expected utility theory (Gilboa and Schmeidler 1989), α-maxmin expected utility theory (Ghirardato et al. 2004), multiplier preferences (Hansen and Sargent 2001; Strzalecki 2011), smooth ambiguity preferences (Klibanoff et al. 2005), variational preferences (Maccheroni et al. 2006), uncertainty averse preferences (Cerreia-Vioglioa et al. 2011), and expected uncertain utility theory (Gul and Pesendorfer 2014). Surveys of non-expected utility theories include Starmer (2000), Sugden (2004), Etner et al. (2012), and Keith and Ahner (2021). Textbook treatments include Wakker (2010) and Dhami (2016).

Cross-References

► Choice Under Risk and Uncertainty

References

Allais M (1979) The foundations of a positive theory of choice involving risk and a criticism of the postulates and axioms of the American school. In: Allais M, Hagen O (eds) Expected utility hypothesis and the Allais Paradox. D. Reidel, Dordrecht, pp 27–145. [Originally published as “Fondements d’une Théorie Positive des Choix Comportant un Risque et Critique des Postulats et Axiomes de l’Ecole Américaine,” in *Econométrie*, Colloques Internationaux du Centre National de la Recherche Scientifique, vol. 40 (1953), Paris: CNRS, pp. 257–332. Summarized in “Le Comportement de l’Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l’Ecole Américaine,” *Econometrica*, vol. 21 (1953), pp. 503–546]

Anscombe FJ, Aumann RJ (1963) A definition of subjective probability. Ann Math Stat 34:199–205

Arrow KJ (1971) Essays in the theory of risk-bearing. Markham, Chicago

Bell DE (1982) Regret in decision making under uncertainty. Oper Res 30:961–981

Bell DE (1985) Disappointment in decision making under uncertainty. Oper Res 33:1–27

Bernoulli D (1954) Exposition of a new theory on the measurement of risk. Econometrica 22:23–36. [Originally published as “Specimen Theoriae Novae de Mensura Sortis,” in *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, vol. 5 (1738), pp. 175–192]

Bleichrodt H, Doctor JN, Gao Y, Li C, Meeker D, Wakker PP (2019) Resolving Rabin’s paradox. J Risk Uncertain 59:239–260

Camerer C (1995) Individual decision making. In: Hagel JH, Roth AE (eds) The handbook of experimental economics. Princeton University Press, Princeton, pp 587–703

Cerreia-Vioglioa S, Maccheroni F, Marinacci M, Montrucchio L (2011) Uncertainty averse preferences. J Econ Theory 146:1275–1330

Cox JC, Sadiraj V, Vogt B, Dasgupta U (2013) Is there a plausible theory for decision under risk? A dual calibration critique. Econ Theory 54:305–333

de Finetti B (1964) Foresight: its logical laws, its subjective sources. In: Kyburg HE Jr, Smokler HE (eds) Studies in subjective probability. Wiley, New York, pp 93–158. [Originally published as “La Prévision: Ses Lois Logiques, Ses Sources Subjectives,” in *Annales de l’Institut Henri Poincaré*, vol. 7 (1937), pp. 1–68]

Dhami S (2016) The foundations of behavioral economic analysis. Oxford University Press, Oxford

Ellsberg D (1961) Risk, ambiguity, and the savage axioms. Q J Econ 75:643–669

Etner J, Jeleva M, Tallon J-M (2012) Decision theory under ambiguity. J Econ Surv 26:234–270

Fishburn PC (1970) Utility theory for decision making. Wiley, New York

Fishburn PC (1982) Nontransitive measurable utility. J Math Psychol 26:31–67

Fishburn PC (1988) Nonlinear preference and utility theory. Johns Hopkins University Press, Baltimore

Ghirardato P (2002) Revisiting savage in a conditional world. Econ Theory 20:83–92

Ghirardato P, Maccheroni F, Marinacci M (2004) Differentiating ambiguity from ambiguity attitude. J Econ Theory 118:133–173

Gilboa I (1987) Expected utility with purely subjective non-additive probabilities. J Math Econ 16:65–88

Gilboa I, Schmeidler D (1989) Maxmin expected utility with a non-unique prior. J Math Econ 18:141–153

Gul F (1991) A theory of disappointment aversion. Econometrica 59:667–686

Gul F, Pesendorfer W (2014) Expected uncertain utility theory. Econometrica 82:1–39

Hacking I (1980) Strange expectations. Philos Sci 47:562–567

Hadar J, Russell WR (1969) Rules for ordering uncertain prospects. Am Econ Rev 59:25–34

Hammond PJ (1988) Consequentialist foundations for expected utility. Theor Decis 25:25–78

Hanoch G, Levy H (1969) The efficiency analysis of choices involving risk. Rev Econ Stud 36:335–346

Hansen LP, Sargent TJ (2001) Robust control and model uncertainty. Am Econ Rev 91:60–66

Hernstein IN, Milnor J (1953) An axiomatic approach to measurable utility. Econometrica 21:291–297

Jensen NE (1967) An introduction to Bernoullian utility theory: I. Utility functions. Swed J Econ 69:163–183

Kahneman D, Tversky A (1979) Prospect theory: an analysis of decision under risk. Econometrica 47:263–292

Karni E, Schmeidler D (1991) Atemporal dynamic consistency and expected utility theory. J Econ Theory 54: 401–408
Keith AJ, Ahner DK (2021) A survey of decision making and optimization under uncertainty. Ann Oper Res 300: 319–353
Klibanoff P, Marinacci M, Mukerji S (2005) A smooth model of decision making under ambiguity. Econometrica 73:1849–1892
Kreps DM (1988) Notes on the theory of choice. Westview Press, Boulder
Loomes G, Sugden R (1982) Regret theory: an alternative theory of rational choice under uncertainty. Econ J 92: 805–824
Loomes G, Sugden R (1986) Disappointment and dynamic consistency in choice under uncertainty. Rev Econ Stud 53:271–282
Loomes G, Sugden R (1987) Some implications of a more general form of regret theory. J Econ Theory 41:270–287
Luce RD, Raiffa H (1957) Games and decisions: introduction and critical survey. Wiley, New York
Maccheroni F, Marinacci M, Rustichini A (2006) Ambiguity aversion, robustness, and the variational representation of preferences. Econometrica 74:1447–1498
Menezes CF, Hanson DL (1970) On the theory of risk aversion. Int Econ Rev 11:481–487
Myerson RB (1979) An axiomatic derivation of subjective probability, utility, and evaluation functions. Theor Decis 11:339–352
Nebout A (2014) Sequential decision making without independence: a new conceptual approach. Theor Decis 77: 85–110
Pratt JW (1964) Risk aversion in the small and in the large. Econometrica 32:122–136
Quiggin J (1982) A theory of anticipated utility. J Econ Behav Organ 3:323–343
Rabin M (2000) Risk aversion and expected-utility theory: a calibration theorem. Econometrica 68:1281–1292
Rabin M, Thaler RH (2001) Anomalies: risk aversion. J Econ Perspect 15:219–232
Ramsey FP (1931) Truth and probability. In: Braithwaite RB (ed) The foundations of mathematics and other logical essays. Routledge, London, pp 156–198
Rothschild M, Stiglitz JE (1970) Increasing risk: I. A definition. J Econ Theory 2:225–243
Samuelson PA (1963) Risk and uncertainty: a fallacy of large numbers. Scientia 98:108–113
Savage LJ (1972) The foundation of statistics, 2nd revised edition. Dover, New York. [First edition 1954, Wiley]
Schmeidler D (1989) Subjective probability and expected utility without additivity. Econometrica 57:571–587
Starmer C (2000) Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk. J Econ Lit 38:332–382
Strzalecki T (2011) Axiomatic foundations of multiplier preferences. Econometrica 79:47–73
Sugden R (2004) Alternatives to expected utility: foundations. In: Barberà S, Hammond PJ, Seidl C (eds) Handbook of utility theory, vol 2. Kluwer Academic, Boston, pp 685–755
Tversky A, Kahneman D (1992) Advances in Prospect theory: cumulative representation under uncertainty. J Risk Uncertain 5:297–323
Volij O (1994) Dynamic consistency, consequentialism and the reduction of compound lotteries. Econ Lett 46:121–129
von Neumann J, Morgenstern O (1953) Theory of games and economic behavior, 3rd edn. Princeton University Press, Princeton. [1st edition 1944. 2nd edition 1947]
Wakker PP (2010) Prospect theory: for risk and ambiguity. Cambridge University Press, Cambridge
Yaari ME (1987) The dual theory of choice under risk. Econometrica 55:95–115