

Alternatives to Subjective Expected Utility:

Ambiguity Aversion (MEU), Capacities (CEU), and Smooth Ambiguity Aversion (SAA)

The observed behavior in the Ellsberg paradox led many scholars to suggest that individuals have ambiguous (unclear) beliefs, rather than objective or subjective beliefs as previous theories considered. In order to allow for such ambiguity, and to explain the Ellsberg paradox, the literature has offered alternative theories. Given their importance and subsequent use in applications, we next present three of them: expected utility theory with multiple priors (also referred to as maxmin expected utility, MEU), rank-dependent expected utility (or Choquet expected utility, CEU), and smooth ambiguity aversion (SAA). In our discussion, we let F denote an act $F : S \rightarrow X$ from the set of states to the set of outcomes (which can be monetary payoffs or any other prize).

1.1 Maxmin Expected Utility (MEU)

Following Gilboa and Schmeidler (1989), consider a setting where subjects have too little information to form their priors, and allow the subjects to consider a *set of priors*. Now, if an individual is uncertainty averse (to be defined below), he will choose lottery F over another lottery G if the former provides a higher expected utility than the latter according to a worst possible prior. In this context, Gilboa and Schmeidler replaced the IA with their “uncertainty aversion” axiom and their “certainty-independence” axiom, which we define next.

Uncertainty Aversion. Consider an individual who is indifferent between two lotteries F and G . Then he is *uncertainty averse* if he weakly prefers the compound lottery $\alpha F + (1 - \alpha)G$ to simple lottery G , where $\alpha \in (0, 1)$.

Intuitively, a decision maker who is uncertainty averse has a preference for mixing (or hedging), since the compound lottery becomes at least as valuable as either of the two lotteries alone.¹ In particular, that may occur when the two acts have negatively correlated payoffs, so their mixture yields a more even payoff distribution across states.² Let us next define the second axiom, certainty independence, which weakens the IA.

Certainty Independence. For any two lotteries F and G and a constant act K (e.g., a certain outcome, or a lottery that remains constant across all states), the decision maker weakly prefers lottery F to G if and only if he prefers $\alpha F + (1 - \alpha)K$ to $\alpha G + (1 - \alpha)K$, where $\alpha \in (0, 1)$.

¹ The standard IA is actually a special case of the uncertainty aversion axiom if we restrict the decision maker to exhibit indifference between the compound lottery and the simple lotteries. Indeed, if the individual is indifferent between F and G , then by the IA he is also indifferent between $\alpha F + (1 - \alpha)G$ and $\alpha G + (1 - \alpha)G$, that is, between $\alpha F + (1 - \alpha)G$ and G ; thus becoming a special case of the above definition of uncertainty aversion.

² Intuitively, think about two investments, one being attractive when the economy grows (e.g., shares of a diamond company, or luxury cars) and another when the economy enters a crisis (e.g., shares of a dollar store). If you are indifferent between both of them, you might prefer their mixture rather than one of the two shares alone.

Note that the certainty-independence axiom relaxes the IA as it only requires that preferences over two lotteries F and G be unaffected when each lottery is mixed with a certain outcome K (rather than when mixed with a third lottery, as required by the IA). In other words, certainty independence can be understood as a special case of the IA, where the third lottery has degenerate probability distribution, thus describing an outcome that occurs with certainty.

Importantly, when we replace the IA with uncertainty aversion and certainty independence, we obtain a nice utility representation of preferences that allows for the possibility of the individual decision maker sustaining multiple prior beliefs. In particular, a decision maker weakly prefers lottery F to G if and only if (i.e., it is equivalent to)

$$\min_{p \in C} \int_S u(F(s)) dp(s) \geq \min_{p \in C} \int_S u(G(s)) dp(s).$$

That is to say, the individual decision maker first evaluates the expected utility of lotteries F and G according to each of his multiple priors $p \in C$; he then identifies the prior that yields the lowest expected utility when playing lottery F , and similarly for lottery G ; last, he selects the lottery with the highest expected utility. Intuitively, he separately evaluates each lottery at his worst possible prior, and then chooses the lottery yielding the highest expected utility, which explains why this approach is also referred to as “maxmin expected utility,” or MEU. (For a proof of this expected utility representation result, see Gilboa and Schmeidler (1989), pp. 145–49; or the more didactic presentation in Ok (2007)’s textbook, pp. 513–19).

Example 1: Ambiguity aversion Consider a decision maker with Bernoulli utility function $u(x) = \sqrt{x}$, where $x \geq 0$ denotes monetary amounts. Assume that he faces two lotteries $L_A = (\$1, \$100)$ and $L_B = (\$3, \$5)$, and allow the decision maker to hold priors $(p_A, 1 - p_A)$ for lottery L_A , and $(p_B, 1 - p_B)$ for lottery L_B . According to MEU, he chooses lottery L_B if

$$\min_{p_B} [p_B \sqrt{3} + (1 - p_B) \sqrt{5}] \geq \min_{p_A} [p_A \sqrt{1} + (1 - p_A) \sqrt{100}].$$

For instance, priors are left unrestricted if the decision maker does not have any available information with which to update his priors, so they can take values $p_A, p_B \in [0, 1]$. In this setting, it is possible that his most pessimistic belief is that by which he receives the lowest monetary amount with probability one in both lotteries, implying that in lottery L_B ,

$$\min_{p_B} [p_B \sqrt{3} + (1 - p_B) \sqrt{5}] = \sqrt{3}$$

with argmin $p_B = 1$, and similarly for lottery L_A ,

$$\min_{p_A} [p_A \sqrt{1} + (1 - p_A) \sqrt{100}] = 1$$

with $\text{argmin } p_A = 1$. Hence a decision maker with MEU preferences selects lottery L_B because $\sqrt{3} > 1$. Note that such a choice is somewhat extreme, as the decision maker first focuses on the worst-case scenario within each lottery, and then selects the lottery that provides the best worst-case scenario. \square

Next we will show that the Ellsberg paradox, despite not being rationalized by subjective probability theory, can be explained if the decision maker exhibits MEU preferences. Recall from the Ellsberg paradox that gamble A (B) rewards the individual with \$1,000 if the ball he extracts is red (blue, respectively). While the proportion of red balls in the urn is known, $1/3$, the decision maker does not receive any information about the number of blue balls. Hence an individual with MEU preferences finds

$$U(A) = \frac{1}{3}u(\$1,000) \quad \text{and} \quad U(B) = \min_{p_{Blue}} p_{Blue}u(\$1,000).$$

Thus, if the individual selects gamble A , as required by the Ellsberg paradox, he reveals that p_{Blue} satisfies

$$\frac{1}{3} > \min_{p_{Blue}} p_{Blue}, \tag{1}$$

where the right-hand side indicates the worst possible prior among all those considered by the decision maker. A similar argument applies for gamble C (extracting a ball that is not red) and gamble D (extracting a ball that is not blue), where

$$U(C) = \frac{2}{3}u(\$1,000) \quad \text{and} \quad U(D) = \min_{p_{Not\ Blue}} p_{Not\ Blue}u(\$1,000).$$

Hence, if the decision maker selects gamble C , as described in the Ellsberg paradox, he reveals that

$$\frac{2}{3} > \min_{p_{Not\ Blue}} p_{Not\ Blue}. \tag{2}$$

Conditions (1) and (2) are compatible, since

$$\min_{p_{Blue}} p_{Blue} \neq 1 - \min_{p_{Not\ Blue}} p_{Not\ Blue},$$

can easily occur. For instance, if the decision maker has two possible priors $(p_{Blue}, p_{Not\ Blue}) = (0.1, 0.9)$ and $(0.4, 0.6)$, the lowest prior p_{Blue} is $\min_{p_{Blue}} p_{Blue} = \min\{0.1, 0.4\} = 0.1$, while the lowest for $p_{Not\ Blue}$ is

$\min_{p_{Not\ Blue}} p_{Not\ Blue} = \min\{0.9, 0.6\} = 0.6$, entailing that $\min_{p_{Blue}} p_{Blue} \neq 1 - \min_{p_{Not\ Blue}} p_{Not\ Blue}$ since $0.1 \neq 0.6$.

Hence the Ellsberg paradox can be explained according to MEU.

1.2 Choquet Expected Utility (CEU)

Schmeidler (1989) provided an alternative theory to account for the anomaly in the Ellsberg paradox by defining beliefs with the use of capacities. Specifically, a capacity is defined as a real-valued function $\nu(\cdot)$ from a subset of the state space S to $[0, 1]$, with the normalization $\nu(\emptyset) = 0$ and $\nu(S) = 1$. In addition, we say that capacity $\nu(\cdot)$ satisfies *comonotonicity* if $\nu(A) \geq \nu(B)$, where A is a superset of B ; in words, the probability weight assigned to the larger set A cannot be smaller than that assigned to its subset B . While a probability distribution assigns weights that are additive across two mutually exclusive events, such as $a, b \in S$ (their probabilities are additive), their capacities allow for nonadditivity. For instance, if $S = \{a, b\}$, then an example of a capacity is $\nu(\emptyset) = 0$, $\nu(a) = 0.4$, $\nu(b) = 0.5$, and $\nu(a, b) = 1$, where $\nu(a) + \nu(b) \neq \nu(a, b)$.

In this context, we cannot simply use a standard integral over states, since capacity $\nu(\cdot)$ does not correspond to our notion of beliefs. However, Choquet (1953) developed integrals specifically for capacities, which Schmeidler (1989) used in his representation result. In particular, a decision maker weakly prefers F to G if the Choquet integrals satisfy

$$\int_S u(F(S)) d\nu(S) \geq \int_S u(G(S)) d\nu(S).$$

Moreover the CEU and MEU models are connected if we impose the uncertainty aversion axiom in this CEU context. In order to do that, we first need that capacity $\nu(\cdot)$ satisfies supermodularity (also referred to as complementarity in this context), in the sense that the marginal contribution of a subset of events B is larger when it joins a large set of events A than when it joins a small set C . To see this, let us take two sets A and C , where C is a subset of A , that is, $C \subset A$. In this context, the marginal contribution of B to the large set A is larger than its marginal contribution to the small group C :

$$\nu(A \cup B) - \nu(A) \geq \nu(C \cup B) - \nu(B),$$

where A and B do not intersect, $A \cap B = \emptyset$.³

Example 2: Simple capacities While the use of Choquet integrals is involved, the literature often uses “simple” capacities. Intuitively, a *simple capacity* on state space S can be understood as a convex combination between two extreme capacities: (1) a standard probability weight on A , $p(A) \in [0, 1]$, which

³ Schmeidler (1986) showed that CEU and MEU are mathematically related. In particular, the Choquet integral of F with respect to a supermodular capacity ν coincides with the evaluation of F according to MEU when the decision maker’s priors belong to the core of capacity ν .

is a capacity with additive probabilities; and (2) the “complete ignorance” capacity w , which assigns $w(S)=1$ and $w(A)=0$ for every $A \subseteq S$. (In other words, in w the decision maker assigns a zero probability weight to all events, except to the entire state space S .) Formally, simple capacities are then defined as

$$v(A) = \lambda p(A) + (1 - \lambda)w(A)$$

for every $A \subseteq S$ and where $\lambda \in [0,1]$. In this context, parameter λ denotes the individual’s degree of confidence on $p(A)$, while $1 - \lambda$ captures his “degree of ambiguity” about $p(A)$ (also referred to as his “degree of uncertainty aversion”). For applications of simple capacities to noncooperative games, see [Haller \(2000\)](#), and to common pool resources, see [Aflaki \(2013\)](#). \square

1.3 Smooth Ambiguity Aversion (SAA)

Klibanoff (2005) offered a framework that accounts for observed behavior in the Ellsberg paradox while achieving separation between ambiguity (defined as the decision maker considering multiple priors) and his ambiguity attitude (a reflection of the decision maker’s taste).⁴ A decision maker weakly prefers lottery F to G if and only if

$$\int_p \left[\int_S u(F(s)) dp(s) \right] d\mu(p) \geq \int_p \left[\int_S u(G(s)) dp(s) \right] d\mu(p).$$

That is to say, the individual decision maker, first, evaluates the expected utility of lottery F for each prior $p \in C$; and second, he aggregates using prior μ over each of his multiple priors $p \in C$ given the increasing function $\phi(\cdot)$. Performing a similar calculation for lottery G , he can choose the lottery with the highest expected value. Intuitively, priors μ capture the ambiguity the decision maker faces, as a more concentrated support of μ indicates that he does not face much ambiguity on the possible priors $p \in C$. In the extreme, when μ is degenerate on a single prior p , the SAA model reduces to the standard EU model. Moreover the shape of function $\phi(\cdot)$ captures his attitude toward ambiguity. For example, if $\phi(x) = -e^{-\alpha x}$, parameter α describes the concavity of the function, as depicted in figure 1. When $\alpha \rightarrow 0$, function $\phi(x) \rightarrow -1$ for every prior p , assigning the same weight on all priors. However, when α increases, so does the concavity of function $\phi(x)$, assigning a zero probability weight on a larger support of priors. In the limit, when $\alpha \rightarrow \infty$, function $\phi(x) \rightarrow 0$ for all priors, except for the lowest (i.e., the prior corresponding to the worst case scenario); indicating that in such limiting case SAA collapses to the MEU model we discussed above.

⁴ Thanks to PhD candidate Casey Bolt for his suggestions in this subsection, especially in example 3.

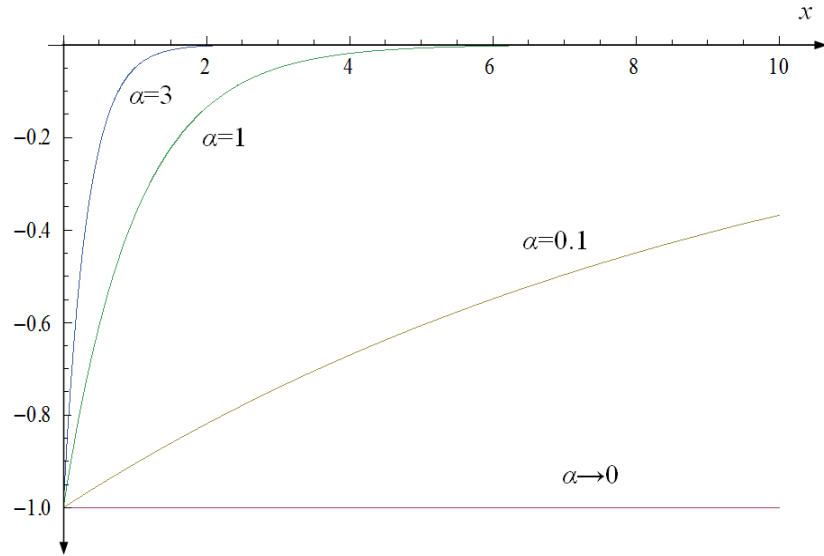


Figure 1. Function $\phi(x) = -e^{-\alpha x}$ evaluated at different values of α

Example 3: Investing under ambiguity Consider a decision maker with utility function $u(W) = aW - bW^2$, where $a, b > 0$ and W is exogenous wealth.⁵ He chooses his investment $\$x$ in a risky asset, which yields a wealth of $W + Rx$, where the realization of return R is either low, $R_L = 0.75$, or high, $R_H = 2$. (For simplicity, we can assume that the decision maker's budget constraint with respect to investing is not binding.) Due to ambiguity, this individual forms two priors over the low return, $\alpha_1 = 0.25$ or $\alpha_2 = 0.90$, which entails that his priors on the high return are $1 - \alpha_1 = 0.75$ or $1 - \alpha_2 = 0.10$. The decision maker evaluates ambiguity with the increasing function $\phi(u) = u^\rho$, which displays constant relative ambiguity aversion since

$$A_R(u, \phi) = -\frac{\phi''(u)}{\phi'(u)}u = 1 - \rho.$$

(Note that expression $A_R(u, \phi)$ is analogous to the Arrow-Pratt coefficient of relative risk aversion, but evaluating the function that the individual uses to evaluate the ambiguity of priors $\phi(u)$, rather than his Bernoulli utility function.) If $\rho = 1$, then $A_R(u, \phi) = 0$, which implies that the decision maker is ambiguity

⁵ This functional form has a coefficient of relative risk aversion of $r_R(W, u) = -W \frac{u''(W)}{u'(W)} = \frac{2bW}{a - 2bW}$ with

derivative $\frac{\partial r_R(W, u)}{\partial W} = \frac{2ab}{(a - 2bW)^2} > 0$, since parameters satisfy $a, b > 0$, by definition.

neutral while he can still exhibit risk aversion. As ρ decreases, relative ambiguity aversion increases, so for $\rho < 1$ the individual is ambiguity averse. If instead $\rho > 1$, the decision maker is ambiguity loving. Assuming that the individual places equal weight on his priors α_1 or α_2 being correct, his problem becomes

$$\max_x 0.5[\alpha_1 u(x; R_L) + (1 - \alpha_1)u(x; R_H)]^\rho + 0.5[\alpha_2 u(x; R_L) + (1 - \alpha_2)u(x; R_H)]^\rho.$$

We can now plug in parameters α_1 , α_2 , and utility function $u(\cdot)$. For simplicity, we also consider parameters $a = 10, b = 1, \rho = 0.5$, and wealth $W = 1$ (so that the low return $R_L = 0.75$ entails a 25 percent loss on every dollar invested in the asset). The above problem becomes

$$\begin{aligned} \max_x 0.5 & \left[2.5(1 + 0.75x) - 0.25(1 + 0.75x)^2 + 75(1 + 2x) - 0.75(1 + 2x)^2 \right]^{0.5} \\ & + 0.5 \left[9(1 + 0.75x) - 0.9(1 + 0.75x)^2 + (1 + 2x) - 0.1(1 + 2x)^2 \right]^{0.5}. \end{aligned}$$

Taking first-order condition with respect to x yields

$$\begin{aligned} & (9.375 - 0.125[1.5 + 1.125x] + 7.5 - 0.375[4 + 8x])Eu(x; \alpha_1)^{-0.5} \\ & + (33.75 - 0.45[1.5 + 1.125x] + 1 - 0.05[4 + 8x])Eu(x; \alpha_2)^{-0.5} = 0. \end{aligned}$$

And solving for x , we obtain an investment of $x \approx 2.3$. If we replicate this problem for a ambiguity neutral individual, where $\rho = 1$, we find a higher investment of $x \approx 2.53$. Thus ambiguity aversion reduces his investment in the risky asset. \square

Further reading The preceding two models have been used in many applications. Dow and Werlang (1992) consider individuals with supermodular capacities in the CEU framework analyzed above to show that individuals do not have incentives to buy or sell assets for a potentially large range of prices. A similar finding emerged in variations on the model developed by Epstein and Wang (1994) and Uppal and Wang (2003). The multiple priors model in MEU has also been applied to labor economics by Nishimura and Ozaki (2004), and to macroeconomic policy by Hansen et al. (1999) and Hansen and Sargent (2001, 2008). For a detailed review of the different models proposed in this literature that build on MEU and CEU, see Machina and Siniscalchi (2014), and for an extensive survey of applications of these models to economics and finance, see Mukenji and Tallon (2004). Last, the preceding models have been recently tested in experimental labs, providing evidence of support for MEU and CEU in certain contexts, but relatively more consistent evidence that supports prospect theory; see Abdellaoui et al. (2014), Baillon, and Bleichrodt (2015), and Dimmock et al. (2015).