# Ambiguity Aversion, Risk Aversion, and the Weight of Evidence

Edi Karni
Johns Hopkins University*

March 9, 2023

**Abstract**

Wakker (1990) showed that under natural conditions the Choquet expected utility (CEU) and the rank-dependent utility (RDU) models are identical. Invoking Wakker's result and applying the certainty equivalent reduction procedure, this paper shows that risk aversion in the RDU model implies ambiguity aversion in the corresponding CEU model. Consequently, the pattern of choice depicted by Ellsberg's experiments and, more generally, preference for evidence to support their beliefs is an expression of decision makers' risk aversion. This paper also shows that, properly formulated, smooth ambiguity aversion may also be regarded as risk aversion.

**Keywords:** Ambiguity aversion; Risk aversion; Choquet expected utility, Rank-dependent utility; Bayesianism;

**JEL classification:** D8, D81, D83

*Department of Economics, Wyman Park Building, 3100 Wyman Park Dr, Baltimore, MD 21211, USA. E-mail: karni@jhu.edu. I am very grateful to Peter Wakker for his insightful comments.

# 1 Introduction

Bayesian decision theory, pioneered by Ramsey (1931) and de Finetti (1937), culminated in Savage's (1954) and Anscombe and Aumann (1963) subjective expected utility (SEU) models. These models presume that decision makers entertain beliefs about the likelihoods of events that are quantifiable by probability measures. Moreover, these beliefs are manifested in, and can be inferred form, decision makers choice behavior. Using different choice sets, both Savage and Anscombe and Aumann depict structures of preference relations (i.e., patterns of choice behaviors) that are necessary and sufficient to quantify a decision maker's beliefs by a unique subjective probability measure.

Consider a decision maker whose belief that an event, say $E$, obtains is quantified by a probability $\pi(E)$. Presumably, this belief incorporates the information of the decision maker regarding the plausibility of this event. Moreover, it is natural to suppose that the quality of the information and the confidence in the belief it inspires, should affect the decision maker behavior. The subjective expected utility models, however, fail to account for the weight of evidence, or information, that support the decision makers' probabilistic belief. In particular, given a posterior probability a decision maker's choices are the same regardless of the evidence supporting that posterior. Put differently, SEU theories, accords no weight to the evidence supporting the decision maker's beliefs.

To grasp the issue, consider an urn containing a hundred balls that are known to be either red or black. No other prior information about the color composition of the balls in the urn is available. A ball is about to be drawn at random and a decision maker is contemplating placing a bet on its color. Compare the following two scenarios; in the first scenario, before placing the bet, a sample of balls are drawn from the urn repeatedly, with replacement, and their colors observed. Suppose that after repeated draws, it so happens that the number of red and black balls in the sample is the same. The decision maker concludes that observing either color in the next draw is equally likely. In the second scenario no balls are drawn before the decision maker is required to place the bet. It seems reasonable that, by reason of symmetry (or insufficient reason), the decision maker believes that the events drawing a red ball and drawing a black ball are equally likely. According to subjective expected utility theory, in both scenarios, the decision maker's subjective probabilities of the two events are the same and equal 1/2. Consequently, the decision maker should be indifferent between betting on red and betting on black, and he should also be indifferent to between betting on red on a draw from the tested and the untested urn. The greater confident in his belief about the likely outcome if the ball is drawn form the tested urn is completely disregarded.

This lack of consideration of the weight of evidence supporting the decision making beliefs has been long recognized. In *Treatise on Probability* Keynes discusses the scenario similar to the one described above. In Keynes' words, "...in the first case we know that the urn contains black and white in equal proportions; in the second case the proportion of each color is unknown, and each ball is as likely to be black as white. It is evident that in either case the probability of drawing a white ball is 1/2, but that the weight of the argument in favour of this conclusion is greater in the first case." Keynes (1921 [1973], p. 82).[1]

Ellsberg (1961) argued that, facing choices between the bets described in the example above, decision makers exhibit strict preference for betting on either color of a ball drawn form an urn containing equal number of black and white balls over betting on either color of a ball drawn from an urn in which proportion of each color is unknown. These preferences are inconsistent with the existence of additive subjective probability à la Savage.

Schmeidler (1982, 1986, 1989) proposed a novel decision model, that later became known as Choquet expected utility (CEU), designed to accommodate choice behavior that accounts for the information supporting the decision maker's beliefs.[2] According to the CEU model, decision makers maximize the expectations of a utility function with respect to a non-additive probability measure, or capacity. To capture the preference for being better informed, Schmeidler introduced the notion of ambiguity aversion and characterized it by convex capacity.[3]

About the same time that Schmeidler developed the CEU model, a class of theories of decision making under risk, dubbed rank-dependent utility models (RDU) were introduced.

[1] Zappia (2020) contains a detailed review of this issue including an exchange between Savage and Popper prompted by it. Zappia (2020) mentioned that Popper (1958) "...argued that he found it paradoxical that two apparently similar events should be attributed the same subjective probability even though the evidence supporting judgment in one case was stronger than in the other case."

[2] For alternative modeling of Choquet expected utility see Gilboa (1987), Wakker (1989a, 1989b), Nakamura (1990) and Chew and Karni (1994).

[3] Schmeidler (1989) dubbed the preference for betting on events whose probability is supported by more evidence uncertainty aversion. This term was replaced by abiguity aversion in the nomenclature of decision theory. Schmeidler also discusses several equivalent characterizations of ambiguity aversion.

These include Quiggin's (1982) Anticipated Utility, Yaari's (1987) Dual Theory as well as the general RDU model of Chew (1989). According to these models preference relations on objective risks, or lotteries, are represented by the inner product of a utility function on the set of outcomes and a transformation function on the corresponding probabilities, where the transformation of the objective probabilities depends on the ranks of the outcomes in the set of feasible outcomes.

Wakker (1990) showed that under natural conditions the CEU and RDU models are identical. Invoking Wakker's result, I argue that the representation of risk aversion by a RDU model implies the representation of ambiguity aversion by the corresponding to a CEU model.

A different formalization of the idea of ambiguity aversion proposed by Klibanoff, Marinacci, and Mukerji (2005) and Seo (2009). In their models ambiguity aversion is captured by a concave real-valued function on the expected utilities associated with the set of all conceivable priors. In this paper I argue that, properly formulated, smooth ambiguity aversion may also be regarded as risk aversion.

The next section sets the stage, describing briefly the rank-dependent utility and Choquet expected utility representations. Section 3 discusses two procedure of reduction of compound lotteries and introduces the lead example. Section 4 includes the statement and proof of the main result. Section 5 applies the main idea to the models based on second-order beliefs. The concluding section includes discussion and interpretation of the findings of this work.

## 2 Setting the Stage

### 2.1 Subjective expected utility and rank-dependent utility

Let $S$ be a finite set of *states* and let $X$ be an interval in the real line. Denote by $\mathcal{G}(X)$ the set of cumulative distribution functions (CDF) on $X$ with finite supports. The choice set, $A := \{f : S \to \mathcal{G}(X)\}$, consists of elements representing alternative courses of action and are referred to as *acts*. For all $f, g \in A$ and $\alpha \in [0,1]$, define $(\alpha f + (1-\alpha) g) \in A$ by $(\alpha f + (1-\alpha) g)(s) = \alpha f(s) + (1-\alpha) g(s)$, for all $s \in S$. Thus, $A$ is a convex set in a linear space.

A preference relation, denoted $\succcurlyeq$, is a complete and transitive binary relation on $A$. The asymmetric and symmetric parts of $\succcurlyeq$, are denoted by $\succ$ and $\sim$, respectively. Anscombe and Aumann (1963) provide necessary and sufficient conditions for the representation of $\succcurlyeq$ by a subjective expected utility functional. Formally, for all $f, g \in A$,

$$f \succcurlyeq g \Leftrightarrow \Sigma_{s \in S} U\left(f\left(s\right)\right) \pi\left(s\right) \geq \Sigma_{s \in S} U\left(g\left(s\right)\right) \pi\left(s\right), \tag{1}$$

where the *utility function* $U$ is affine real-valued function on $X$ and $\pi$ is an additive probability measure on $2^S$. The utility-probability pair $(U, \pi)$ is *jointly unique* (i.e., given $\pi$ the function $U$ is unique up to positive affine transformation and given $u$, $\pi$ is unique).

The RDU models – Quiggin's (1982, 1993) Anticipated Utility,[4] Yaari's (1987) Dual Theory as well as the general RDU model of Chew (1989) – are theories of decision making under risk (that is, the domain of the preference relation is the set, $\mathcal{G}(X) \subset A$, of constant acts that deliver the same CDF in every state). According to these models a preference relation $\succcurlyeq$ on $\mathcal{G}(X)$ is represented by RDU. Formally, for all $G, F \in \mathcal{G}(X)$,

$$G \succcurlyeq F \Leftrightarrow \int_X u\left(x\right) d\left(\zeta \circ G\right)\left(x\right) \geq \int_X u\left(x\right) d\left(\zeta \circ F\right)\left(x\right), \tag{2}$$

where the utility, $u$, is a real-valued function on $X$, unique up to positive affine transformation, and $\zeta : [0, 1] \to [0, 1]$ is nondecreasing, continuous, and onto function, dubbed *probability transformation function.*

### 2.2 Rank-dependent and Choquet expected utility

To analyze the relationships between ambiguity aversion and risk aversion we need to convert the RDU models to theories of decision making under uncertainty. This conversion was attained by Wakker (1990). Invoking a capacity of the CEU model, Wakker induced a probability transformation function in the RDU model such that when composed with an additive probability measure, mimics the capacity of the CEU model. Formally, a preference relation, $\hat{\succcurlyeq}$ on $A$, is said to have a CEU representation if for all $f, g \in A$,

$$f \hat{\succcurlyeq} g \Leftrightarrow \int_S U\left(f\left(s\right)\right) d\varphi\left(s\right) \geq \int_S U\left(g\left(s\right)\right) d\varphi\left(s\right), \tag{3}$$

where $U$ is a real-valued affine function on $\mathcal{G}(X)$, unique up to positive affine transformation, and $\varphi : 2^S \to [0, 1]$ is a capacity measure.[5]

[4] See also Segal (1989) and (1993).

[5] A capacity measure is a set function $\varphi$ on a measurable space $(S, \Sigma)$ such that $\varphi(\varnothing) = 0$, $\varphi(S) = 1$ and $\varphi(A) \leq \varphi(B)$ for all $A, B \in \Sigma$ such that $A \subseteq B$.

Given a preference relation $\hat{\succcurlyeq}$ on $A$ that has a CEU representation, define a binary relation $\succcurlyeq^*$on $2^S$ as follows: $C \succcurlyeq^* B$ if and only if $\varphi(C) \geq \varphi(B)$. Consider a decision maker whose preference relation $\hat{\succcurlyeq}$ on $A$ incorporates beliefs that are represented by a finitely-additive subjective probability measure, $\pi$ on $2^S$, such that $C \succcurlyeq^* B$ if and only if $\pi(C) \geq \pi(B)$. Thus, $\varphi(C) \geq \varphi(B)$ if and only if $\pi(C) \geq \pi(B)$. Therefore, there exists strictly increasing and onto probability transformation function $\zeta : [0,1] \to [0,1]$ defined by $\varphi(E) = (\zeta \circ \pi)(E)$, for all $E \in 2^S$.[6] Then

$$\int_S U(f(s))\, d\varphi(s) = \int_S U(f(s))\, d(\zeta \circ \pi)(s). \tag{4}$$

For each $f \in A$ let $c_f := \left(c_{f(s)}\right)_{s \in S}$, where $c_{f(s)}$ is the certainty equivalent of $f(s)$ (i.e., $U(f(s)) = U\left(\delta_{c_{f(s)}}\right)$, where $\delta$ is the Dirac measure). Then, by (3), $f \hat{\sim} c_f$ and, by (1), $f \sim c_f$. For each $f \in A$ define $E_f(x) = \{s \in S \mid c_{f(s)} \leq x)\}$, $x \in X$, and let $G_f(x) := \pi(E_f(x))$, for all $x \in X$. For each $f \in A$,

$$\int_S U(f(s))\, d(\zeta \circ \pi)(s) = \int_S u\left(c_{f(s)}\right) d(\zeta \circ \pi)(s) = \int_X u(x)\,(\zeta \circ \pi)(E_f(x))\, dx = \int_X u(x)\, d(\zeta \circ G_f)(x). \tag{5}$$

Therefore, we have a RDU model based on (additive) subjective probability measure $\pi$ on $S$ that mimics the CEU model.

# 3 The Weight of Evidence and Risk Aversion

## 3.1 Two reduction procedures

Resolutions of uncertainties may occur in stages as information may be obtained sequentially. If acting upon the receipt of early information is not feasible, the situation is represented by compound lotteries that may involve subjective and objective uncertainties. A prime example of such compound lotteries is the Anscombe and Aumann (1963) model in which the decision maker's beliefs about the likely outcomes of a horse race are represented by subjective probabilities and the payoffs, contingent on the outcomes of the horse race, are 'roulette lotteries' that assign objective probabilities to a set of prizes.

[6] Wakker (1990) Lemma 6 gives necessary and sufficient condition for the existence of such probability transformation function.

The literature dealing with decision making under uncertainty includes alternative procedures of reducing compound lotteries to equivalent (i.e., indifferent) one stage lotteries. To discuss the reduction procedures formally, let $\mathcal{G}(X)$ denote the set of one stage lotteries. A two-stage compound lottery is a probability distribution on $\mathcal{G}(X)$. Formally, $L_2 = [\pi(1), G_1; \pi(2), G_2; ...; \pi(n), G_n]$, where $G_s \in \mathcal{G}(X)$ and $\pi(\cdot)$ is a possible subjective probability measure on $S$. Denote by $\mathcal{L}_2$ the set of two-stage compound lotteries.

The reduction of compound lotteries axiom requires that the probabilities of the ultimate outcomes, or prizes, be the product of the probabilities of reaching the nodes along the path leading to the outcome. For our purpose we only need to two-stage compound lotteries.

**Reduction of Compound Lotteries Axiom (RCLA):** For all $L_2 \in \mathcal{L}_2$, $L_2 \sim [\Sigma_{s\in S}\pi(s) g_s(x_1), x_1; ...; \Sigma_{s\in S}\pi(s) g_s(x_m), x_m]$, where $g_s(x_i) := G_s(x_i) - G_s(x_{i-1})$, $i = 1, ..., m$.

An alternative reduction procedure calls for the replacement of the second-stage lottery by its certainty equivalent to obtain one stage lottery. Formally, denote by $G_s$ the constant act that assigns the lottery $G_s \in \mathcal{G}(X)$ to every $s \in S$. The certainty equivalent of $G_s$ is $c(G_s) \in \mathbb{R}$ defined by $\delta_{c(G_s)} \sim G_s$.

**Certainty Equivalence Reduction Axiom (CERA):** For all $L_2 \in \mathcal{L}_2$, $L_2 \sim \left[\pi(1), \delta_{c(G_1)}; \pi(2), \delta_{c(G_2)}; ...; \pi(n), \delta_{c(G_n)}\right]$.

Preference relations that admit expected utility representation are consistent with both reduction procedures. By contrast, preference relations that admit rank-dependent utility representations with nonlinear probability transformation functions are consistent with CERA but not with RCLA (See Segal [1990]).

## 3.2 The lead example

Consider the urn, described in the introduction, containing 100 balls that are either red or black. The color composition of the balls in the urn admits 101 feasible states. Imagine that it is possible to inspect the content of the urn to verify its composition. It is then possible to ask decision makers to choose among acts whose payoffs are contingent on the color composition of the balls in the urn. To formalize this decision problem, let $S = \{0, 1, ..., 100\}$, where $s \in S$ is the number of red balls in the urn. Consider a bet that pays off \$100 if a ball drawn at random is red and \$0 otherwise. Then, from the decision maker's viewpoint, the bet is an act, $f_R := (f_R(s))_{s\in S}$ whose payoffs are state-contingent lotteries $f_R(s) = (\$100, s/100; \$0, (1-s)/100)$. Let $c(f_R(s))$ denote the certainty equivalent of $f_R(s)$ (i.e., the decision maker is indifferent between the constant act that pays $c(f_R(s)) \in \mathbb{R}$ in every state and the constant act that pays off \$100 with probability $s/100$ and \$0 otherwise). Define the induced act $c_R = (c(f_R(0)), c(f_R(1)), ..., c(f_R(100)))$.

Suppose that a decision maker's belief regarding the color composition of the balls is represented by a subjective probability measure $\pi(\cdot \mid y^n)$ on $S$, where $y^n = (y_1, ..., y_n)$, $n \geq 1$, denotes the colors of a sample of $n$ balls drawn at random from the urn with replacement, and let $y^0$ indicates that no ball is drawn. Then $\pi\left(\cdot \mid y^0\right) = \pi(\cdot)$ is the subjective probability representing the prior belief.

Consider the decision maker's choice between the acts, $f_R|_{y^n}$ and $f_R|_{y^m}$, of betting on red conditional on observing the samples $y^n$ and $y^m$, respectively, where $0 \leq m \leq n$. Clearly, $\pi(\cdot \mid y^n)$ is a mean-preserving squeeze of $\pi(\cdot \mid y^m)$.[7] By definition every risk-averse decision maker prefers $f_R|_{y^n}$ over $f_R|_{y^m}$.

Let there be two identical urns, say $A$ and $B$, containing 100 balls each that are either red or black. Suppose that random samples of sizes $2m$ and $2n$, $m < n$, were drawn from $A$ and $B$, respectively, with replacement. Assume that the prior $\pi\left(\cdot \mid y^0\right)$ is symmetric and suppose that both samples happen to contain equal number of red and black balls.

Consider a decision maker whose preference relation has a RDU representation as in (4) and (5) facing a choice between two bets; $f_R|_A$, that pays off \$100 if the next ball drawn at random from urn $A$ is red, and $f_R|_B$, that pays off \$100 if the next ball drawn at random from urn $B$ is red. By Bayes rule, given a state $s$ and a sample of size $r = 2k$, by the binomial distribution, the probability of $k$ red balls is

$$\Pr(k \mid s) = \frac{k!}{k!(r-k)!}\pi\left(s \mid y^0\right)^r\left(1-\pi\left(s \mid y^0\right)\right)^{r-k} = \left[\frac{1}{k!}\pi\left(s \mid y^0\right)\left(1-\pi\left(s \mid y^0\right)\right)\right]^k$$

By Bayes rule, the posterior probabilities satisfy, for all $r = 2k$, $k = 1, 2, ...$,

$$\pi(s \mid k) = \frac{\left[\frac{1}{k!}\pi\left(s \mid y^0\right)\left(1-\pi\left(s \mid y^0\right)\right)\right]^k \pi\left(s \mid y^0\right)}{\Sigma_{s'\in S}\left[\frac{1}{k!}\pi\left(s' \mid y^0\right)\left(1-\pi\left(s' \mid y^0\right)\right)\right]^k \pi\left(s' \mid y^0\right)}.$$

An expected utility-maximizing decision maker would be indifferent between the two bets. To grasp this suffices it to note that an expected utility maximizer abides by the

[7] Formally, $\sum_{s=0}^{t}\pi\left(s \mid y^k\right) := \Pi\left(t \mid y^k\right), t \in \{0, 1, ..., 101\}$. Then, $\sum_{s=0}^{t}\left[\Pi\left(s \mid y^n\right) - \Pi\left(s \mid y^m\right)\right] \leq 0$ for all $t \in \{0, 1, ..., 101\}$ and $\sum_{s=0}^{100}\left[\Pi\left(s \mid y^n\right) - \Pi\left(s \mid y^m\right)\right] = 0$.

reduction of compound lotteries axiom. Letting $u(100) = 1$ and $u(0) = 0$, an expected utility maximizing decision maker evaluates the bets as follows:

$$U(f_R|_A) = \Sigma_{s=0}^{100} \frac{s}{100} \pi(s \mid y^m) = \frac{1}{2} = \Sigma_{s=0}^{100} \frac{s}{100} \pi(s \mid y^n) = U(f_R|_B).^8$$

However, a risk-averse RDU maximizing decision maker would strictly prefer to bet on red form urn $B$. To grasp this, recall that a RDU maximizing decision maker who is not an expected utility maximizer does not abide by the RCLA, displaying CERA instead. Moreover, risk-aversion in the RDU model requires that the probability transformation function, $\zeta$, that figures in the representation (5) be concave (see Chew, et al [1987]). Since $\pi(\cdot \mid y^m)$ is a mean-preserving spread of $\pi(\cdot \mid y^n)$, the RDU valuations of the bets are:

$$V(f_R|_A) = \sum_{s \in S} u(c(f_R(s))) \left[\zeta\left(\Sigma_{t=0}^{s} \pi(t \mid m/2)\right) - \zeta\left(\Sigma_{t=0}^{s-1} \pi(t \mid m/2)\right)\right]$$

and

$$V(f_R|_B) = \sum_{s \in S} u(c(f_R(s))) \left[\zeta\left(\Sigma_{t=0}^{s} \pi(t \mid n/2)\right) - \zeta\left(\Sigma_{t=0}^{s-1} \pi(t \mid n/2)\right)\right].$$

Hence, by risk-aversion, $V(f_R|_A) < V(f_R|_B)$. The weight of evidence tilts the bet in favor of the option in which the belief is based on more informative experiment. This observation underscores the point on which I elaborate in the analysis below, that it is risk-aversion in the context of the RDU model that captures the effect of the weight of evidence on the decision maker's preferences. The Ellsberg paradox is a special case of the argument above in $m = 0$ and $n \to \infty$.

[8] To grasp this claim observe that, by symmetry, $\pi(s \mid y^0) = \pi(100 - s \mid y^0)$, for all $s \in S$. Thus,

$$\pi(s \mid y^0)(1 - \pi(s \mid y^0)) = \pi(100 - s \mid y^0)(1 - \pi(100 - s \mid y^0)),$$

and

$$U(f_R|_A) = \frac{1}{100} \sum_{s=0}^{100} \frac{\left[\pi(s \mid y^0)(1 - \pi(s \mid y^0))\right]^{r/2} \pi(s \mid y^0)\, s}{\Sigma_{s' \in S} [\pi(s' \mid y^0)(1 - \pi(s' \mid y^0))]^{r/2} \pi(s' \mid y^0)} =$$

$$\frac{1}{100} \sum_{s=0}^{50} \frac{\left[2\pi(s \mid y^0)(1 - \pi(s \mid y^0))\right]^{r/2} \pi(s \mid y^0) \left( \frac{\pi(s|y^0)(1-\pi(s|y^0))s + \pi(100-s|y^0)(1-\pi(100-s|y^0))(100-s)}{2\pi(s|y^0)(1-\pi(s|y^0))} \right)}{\Sigma_{s' \in S} [\pi(s' \mid y^0)(1 - \pi(s' \mid y^0))]^{r/2} \pi(s' \mid y^0)} =$$

$$\frac{50}{100} \sum_{s=0}^{50} \frac{2\left[\pi(s \mid y^0)(1 - \pi(s \mid y^0))\right]^{r/2} \pi(s \mid y^0)}{\Sigma_{s' \in S} [\pi(s' \mid y^0)(1 - \pi(s' \mid y^0))]^{r/2} \pi(s' \mid y^0)} = \frac{1}{2}.$$

This is true for all $r = 2k, .k = 0, 1...$

# 4 Risk aversion and ambiguity aversion

## 4.1 Preliminaries

A preference relation $\succcurlyeq$ on $\mathcal{G}(X)$ is said to display risk aversion if, for all $F, G \in \mathcal{G}(X)$, $F \succcurlyeq G$ whenever $G$ differs from $F$ by a mean-preserving spread. It is said to display strict risk aversion if $F \succ G$.

The theory of risk aversion in the RDU model was developed in Chew et al (1987). Corollary 2 of Chew et al (1987) asserts that a preference relation $\succcurlyeq$ on $\mathcal{G}(X)$ that admits a RDU presentation $G \to \int_X u(x)\, d(\zeta \circ G)(x)$ displays risk aversion if and only if both the utility function, $u$, and the probability transformation function, $\zeta$, are concave. It displays strict risk aversion if and only if it displays risk aversion and either $u$ or $\zeta$ or both are strictly concave.

The notion of ambiguity aversion was first formulated and characterized by Schmeidler (1982) and was further developed in Schmeidler (1989). Formally, a preference relation $\hat{\succcurlyeq}$ on $A$ is said to display ambiguity aversion if, for all $f, g \in A$ and $\alpha \in [0,1]$, $f \hat{\succcurlyeq} g$ implies that $\alpha f + (1-\alpha) g \hat{\succcurlyeq} g$. It is said to display strict ambiguity aversion if $\alpha f + (1-\alpha) g \hat{\succ} g$. If $\hat{\succcurlyeq}$ admits CEU representation then ambiguity aversion is characterized by a convex capacity (i.e., a capacity satisfying $\varphi(B \cup C) + \varphi(B \cap C) \geq \varphi(B) + \varphi(C)$, for all $B, C \in 2^S$). Equivalently, it is characterized by concavity of the CEU functional $I(f) = \int_S u(f(s))\, d\varphi(s)$.[9]

To set the stage for the main result we need the following.

**Lemma:** *For all $f, g \in A$, $\alpha \in [0,1]$ and $x \in X$,*

$$\pi\left(E_{\alpha f + (1-\alpha) g}(x)\right) = \alpha \pi\left(E_f(x)\right) + (1-\alpha)\pi\left(E_g(x)\right).$$

*Proof.* By the theorem of Anscombe and Aumann (1963), the affinity of the utility function implies that, for all $f, g \in A$ and $\alpha \in [0,1]$,

$$\Sigma_{s \in S} u\left(\alpha f(s) + (1-\alpha) g(s)\right)\pi(s) = \alpha \Sigma_{s \in S} u\left(f(s)\right)\pi(s) + (1-\alpha)\Sigma_{s \in S} u\left(g(s)\right)\pi(s). \tag{6}$$

Equivalently, by the certainty equivalence reduction axiom,

$$\Sigma_{s \in S} u\left(c_{\alpha f + (1-\alpha) g}(s)\right)\pi(s) = \alpha \Sigma_{s \in S} u\left(c_f(s)\right)\pi(s) + (1-\alpha)\Sigma_{s \in S} u\left(c_g(s)\right)\pi(s). \tag{7}$$

[9] See the Proposition in Schmeidler (1989).

But, for all $f \in A$,

$$\Sigma_{s\in S} u\left(c_f\left(s\right)\right)\pi\left(s\right) = \int_X u\left(x\right) d\pi\left(E_f\left(x\right)\right). \tag{8}$$

Hence, by (7) and (8), for all $f, g \in A$ and $\alpha \in [0,1]$,

$$\int_X u\left(x\right) d\pi\left(E_{\alpha f+(1-\alpha)g}\left(x\right)\right) = \alpha \int_X u\left(x\right) d\pi\left(E_f\left(x\right)\right) + (1-\alpha)\int_X u\left(x\right) d\pi\left(E_g\left(x\right)\right) \tag{9}$$

$$= \int_X u\left(x\right) d\left(\alpha\pi\left(E_f\left(x\right)\right) + (1-\alpha)\pi\left(E_g\left(x\right)\right)\right).$$

By the uniqueness of $\pi$, we get that

$$\pi\left(E_{\alpha f+(1-\alpha)g}\left(x\right)\right) = \alpha\pi\left(E_f\left(x\right)\right) + (1-\alpha)\pi\left(E_g\left(x\right)\right). \tag{10}$$

for all $f, g \in A$, $\alpha \in [0,1]$ and $x \in X$. ■

## 4.2 The Main Result

The main result of this paper is the demonstration that risk aversion in the RDU model implies ambiguity aversion in the corresponding CEU model.

**Theorem:** *Let $\hat{\succcurlyeq}$ and $\succcurlyeq$ on A have CEU and RDU representations $f \to \int_S u\left(f\left(s\right)\right) d\varphi\left(s\right)$ and $f \to \int_S u\left(f\left(s\right)\right) d\left(\zeta\circ\pi\right)\left(s\right)$, respectively and suppose that $\int_S u\left(f\left(s\right)\right) d\varphi\left(s\right) = \int_S u\left(f\left(s\right)\right) d\left(\zeta\circ\pi\right)\left(s\right)$. If $\succcurlyeq$ displays (strict) risk aversion then $\hat{\succcurlyeq}$ displays (strict) ambiguity aversion.*

*Proof.* By the hypothesis and (4) and (5),

$$\int_S u\left(f\left(s\right)\right) d\varphi\left(s\right) = \int_X u\left(f\left(s\right)\right) d\left(\zeta\circ\pi\right)\left(E_f\left(x\right)\right) = \int_X u\left(x\right) d\left(\zeta\circ G_f\right)\left(x\right). \tag{11}$$

Corollary 2 of Chew et al (1987) asserts that a necessary condition for $\succcurlyeq$ to display (strict) risk aversion is that $\zeta$ is (strictly) concave. Thus, if $\succcurlyeq$ displays (strict) risk aversion then

$$\int_X u\left(x\right) d\left(\zeta\circ\left(\alpha G_f + (1-\alpha) G_g\right)\right)\left(x\right) \geq (>)\, \alpha\int_X u\left(x\right) d\left(\zeta\circ G_f\right)\left(x\right) + (1-\alpha)\int_X u\left(x\right) d\left(\zeta\circ G_g\right)\left(x\right). \tag{12}$$

The convexity of $A$ and the affinity of $u$ imply that

$$u\left(\alpha f\left(s\right) + (1-\alpha) g\left(s\right)\right) = \alpha u\left(f\left(s\right)\right) + (1-\alpha) u(g\left(s\right)),\ \forall s \in S. \tag{13}$$

By the Lemma and the definition of $G_f\left(x\right)$,

$$\int_S \left(\alpha u\left(f\left(s\right)\right) + (1-\alpha) u\left(g\left(s\right)\right)\right) d\varphi\left(s\right) = \int_X u\left(x\right) d\left(\zeta\circ\pi\right)\left(E_{\alpha f+(1-\alpha)g}\left(x\right)\right) \tag{14}$$

$$= \int_X u(x)\, d(\alpha\pi(E_f(x)) + (1-\alpha)\pi(E_g(x))) = \int_X u(x)\, d(\zeta \circ (\alpha G_f + (1-\alpha) G_g))(x),$$

and

$$\alpha \int_S u(f(s))\, d\varphi(s) + (1-\alpha) \int_S u(f(s))\, d\varphi(s) = \alpha \int_X u(x)\, d(\zeta \circ G_f)(x) + (1-\alpha) \int_X u(x)\, d(\zeta \circ G_g)(x). \tag{15}$$

Thus, (11) and (12) implies that

$$\int_S (\alpha u(f(s)) + (1-\alpha) u(g(s)))\, d\varphi(s) \geq (>)\, \alpha \int_S u(f(s))\, d\varphi(s) + (1-\alpha) \int_S u(f(s))\, d\varphi(s). \tag{16}$$

By the Proposition of Schmeidler (1989), $\hat{\succcurlyeq}$ displays (strict) ambiguity aversion if and only if $I(u(f)) := \int_S u(f(s))\, d\varphi(s)$ is concave. Hence, if $\succcurlyeq$ displays risk aversion then $\hat{\succcurlyeq}$ displays ambiguity aversion. ■

### 4.3 Bayesian updating

The issue of updating ambiguous beliefs has been thoroughly studied, and various axiomatic-based updating rules have been proposed both for Choquet expected utility (see, for example, Gilboa and Schmeidler (1993), Eichenberger, Grant and Kelsey (2007)).

The approach taken in this paper suggests a new and natural procedure for updating the capacities by the application of Bayes rule in situations in which the Choquet expected utility has an equivalent rank-dependent utility representation. The idea is as follows: Under the conditions specified in Wakker's (1990), capacities may be expressed as transformations of additive probability measure. Update the additive probability measure using Bayes rule and update the corresponding capacity by setting the capacity of each event equal to the transformation of the updated additive probability of the same event. Formally, let $y^n$ denote a sample of observations. For each event $E$, set $\varphi(E \mid y^n) = \zeta(\pi(E \mid y^n))$, where $\pi(E \mid y^n) = \pi(y^n \mid E)\pi(E) / [\pi(y^n \mid E)\pi(E) + \pi(y^n \mid S \backslash E)(1-\pi(E)]$. In particular, $\varphi(E_f(x) \mid y^n) = \zeta(\pi(E_f(x) \mid y^n) = \zeta \circ G_f(x \mid y^n)$, $x \in X$.

## 5 Ambiguity and Risk Aversion with Second-Order Beliefs

Klibanoff et. al (2005) and Seo (2009) provide alternative axiomatizations of a model of decision making under uncertainty in which ambiguity is expressed by the set of conceivable priors and ambiguity attitudes are captured by a real-valued function on the reals representing the expected utilities of acts under these priors. A decision maker's beliefs about the likelihoods of the priors is represented by a probability measure referred to as *second-order belief.* Formally, a preference relations exhibiting smooth ambiguity aversion has the representation

$$f \mapsto \int_{\Pi} v\left(\int_{S} u\left(f\left(s\right)\right) d\pi\left(s\right)\right) d\Phi\left(\pi\right), \tag{17}$$

where $\Pi := \{\pi \in [0,1]^{|S|} \mid \Sigma_{s\in S}\pi\left(s\right) = 1\}$, is the set of all probability distributions on the set of states, $v$ is a real-valued function on $\mathbb{R}$, $u$ is a real-valued function on $\Delta\left(X\right)$, and $\Phi$ a probability measure on $\Pi$, representing the second order beliefs. In the usual interpretation risk aversion is captured by the concavity of $u$ and ambiguity aversion by that of $v$.

Adopting the approach of the preceding analysis I show below that smooth ambiguity aversion may be interpreted as (extra) layer of risk aversion. To being with, I use the CERA to translate the ambiguity to risk. With this objective in mind, consider the risk represented by a choice of an act, $f$, under a prior $\pi \in \Pi$, and define the certainty equivalent of this risk, denoted $c_f\left(\pi\right)$, as the solution of the equation

$$u\left(c_f\left(\pi\right)\right) = \int_{S} u\left(f\left(s\right)\right) d\pi\left(s\right). \tag{18}$$

Then the representation (17) may be written as

$$f \mapsto \int_{\Pi} \left(v \circ u\right)\left(c_f\left(\pi\right)\right) d\Phi\left(\pi\right). \tag{19}$$

For every $x \in \mathbb{R}$, define $E_f\left(x\right) = \{\pi \in \Pi \mid c_f\left(\pi\right) \leq x\}$ and let $G_f\left(x\right) := \Phi\left(E_f\left(x\right)\right), x \in \mathbb{R}$. Then

$$\int_{\Pi} v\left(\int_{S} u\left(f\left(s\right)\right) d\pi\left(s\right)\right) d\Phi\left(\pi\right) = \int \left(v \circ u\right)\left(x\right) dG_f\left(x\right). \tag{20}$$

Given an act $f$ let $\hat{G}_f$ be a mean-preserving spread of $G_f$, then the decision maker displays risk aversion if and only if $G_f \succcurlyeq \hat{G}_f$ if and only if the composition $v \circ u$ is a concave function. In particular, the concavity of $v$, which presumably captures the decision maker's ambiguity aversion may also be an expression a second layer of risk aversion. Indeed, since $v$ is a concave transformation of $u$, by a theorem of Pratt (1964), $v \circ u$ exhibits greater risk aversion than $u$.

Consider next the effect of the weight of evidence. Returning to the lead example, the utility of betting on red from urn $A$ is

$$V\left(f_R \mid_A\right) = \Sigma_{i=1}^{100}\left(v \circ u\right)\left(c_f\left(\pi_i\right)\right)\left[G_f\left(c_f\left(\pi_i\right) \mid m/2\right) - G_f\left(c_f\left(\pi_{i-1}\right) \mid m/2\right)\right],$$

and that of betting on red from urn $B$ is

$$V\left(f_R\mid_B\right)=\Sigma_{i=1}^{100}\left(v\circ u\right)\left(c_f\left(\pi_i\right)\right)\left[G_f\left(c_f\left(\pi_i\right)\mid n/2\right)-G_f\left(c_f\left(\pi_{i-1}\right)\mid n/2\right)\right].$$

But $n>m$ implies that $G_f\left(\cdot\mid n/2\right)$ is a mean-preserving squeeze of $G_f\left(\cdot\mid m/2\right)$, and (strict) risk aversion implies that $V\left(f_R\mid_B\right)>V\left(f_R\mid_A\right)$. The weight of evidence makes the betting on red from the less ambiguous urn preferable.

# 6 Discussion

Decision makers' preference to base their beliefs (about the likely realization of the events that underlie the risks they are facing) on more information, dubbed ambiguity aversion, is captured by convex capacity in the CEU model. By contrast, risk aversion depicts the decision maker's preference to avoid larger spread of payoffs of the risks, is captured by the concavity of the utility and the probability transformation functions in the RDU models. Generally speaking, ambiguity aversion and risk aversion are distinct concepts as the following example illustrates.

Consider the Ellsberg (1961) two urn experiment mentioned in the introduction. For the sake of simplifying the exposition, consider a decision maker whose utility function is linear and normalized so that $u(\$0)=0$ and $u(\$100)=1$.[10] Suppose that the decision maker faces a choice between betting on the event, $R$, drawing a red ball, and the event $B$, drawing a black ball, from the ambiguous urn (i.e., being paid \$100 if the ball drawn at random from the ambiguous urn is red and \$0 otherwise). According to the CEU model symmetry and ambiguity aversion imply that $\varphi(R)=\varphi(B)<1/2$. According to Wakker (1990), by the additivity of the prior probability measure $\pi\left(\cdot\mid y^0\right)$, we have $\pi\left(R\mid y^0\right)=\pi\left(B\mid y^0\right)=1/2$. By (4), $\varphi(E)=(\zeta\circ\pi)(E)$, $E\in\{A,B\}$.

Ambiguity aversion implies that the decision maker prefers betting on red from the unambiguous urn over betting on red from the ambiguous urn. According the CEU model this is equivalent to $0\varphi(B)+1\varphi(R)<1/2$. For instance, letting $\varphi(R)=\varphi(B)=1/4$ implies, according to Wakker (1990), that $\varsigma\left(\pi\left(B\mid y^0\right)\right)=\varsigma\left(\pi\left(R\mid y^0\right)\right)=\varsigma(1/2)=1/4$ (e.g., $\varsigma(p)=p^2$). According to Chew et al. (1987) the corresponding RDU model exhibits risk proclivity. Indeed, $0\times\varsigma\left(\pi\left(B\mid y^0\right)\right)+1\times\left(1-\varsigma\left(\pi\left(B\mid y^0\right)\right)\right)=3/4$ which is smaller

[10] In the case of RDU theory this corresponds to Yaari's (1987) dual theory. The CEU representation is: $f\mapsto\int_S\left[\Sigma_{x\in X}xf(s)(x)\right]d\varphi(s)$.

than the certainty equivalent of the risk represented by $\pi\left(\cdot \mid y^0\right)$. Ambiguity averse decision maker prefers betting on either color from the ambiguous urn over that from the ambiguous urn whereas the RDU decision maker displays the opposite preferences.[11]

How do we reconcile these examples with the main result of this paper? The answer is to be found in the way we conceive ambiguity. According to the analysis in this paper, a bet on any color form the ambiguous urn is modeled as two-stage compound lottery in which the beliefs about the likely realization of the states in the first stage are represented by a convex capacity measure. Risk aversion in the corresponding RDU model is captured by concave probability transformation function of additive probability measures that ranks the probability of event in the same way that the convex capacity does. The *state-dependent second stage risks* represented by the bets, either or red or black, are replaced by their state-dependent certainty equivalents. For instance, because no ambiguity is associated with the unambiguous urn, the capacity assigned to the event red ball drawn at random form that urn is $\varphi(R) = 1/2$. Consequently, $\pi\left(R \mid y^0\right) = \varphi(R)$. Thus, the second-stage risk of betting on red from the unambiguous urn is reduced (by CERA) to a certainty equivalent outcome equal to \$50 (given the linear utility, $0\pi\left(B \mid y^0\right) + 1\pi\left(R \mid y^0\right) = 1/2 = u(\$50)$). If betting on the color of a ball from the ambiguous urn display ambiguity averse behavior according to the CEU model then, as is shown by the theorem, the corresponding RDU model displays risk-averse behavior. Let $0\pi\left(B \mid y^0\right) + 1\pi\left(R \mid y^0\right) = s/100$, be the expected value of betting on red if the state (i.e., the number of red balls in the ambiguous urn) is $s = 0, 1, ..., 100$, then the certainty equivalent a bet on red, $c(R)$, is given by the equation

$$c(R) = \frac{1}{100}\sum_{s\in 0}^{100} s\left[\zeta\left(\Sigma_{t=0}^{s}\pi(t \mid y^0)\right) - \zeta\left(\Sigma_{t=0}^{s-1}\pi(t \mid y^0)\right)\right] < \$50,$$

where the inequality is an implication of the concavity of $\varsigma$. Thus, given our way of modeling ambiguity there is no contradiction between ambiguity aversion in the CEU model and risk aversion in the corresponding RDU model. Moreover, the models agree on the ranking of bet on the same color from the ambiguous and unambiguous urn.

The analysis of section 5 reveals that, under CERA, risk and ambiguity aversion are fundamentally the same, describing attitudes towards mean-preserving spreads of the ul-

[11] A similar situation in which the CEU decision maker exhibits ambiguity seeking whereas the corresponding RDU decision maker exhibits risk aversion is easy to construct (e.g., let $\zeta(p) = \sqrt{p}$ and $\varphi(R) = \varphi(B) = \sqrt{1/2}$).

timate payoffs. The spread may accounted for by irreducible risks or by risks that are reducible through the acquisition of information. The reduction of compound risks procedures determine whether the attitudes towards sources of the spread are treated as separate factors whose modeling requires distinct theoretical concepts or combined by the CERA procedure and treated as unified notion of risk aversion.

I have long been puzzled by the fact that institutions designed to better allocate risk-bearing (such as insurance and financial markets) are prevalent whereas no institutions appear to improve the allocation of ambiguity. Perhaps the answer is that ambiguity aversion is an aspect of risk aversion and therefore does not require special institutions.

## References


[1] Anscombe, Francis J. and Aumann, Robert J. (1963) "A definition of subjective probability," *Annals of Mathematical Statistics* 43: 199–205.

[2] Chew, Soo Hong (1989) "An Axiomatic Generalization of the Quasilinear Mean and Gini Mean with Applications to Decision Theory," Johns Hopkins University and Tulane University; rewritten version of Chew, S. H. (1985), 'An Axiomatization of the Rank-Dependent Quasilinear Mean Generalizing the Gini Mean and the Quasilinear Mean', Economics Working Paper #156, Johns Hopkins University.

[3] Chew, Soo Hong, and Karni, Edi (1994) "Choquet Expected Utility with Finite State Space: Commutativity and Act-Independence," *Journal of Economic Theroy* 62: 469-479.

[4] Chew, Soo Hong, Karni, Edi and Safra, Zvi (1987) "Risk Aversion in the Theory of Expected Utility with Rank Dependent Probabilities," *Journal of Economic Theory* 42: 370–381.

[5] de Finetti, Bruno (1937) "La prévision: Ses lois logiques, ses sources subjectives," *Annals de l'Institute Henri Poincare,* Vol. 7, 1-68. (English translation, by H. E. Kyburg, appears in H. E. Kyburg and H. E. Smokler (eds.) (1964) *Studies in Subjective Probabilities.* New York. John Wiley and Sons.)

[6] Eichenberger, Jurgen, Grant, Simon and Kelsey, David (2007) "Updating Choquet Beliefs," *Journal of Mathematical Economics*, 43: 888-889.

[7] Ellsberg, Daniel (1961) "Risk, ambiguity and the Savage axioms," *Quarterly Journal of Economics,* 75: 643-669.

[8] Gilboa, Itzhak (1987), "Expected Utility with Purely Subjective Non-Additive Probabilities," *Journal of Mathematical Economics*, 16: 65-88.

[9] Gilboa, Itzhak and Schmeidler, David (1993) "Updating Ambiguous Beliefs," *Journal of Economic Theory* 59: 33-49.

[10] Keynes, John M. (1921) *Treatise on Probability.* London: Macmillan and Co.

[11] Klibanoff, Peter, Marinacci, Massimo and Mukerji, Sujoy (2005): “A Smooth Model of Decision Making Under Ambiguity,” *Econometrica*, 73: 1849–1892.

[12] Nakamura, Yutaka (1990) “Subjective Expected Utility with Non-Additive Probabilities on Finite State Space,” *Journal of Economic Theory* 51: 346-366.

[13] Popper, Karl R. (1958) “A third note on degree of corroboration or confirmation,” *British Journal for the Philosophy of Science*, 8(32): 294–302.

[14] Pratt, John W. (1964) “Risk Aversion in the Small and in the Large,” *Econometrica*, 32: 122-135.

[15] Quiggin, John (1982) “A Theory of Anticipated Utility,” *Journal of Economic Behaviour and Organization* 3: 323–343.

[16] Quiggin, John (1993) *Generalized Expected Utility Theory - The Rank-Dependent Model.* Kluwer Academic Publishers, Dordrecht.

[17] Ramsey, Frank P. (1931) “Truth and Probability,” In *The Foundations of Mathematics and Other Logical Essays.* R. B. Braithwaite and F. Plumpton (Eds.) London: K. Paul, Trench, Truber and Co.

[18] Savage, Leonard. J. (1954) *The Foundations of Statistics.* New York: John Wiley.

[19] Schmeidler, David (1986) “Integral representation without additivity,” *Proceedings of the American Mathematical Society*, 97: 255-261.

[20] Schmeidler, David (1989) “Subjective probability and expected utility without additivity,” *Econometrica*, 57: 571-87. First version dates 1982.

[21] Segal, Uzi (1989) “Anticipated Utility: A Measure Representation Approach,” *Annals of Operations Research* 19: 359–373.

[22] Segal, Uzi (1990) “Two-Stage Lotteries without the Reduction Axiom,” *Econometrica* 58, 349–377.

[23] Segal, Uzi (1993) “The Measure Representation: A Correction,” *Journal of Risk and Uncertainty* 6: 99-107.

[24] Seo, Kyoungwon (2009) "Ambiguity and Second-Order Beliefs," *Econometrica*, 77: 1575–1605.

[25] Wakker, Peter, P. (1989a) "Continuous Subjective Expected Utility with Nonadditive Probabilities," *Journal of Mathematical Economics* 18: 1-27.

[26] Wakker, Peter P. (1989b) *Additive Representations of Preferences, A New Foundation of Decision Analysis*, Kluwer (Academic Publishers), Dordrecht.

[27] Wakker, Peter, P. (1990) "Under Stochastic Dominance Choquet-Expected Utility and Anticipated Utility are Identical," *Theory and Decision* 29: 119-132.

[28] Yaari, Menahem E. (1987) "The Dual Theory of Choice under Risk," *Econometrica* 55: 95–115.

[29] Zappia, Carlo (2020) "Paradox? What Paradox? On a Brief Correspondence Between Leonard Savage and Karl Popper," *Resaerch in the History of Economic Thought and Methodology* 38C, 161-177.