Ambiguity Measurement

Yehuda Izhakian*†

This Version: July 9, 2012

First Version: January 9, 2012

Abstract

Ordering alternatives by their degree of ambiguity is a crucial element in decision-making processes in general and in asset pricing in particular. Thus far the literature has not provided an applicable measure of ambiguity allowing for such ordering. The current paper addresses this need by introducing a novel empirically applicable *ambiguity measure* derived from a new model of decision making under ambiguity in which probabilities of events are themselves random. In this model a complete distinction is attained between preferences and beliefs and between risk and ambiguity that enables the degree of ambiguity to be measured. A merit of the model is that ambiguous probabilities can be incorporated into asset prices and an *ambiguity premium* can be measured empirically.

Keywords: Ambiguity Measure, Ambiguity Aversion, Ambiguity Premium, Choquet Expected Utility, Cumulative Prospect Theory, Ellsberg Paradox, Knightian Uncertainty, Random Probabilities.

JEL Classification Numbers: C44, D81, D83, G11, G12.

*Department of Economics, Stern School of Business, New York University, yud@stern.nyu.edu

†I thank David Backus, Adam Brandenburger, Luís Cabral, Itamar Drechsler, Ignacio Esponda, Xavier Gabaix, Sergiu Hart, Ruth Kaufman, Peter Klibanoff, Ilan Kremer, Sujoy Mukerji, Efe Ok, David Schmeidler, Laura Veldkamp, Paul Wachtel, Stanley Zin and especially Menachem Brenner, Itzhak Gilboa, Mark Machina and Thomas Sargent for valuable discussions and suggestions. I would also like to thank the seminar participants at New York University, The Interdisciplinary Center (IDC) Herzliya, The Hebrew University of Jerusalem, University of Colorado and Norwegian School of Business. A previous version of this paper was circulated under the title "Shadow Probability Theory and Ambiguity Measurement."

1 Introduction

How should uncertain alternatives be ranked by the criterion of ambiguity? Consider the following example: a *large* urn contains 30 balls which are either black or yellow and a second *smaller* urn contains only 10 balls which are also either black or yellow. In the large urn there are at least 16 black balls, while in the small urn there are at least 5 black balls. Which of the following two bets is more ambiguous? "A ball drawn from the *large* urn is black" or, "A ball drawn from the *small* urn is black." Say you were offered \$10 if a ball drawn from the large urn is black and \$12 if a ball drawn from the small urn is black. Which of these two bets would you choose? Answering questions of these types is part of almost any real-life decision. They imply that decision-making involves the ordering of alternatives by their degree of ambiguity. However, so far the literature has not provided a useable measure of ambiguity that allows for such ordering. The goal of this paper is to provide a theoretical basis and an applicable measure that address this need.

This paper makes three contributions to the existing literature. The first and main contribution is that it introduces a novel, empirically applicable, *ambiguity measure*, underpinned by a new theoretical concept.[1] The second contribution is that it presents a decision-making model to derive this measure. This model completely distinguishes preferences from beliefs and risk from ambiguity, thereby allowing the degree of ambiguity to be measured independently of preferences, particularly in empirical studies. It also allows aspects of preferences concerning ambiguity to be monitored in behavioral studies. Unlike previous decision-making models which resemble cases with only a single Ellsberg urn, this model resembles cases involving multiple urns. The third contribution is that it generalizes classical asset pricing theory to incorporate ambiguity, providing a well-defined ambiguity premium which is clearly distinguished from risk and can be tested empirically.

Assuming that probabilities of events are themselves random, this paper introduces a novel model of decision making, referred to as *expected utility with random probabilities* (henceforth *EURP*), which aims to capture the multi-dimensional nature of uncertainty. In this model, there are two tiers of uncertainty: one with respect to consequences and the other with respect to the probabilities of these consequences. Each tier is modeled by a separate state space. This structure introduces a complete distinction of risk from ambiguity with regard to both beliefs and preferences. The degree of risk and attitudes toward it are then measured with respect to

[1] Throughout this paper the term *ambiguity measure* is used informally.

one space, while ambiguity and ambiguity attitudes are measured with respect to the second space. Ambiguity and preferences concerning it are applied directly to probabilities of events independently of their outcomes.

The main idea of EURP is that perceived probabilities are formed by the "certainty equivalent" probabilities of objective random probabilities.[2] That is to say, the perceived probability is the unique probability value that the DM is willing to accept in exchange for the uncertain probability of a given event. As a consequence of probabilistic sensitivity (i.e., the nonlinear ways in which individuals may interpret probabilities), perceived probabilities are nonadditive. Ambiguity aversion results in a subadditive subjective probability measure, while ambiguity loving results in a superadditive measure. It is her perception of probabilities that ultimately guides her decision-making process.

A natural merit of EURP is that, like measuring the degree of risk by the variance of outcomes, the degree of ambiguity can be measured by the variance of probabilities.[3] However, concerning the variance of probabilities, the question is: to the probability of which event is the variance applied? This paper proves that the degree of ambiguity can be measured by four times the variance of the probability of loss, which is equal to four times the variance of the probability of gain. Formally, our measure of ambiguity is given by

$$\mho^2 = 4\mathrm{Var}\left[\mathrm{P}_L\right] = 4\mathrm{Var}\left[\mathrm{P}_G\right],$$

where P_L and P_G are the random probabilities of loss and gain, respectively, and the variance is taken with respect to second-order probabilities. The intuition behind this new measure is that ambiguity is caused by a perturbation of probabilities with respect to a meaningful reference point. Its main advantage is that it can be computed from the data and can be used in empirical tests.

Measuring the degree of ambiguity allows alternatives to be ranked by the criterion of ambiguity. It provides a way to address important questions that arise regarding the nature of ambiguity, in general, and the nature of the aggregate ambiguity of portfolios, in particular. The nature of ambiguity and the relationship between risk and ambiguity may shed some light on various puzzling financial phenomena. Notable examples are the fact that individuals tend to hold very small portfolios, 3-4 stocks (Goetzmann and Kumar (2008)), the equity premium puzzle (Mehra and Prescott (1985)), the risk-free rate puzzle (Weil (1989)), the phenomenon

[2] In this paper the terms perceived probabilities and subjective probabilities are used interchangeably.

[3] Measuring risk by the variance of outcomes is admissible under some conditions; the same is true for measuring ambiguity by the variance of probabilities.

of the observed equity volatility being too high to be justified by changes in the fundamental (Shiller (1981)) and the home bias puzzle (Coval and Moskowitz (1999)).

To demonstrate the value of EURP and its measure of ambiguity, this paper generalizes asset pricing theory to incorporate ambiguity. Relaxing the assumption that probabilities are known, the price of an asset in our model is determined not only by its degree of risk and the DM's attitude toward risk, but also by its degree of ambiguity and the DM's attitude toward ambiguity. The current paper constructs an *uncertainty premium* and proves that it can be separated into a *risk premium* and an *ambiguity premium*.[4] It provides a well-defined ambiguity premium, completely distinguished from risk and attitude toward risk, and which can be computed from the data. This model has been tested empirically by Brenner and Izhakian (2011), who show that ambiguity, measured by $\mho^2$, has a significant impact on the market portfolio return. We are not aware of any prior study that conducts direct empirical tests of models of decision making under ambiguity other than through parametric fitting and calibrations.[5]

EURP relies on the *Choquet expected utility* (CEU) of Schmeidler (1989), whose axiomatic derivation paved the way for modeling decision making under ambiguity. Gilboa (1987) and Schmeidler (1989), in their pioneering studies, introduce the idea that, in the presence of ambiguity, the probabilities that reflect the DM's willingness to bet cannot be additive, i.e., the sum of the probabilities can be either smaller or greater than 1. EURP combines the concept of nonadditive probabilities with the idea of *reference-dependent beliefs*. Reference dependency is applied to differentiate between the probability of gain and the probability of loss. It allows preferences concerning ambiguity which pertain to these probabilities to be different for losses and for gains.

Tversky and Kahneman's (1992) *cumulative prospect theory* (CPT) also applies a two-sided CEU to gains and to losses.[6] However, CPT focuses on *reference-dependent preferences*. It assumes a DM having different risk attitudes for losses and for gains and asymmetric capacities with different arbitrary weighting schemes for losses and for gains.[7] EURP shows that capacities

[4]This paper uses the term *uncertainty* to describe the aggregation of risk and ambiguity.

[5]Uppal and Wang (2003), Epstein and Schneider (2008), and Ju and Miao (2012), for example calibrate their model to the data. Several papers attribute different explanatory variables to ambiguity. For example, Anderson *et al.* (2009) and Drechsler (2012) attribute the disagreement of professional forecasters to ambiguity.

[6]Using the perception of rank-dependent and cumulative functionals proposed by Weymark (1981), Quiggin (1982), Yaari (1987) and Schmeidler (1989), CPT generalizes the original prospect theory of Kahneman and Tversky (1979) from risk to uncertainty. It modifies the probability weighting functionals of the original prospect theory, such that it always satisfies stochastic dominance and supports an infinite state space.

[7]Capacities are (subjective) nonadditive probabilities. This paper uses the term probability in a broad sense,

are not arbitrary and can be explained by the presence of ambiguity and preferences concerning it.[8] It relies on the axiomatic foundation proposed by Wakker (2010) for both risk and ambiguity preferences. EURP stems from the multiple priors paradigm (Gilboa and Schmeidler (1989)) and results in a two sided variation of CEU (Gilboa (1987) and Schmeidler (1989)).

The rest of the paper is organized as follows. Section 2 establishes the setup. Section 3 introduces a new decision-making model under ambiguity and characterizes DMs' attitudes toward ambiguity. Using this model, Section 4 suggests a measure of ambiguity. To demonstrate an application of EURP for asset pricing, Section 5 models the ambiguity premium and reviews an empirical methodology to test it. Section 6 discusses our results with respect to the related literature, and Section 7 concludes.

2 The setup

EURP assumes two different tiers of uncertainty, one with respect to consequences and the other with respect to the probabilities of these consequences. Each tier is modeled by a separate state space. Uncertainty with respect to consequences (outcomes) is modeled in a *primary outcome space*, while uncertainty with respect to probabilities is modeled in a *secondary probability space*.

2.1 Illustration

To illustrate the idea behind EURP about the nature of ambiguity and the way individuals perceive it, consider an Ellsberg urn with 90 colored balls, 30 of which are red and the other 60 either black or yellow. Drawing (at random) a red ball (R) entitles the DM to a sum of \$0, a yellow ball (Y) entitles her to a sum of \$1, and a black ball (B) entitles her to a sum of \$2.

In terms of EURP, the primary space is defined by the states of drawing different balls, i.e., $\{R, Y, B\}$, and the secondary space is defined by the composition of the urn. The probabilities of drawing different balls are thus defined by this composition. The probability of R is exactly $\frac{1}{3}$, while the probability of B can be one of the values $\frac{0}{90}, \frac{1}{90}, \frac{2}{90}, \ldots, \frac{60}{90}$. The precise probability of B is determined by a second-order unobservable event in the secondary space. Such an event can be, for example, "The experimenter put 30 red balls, 20 black balls and 40 yellow balls in the urn." The DM does not have any information indicating which of the possible probabilities is more likely, and thus she assigns an equal weight to each possibility.

i.e., it can be nonadditive and either subjective or objective.

[8] To explain capacities, EURP does not assume asymmetric risk attitude, different ambiguity preferences for losses and for gains, or loss aversion.

Assume a DM whose preferences concerning ambiguity are given by $\Gamma(\mathrm{P}) = \sqrt{\mathrm{P}}$, and her preferences concerning risk are given by $\mathrm{U}(c) = 1 - e^{-c}$. While making decisions, the DM first forms her perceived probabilities as the certain probabilities she is willing to accept in exchange for the undertrain probabilities. For example, the perceived probability of B, derived from the equally likely objective probabilities $\frac{0}{90}, \frac{1}{90}, \frac{2}{90}, \ldots, \frac{60}{90}$, would be

$$\mathrm{Q}(B) = \left(\frac{1}{61}\sum_{i=0}^{60}\sqrt{\frac{i}{90}}\right)^2 = 0.29.$$

Similarly, the perceived probability of RY (red or yellow) would be $\mathrm{Q}(RY) = \left(\frac{1}{61}\sum_{i=30}^{90}\sqrt{\frac{i}{90}}\right)^2 = 0.65$. Given these probabilities, the DM then assesses the value of the urn using Choquet expected utility:

$$\begin{aligned}
\mathrm{V}(B) &= \mathrm{Q}(R)\left(1-e^{-0}\right) + \left[\mathrm{Q}(RY) - \mathrm{Q}(R)\right]\left(1-e^{-1}\right) + \left[\mathrm{Q}(RYB) - \mathrm{Q}(RY)\right]\left(1-e^{-2}\right) \\
&= \frac{1}{3}\left(1-e^{-0}\right) + \left[\left(\frac{1}{61}\sum_{i=30}^{90}\sqrt{\frac{i}{90}}\right)^2 - \frac{1}{3}\right]\left(1-e^{-1}\right) + \left(1 - \frac{1}{61}\sum_{i=30}^{90}\sqrt{\frac{i}{90}}\right)^2\left(1-e^{-2}\right) \\
&= 0.5.
\end{aligned}$$

Consider now the well-known Ellsberg (1961) experiment in which the subjects are asked to choose between R and B and then to choose between RY and BY. In this experiment individuals usually prefer R over B, but BY over RY.[9] Let the prize of winning a bet be \$1 and otherwise \$0, and assume a DM who considers strictly positive outcomes as a gain and otherwise as a loss. These preferences coincide with the expected utility of each bet, obtained by applying the method above separately for gains and losses. Table 1 is a stylized description of this example.

	R	Y	B	Prob	Q	E	Var	$\mho^2$	V
(R)	\$1	\$0	\$0	$\frac{1}{3}$	0.33	$\$\frac{1}{3}$	$\frac{1}{3}$	0	0.21
(B)	\$0	\$0	\$1	$\frac{0}{90}, \frac{1}{90}, \frac{2}{90}, \ldots, \frac{60}{90}$	0.29	$\$\frac{1}{3}$	$\frac{1}{3}$	0.1530	0.18
(RY)	\$1	\$1	\$0	$\frac{30}{90}, \frac{31}{90}, \frac{32}{90}, \ldots, \frac{90}{90}$	0.65	$\$\frac{2}{3}$	$\frac{1}{3}$	0.1530	0.41
(BY)	\$0	\$1	\$1	$\frac{2}{3}$	0.67	$\$\frac{2}{3}$	$\frac{1}{3}$	0	0.42

Table 1: **The Ellsberg example**

The choices observed in the Ellsberg experiment also coincide with the ordering provided

[9] In expected utility theory, the DM's assessments of the likelihoods of R, B and Y can be described by some probability measure P. The DM is assumed to prefer a greater chance of winning \$1,such that the choices above imply that $\mathrm{P}(R) > \mathrm{P}(B)$ and $\mathrm{P}(B \cup Y) > \mathrm{P}(R \cup Y)$. However, since R, B and Y are mutually exclusive events, no such probability measure exists; hence, it is considered a paradox.

by the measure of ambiguity, $\mho^2$. The degree of ambiguity of B, measured by utilizing the variance of the probability of gain from this bet, is $\mho^2[B] = 0.1530$. The probability of gain attained by the alternative bet, R, is exactly $\frac{1}{3}$, implying a zero degree of ambiguity, $\mho^2[R] = 0$. Since the expected outcomes (E) of R and B are identical, an ambiguity-averse DM prefers R, which has the lower degree of ambiguity, over B. The probability of gain from RY can take one of the values $\frac{30}{90}, \frac{31}{90}, \ldots \frac{90}{90}$, which in turn implies $\mho^2[RY] = 0.1530$. The probability of gain from the alternative BY is exactly $\frac{2}{3}$, which implies $\mho^2[BY] = 0$. Therefore, an ambiguity-averse DM prefers BY over RY. Note that the variance of outcomes, computed using expected probabilities, is identical for all the alternative bets.

2.2 The primary space

Observable events and their consequences are defined by the primary space. The probabilities of these events are uncertain; as a consequence, perceived probabilities are nonadditive. Nonadditive priors are also assumed in the foundation of Schmeidler's (1989) CEU and Tversky and Kahneman's (1992) CPT. The primary space relies on this foundation, while the reasoning for nonadditive priors is provided later by the secondary space.[10]

Let $\mathbb{S}$ be a (finite or infinite) nonempty state space, called the *primary space*, endowed with a σ-algebra, $\mathcal{E}$, of subsets of $\mathbb{S}$.[11,12] Generic elements of this σ-algebra are called *primary events* (*events* for short) and are denoted by E. Define X to be a nonempty convex set of *consequences*, where, since this paper mostly deals with monetary outcomes, consequences are confined to real numbers, $X \subseteq \mathbb{R}$.

A *primary act* (*first-order act*) is a measurable function from states into consequences, $f : \mathbb{S} \to X$, describing the resulting consequence associated with each state $s \in \mathbb{S}$. The set of all primary acts is denoted $\mathscr{F}$. Restricting attention to simple measurable acts, a primary act $f \in \mathscr{F}_0$ is represented as a sequence of pairs

$$f = (E_1 : x_1, \cdots, E_j : x_j, \cdots, E_n : x_n),$$

where x_j is the consequence if event E_j occurs, and $(E_1, \ldots, E_n)$ is a generic partition of the state space $\mathbb{S}$. Sometimes, when the context is clear, the primary act f with a vector of outcomes $(x_1, \ldots, x_n)$ is referred to as a random variable, possibly without specified probabilities, and

[10] CEU and CPT do not characterize the sources shaping the nonadditive priors.

[11] To simplify the exposition, whenever possible our results are proved in static discrete settings; however all of the presented results can be applied to dynamic continuous settings.

[12] Following Wakker and Tversky (1993), Wakker (2010, Appendix G) and Kothiyal *el al.* (2011), the state space, $\mathbb{S}$, can consist of an infinite number of states.

designated $x_j = f(s_j)$ by f_j.

The notation $f = (E_1 : x_1, \cdots, E_n : x_n)$ presupposes that the consequences $x_1, \ldots, x_n$ are listed in a non-decreasing order. All consequences $x \in X$ are interpreted either as a gain or as a loss with respect to the *reference point* $x_k \in X$, which is common to all primary acts.[13] Any consequence $x_j \in X$ is a *loss* if $x_j \leq x_k$ and a *gain* if $x_k < x_j$. The cumulative events of loss and gain are thus defined by $L = E_1 \cup \cdots \cup E_k$ and $G = E_{k+1} \cup \cdots \cup E_n$, respectively. To shorten notations, the convention $E_{j \cdots t} = E_j \cup \cdots \cup E_t$ is used to denote cumulative events.[14]

A *capacity* Q is a function on $2^{\mathbb{S}}$ assigning each event $A \subseteq \mathbb{S}$ with a number $\mathrm{Q}(A)$, satisfying $\mathrm{Q}(\emptyset) = 0$, $\mathrm{Q}(\mathbb{S}) = 1$, and if $A \subset B \subset \mathbb{S}$ then $0 \leq \mathrm{Q}(A) \leq \mathrm{Q}(B)$.[15] The capacity of any loss event $1 \leq j \leq k$ is defined by $\mathrm{Q}(E_1 \cup \cdots \cup E_j) - \mathrm{Q}(E_1 \cup \cdots \cup E_{j-1})$, and the capacity of any gain event $k+1 \leq j \leq n$ is defined by $\mathrm{Q}(E_j \cup \cdots \cup E_n) - \mathrm{Q}(E_{j+1} \cup \cdots \cup E_n)$, where $E_0 = E_{n+1} = \emptyset$. Note that, since Q can be nonadditive, the capacities do not necessarily sum up to 1.

The domain of first-order preference relation, $\succsim^1$, is the set of primary acts, $\mathscr{F}_0$, and the relations $\precsim^1$, $\prec^1$, $\succ^1$ and $\sim^1$ are defined as usual. A primary act yielding the same consequence for any state $s \in \mathbb{S}$ is called a *constant act* and is designated by its constant consequence $x \in X$. The *certainty equivalent* (CE) of a primary act $f \in \mathscr{F}_0$ is a constant act, $x \in \mathscr{F}_0$, such that $f \sim^1 x$.

Let $\mathrm{v} : \mathscr{F}_0 \to \mathbb{R}$ be a function assigning to each primary act f a value

$$\mathrm{v}(f) \equiv \sum_{j=1}^{k} \left[\mathrm{Q}(E_{1 \cdots j}) - \mathrm{Q}(E_{1 \cdots j-1})\right] \mathrm{U}(f_j) + \sum_{j=k+1}^{n} \left[\mathrm{Q}(E_{j \cdots n}) - \mathrm{Q}(E_{j+1 \cdots n})\right] \mathrm{U}(f_j), \quad (1)$$

where $\mathrm{U} : X \to \mathbb{R}$ is a strictly increasing continuous *utility function* satisfying $\mathrm{U}(x_k) = 0$. As usual, U characterizes the DM's preferences concerning risk: a concave function implies risk aversion, and a convex function implies risk loving. Similarly to Equation (1), the value of a primary act with an infinite support is defined by[16]

$$\mathrm{v}(f) \equiv -\int_{-\infty}^{k} \mathrm{Q}\left(\left\{s \in \mathbb{S} \mid \mathrm{U}(f_s) < t\right\}\right) dt + \int_{k}^{\infty} \mathrm{Q}\left(\left\{s \in \mathbb{S} \mid \mathrm{U}(f_s) > t\right\}\right) dt. \quad (2)$$

Assuming that the preference relation $\succsim^1$ on the set of primary acts $\mathscr{F}_0$ satisfies Wakker's (2010, Theorem 12.3.5) axiomatization, $\succsim^1$ can be represented by the function v such that

[13] It is common to assume that the reference point is the status quo, exogenously given.

[14] Similarly, this notational convention is applied to other variables, such as $\mathrm{P}_{j \cdots t} = \mathrm{P}(E_j \cup \cdots \cup E_t)$.

[15] Capacities can be applied directly to consequences without specifying an underlying state space, i.e., $\mathrm{Q}(\underline{c} \leq x \leq \overline{c})$, where $\underline{c}$ and $\overline{c}$ are constants; see Jaffray (1989).

[16] In the case of an infinite support, notations are abused and k stands for x_k.

$\mathrm{v}(f) \geq \mathrm{v}(g)$ *iff* $f \succsim^1 g$, for any $f, g \in \mathscr{F}_0$.[17] The function v applies a two-sided Choquet integration to gains and to losses (relative to the reference point). The representation defined by v and axiomatized in Wakker (2010, Theorem 12.3.5) differs from the Choquet integral in that the latter satisfies a "shift" axiom (i.e., $\int_{\mathbb{S}} (f + c)\, dQ = \int_{\mathbb{S}} f dQ + c$ for every f and c), whereas the former is sensitive to shifts of outcomes crossing the reference point. Note that in both representations lower capacities imply lower values of acts, regardless of whether the capacities are of losses or of gains.

2.3 The secondary space

Probabilities over the primary space are assumed to be random and determined by unobservable events in a separate latent state space, referred to as the *secondary space*. While making her choice, the DM does not know which event will be realized, either in the primary space or in the secondary space. That is, she knows neither the realized outcome nor the realized probabilities of outcomes.

Objective probabilities P of events $E \in \mathcal{E}$ occurring in the primary space are random and determined by secondary events in a (finite or infinite) nonempty state space Ω, called the *secondary space*. A *secondary event*, denoted D, is a generic element of a σ-algebra, $\mathcal{D}$, of subsets of Ω, and $(D_1, \ldots, D_m)$ denotes a generic partition of the secondary space Ω. A consequence of D_i is an *additive* probability measure P_i over the primary space $\mathbb{S}$, where P_i stands for $\mathrm{P}\left(\cdot \mid D_i\right)$. $\mathbb{P}$ denotes the set of consequential probability measures. In a finite primary space the probability measure P_i takes the form of a probability vector $\mathrm{P}_i = (\mathrm{P}_{i,1}, \ldots, \mathrm{P}_{i,j}, \ldots, \mathrm{P}_{i,n})$, assigning to each primary event E_j its possible probabilities, where $\mathrm{P}_{i,j}$ stands for $\mathrm{P}\left(E_j \mid D_i\right)$.

A *secondary act* (*second-order act*) is a measurable function from the secondary state space into the set of probability measures, $\hat{f} : \Omega \to \mathbb{P}$, describing the resulting probability measure associated with each secondary state $\omega \in \Omega$. The set of secondary acts is denoted $\widehat{\mathscr{F}}$. Restricting attention to simple measurable (with respect to Ω) acts, a secondary act $\hat{f} \in \widehat{\mathscr{F}_0}$ is represented as a sequence of pairs

$$\hat{f} = (D_1 : \mathrm{P}_1, \ldots, D_i : \mathrm{P}_i, \ldots, D_m : \mathrm{P}_m),$$

where P_i is the probability measure of event D_i occurring. In this framework, a (roulette) *lottery* is given by a constant $\hat{f}$, assigning every $D \in \mathcal{D}$ with the same probability measure.

[17] Alternative axiomatizations that could be adopted are Tversky and Kahneman (1992), Wakker and Tversky (1993), Chew and Wakker (1996) and Kothiyal *et al.* (2011).

The projection of a secondary act on $\mathcal{E}$ induces *eventwise secondary acts* which assign each primary event $E_j \in \mathcal{E}$ with its possible probabilities. That is,

$$\hat{f}_j = \left(D_1^j : \mathrm{P}_{1,j}, \ldots, D_i^j : \mathrm{P}_{i,j} \ldots, D_m^j : \mathrm{P}_{m,j}\right),$$

where $\hat{f}_j$ stands for $\hat{f}(E_j)$. This notation presupposes that each event E_j has a different arrangement $\left(D_1^j, \ldots, D_m^j\right)$ of the partition $(D_1, \ldots, D_m)$ such that its probabilities $(\mathrm{P}_{1,j}, \ldots, \mathrm{P}_{m,j})$ are listed in a non-decreasing order. The set of E_j's secondary acts is denoted $\widehat{\mathscr{F}_j}$. When the context is clear, $\hat{f}_j$ is referred to as a secondary act and the index j designates the primary event. In this context, $\hat{f}_j$ can be viewed as a random variable describing the probability P_j of E_j. A secondary act $\hat{f}_j$, assigning E_j with the same probability for any $D \in \mathcal{D}$, is called a *constant secondary act* and is designated by its outcome probability Q_j.

A second-order *additive* probability measure χ on 2^Ω assigns to each secondary event $D \subseteq \Omega$ a probability $\chi(D)$. Figure 1 provides a diagrammatic representation of the two spaces: the primary space and the secondary space, and the relation between them.

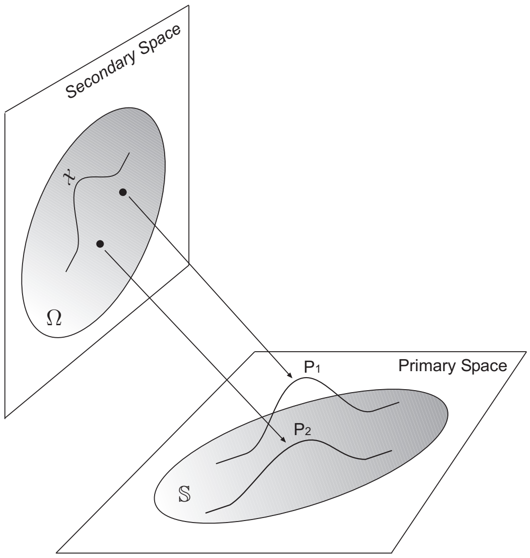

Figure 1: **The two spaces**

The DM is assumed to have a second-order preference relation $\succsim_j^2$, which is a complete order, defined over the set of secondary acts $\widehat{\mathscr{F}_j}$. EURP proposes that $\succsim_j^2$ is utilized solely to form the DM's perceived probability of the primary event E_j. As such, $\succsim_j^2$ can be considered as if it is applied only to secondary acts associated with the same primary act f, i.e., to secondary acts defining the random probability of the same outcome. Given a secondary act $\hat{f}_j$, the preference $\succsim_j^2$ is used only to extract the constant secondary act Q_j which satisfies $\hat{f}_j \sim_j^2 \mathrm{Q}_j$. Each Q_j, referred to as the *perceived probability* (of event E_j), is formed separately. Relying on these perceived probabilities, Q_j, the DM then makes her choices.

2.4 The objects of choice

The objects of choice are acts, denoted $\mathfrak{f}$. An *act* is a measurable function $\mathfrak{f} : \mathbb{S} \times \Pi \to X$, where Π is a set of families of probability distributions, $\mathbb{P}$, and the product space $\mathbb{S} \times \Pi$ is endowed with the product σ-algebra $\mathcal{E} \otimes \mathcal{P}$ of $\mathcal{E}$ and a σ-algebra $\mathcal{P}$ of subsets of Π.[18] The set of all acts is denoted $\mathfrak{F}$. In this framework, an act $\mathfrak{f} \in \mathfrak{F}_0$ is induced by a primary act $f \in \mathscr{F}_0$ and a secondary act $\hat{f} \in \widehat{\mathscr{F}_0}$, where $\hat{f}$ defines the family (subset) of probability measures, $\mathbb{P}$, associated with $\mathfrak{f}$. Thus, an act $\mathfrak{f} \in \mathfrak{F}_0$ can be represented as a sequence of pairs

$$\mathfrak{f} \;=\; ((E_1, \mathbb{P}) : x_1, \cdots, (E_j, \mathbb{P}) : x_j, \cdots, (E_n, \mathbb{P}) : x_n) \,.$$

Given $s \in \mathbb{S}$, we write $\mathfrak{f}_s$ for $\mathfrak{f}\left(s, \hat{f}_s\right)$, where $\hat{f}_s = \mathrm{P}_s = (\mathrm{P}_{1,s}, \ldots, \mathrm{P}_{m,s})$ is the projection of $\mathbb{P}$ on s. The preference $\succsim$, which is a complete order, over the set of acts $\mathfrak{F}_0$ is assumed to be a primitive and is induced by the preferences $\succsim^1$ and $\succsim^2_j$ over $\mathscr{F}_0$ and $\widehat{\mathscr{F}_j}$, respectively.

It might be helpful at this point to illustrate our setup using the following three-color Ellsberg urns. Assume two urns, $\mathfrak{f}$ and $\mathfrak{g}$, each containing 100 balls in red, black and yellow. Drawing a red ball (R) from $\mathfrak{f}$ entitles the DM to a sum of \$0, a yellow ball (Y) entitles her to a sum of \$1, and a black ball (B) entitles her to a sum of \$2. Drawing a red ball (R) from $\mathfrak{g}$ entitles the DM to a sum of \$1, a yellow ball (Y) entitles her to a sum of \$3, and a black ball (B) entitles her to a sum of \$2. It is known that the proportions of the balls R, B, Y in $\mathfrak{f}$ can be either $(40, 20, 40)$ or $(20, 60, 20)$ with equal likelihoods. The former is obtained if the event α occurs and the latter is obtained if the event β occurs. The proportions of balls in $\mathfrak{g}$ can be either $(50, 40, 10)$ or $(10, 40, 50)$, also with equal likelihoods, where the former is obtained if α occurs and the latter is obtained if β occurs. Table 2 summarizes this example.

	Outcome			Probabilities			Distribution	
	R	Y	B	R	Y	B	α	β
$\mathfrak{f}$	\$0	\$1	\$2	$(0.2, 0.4)$	$(0.2, 0.6)$	$(0.2, 0.4)$	$(40, 20, 40)$	$(20, 60, 20)$
$\mathfrak{g}$	\$1	\$3	\$2	$(0.1, 0.5)$	$(0.4, 0.4)$	$(0.1, 0.5)$	$(50, 40, 10)$	$(10, 40, 50)$

Table 2: **Example - objects of choice**

In EURP representation, this example is formulated as follows. The primary state space is $\mathbb{S} = \{R, Y, B\}$, and the secondary state space is $\Omega = \{\alpha, \beta\}$. The primary acts are defined

[18] I thank Mark Machina for his relevant comment.

by $f = (R:0, Y:1, B:2)$ and $g = (R:1, Y:3, B:2)$, and the secondary acts are defined by $\hat{f} = (\alpha : (0.4, 0.2, 0.4), \beta : (0.2, 0.6, 0.2))$ and $\hat{g} = (\alpha : (0.5, 0.4, 0.1), \beta : (0.1, 0.4, 0.5))$. The eventwise secondary acts induced by $\hat{f}$ are thus $\hat{f}_R = (\beta : 0.2, \alpha : 0.4)$, $\hat{f}_Y = (\alpha : 0.2, \beta : 0.6)$ and $\hat{f}_B = (\beta : 0.2, \alpha : 0.4)$, and the eventwise secondary acts induced by $\hat{g}$ are $\hat{g}_R = (\beta : 0.1, \alpha : 0.5)$, $\hat{g}_Y = (\alpha : 0.4, \beta : 0.4)$ and $\hat{g}_B = (\alpha : 0.1, \beta : 0.5)$. The objects of choice (acts) can then be defined by $\mathfrak{f} = ((R, (0.2, 0.4)) : 0, (Y, (0.2, 0.6)) : 1, (B, (0.2, 0.4)) : 2)$ and $\mathfrak{g} = ((R, (0.1, 0.5)) : 1, (Y, (0.4, 0.4)) : 3, (B, (0.1, 0.5)) : 2)$, where for short, we write only the projections of probability measures on events.

3 The decision-making model

The main idea at the basis of EURP is that risk and risk preferences apply to the primary space, whereas ambiguity and ambiguity preferences apply to the secondary space. EURP assumes two differentiated phases in the decision-making process: probability formation and valuation. In the *formation* phase, based on the information she has and her preferences concerning ambiguity, the DM forms a representation of her perceived probabilities for all the events which are relevant to her decision. Then, in the *valuation* phase, based on her preferences concerning risk, the DM assesses the value of each act and chooses accordingly.

3.1 Attitudes toward ambiguity

Attitudes toward ambiguity are defined by preferences over secondary acts in the formation phase. Although the DM does not have an immediate choice over secondary acts, her subjective perception of likelihoods, resulting from an aversion to or love of ambiguity, plays an important role in her decision process. Secondary acts cannot be chosen independently of primary acts; still, in many decision problems second-order preferences can be inferred from observable choices. For example, in the Ellsberg urn experiment, secondary acts may be considered as bets on the composition of the urn. In financial decisions, secondary acts can be viewed as bets about the means, variances and covariances of the investment opportunities. Similarly, secondary acts in model-uncertainty applications can be considered as bets about the true values of the parameters of the model.

Attitudes toward ambiguity are defined by DMs' preferences over random probabilities of primary events and their expected probabilities, i.e., preferences over secondary acts and constant secondary acts, respectively. An ambiguity-averse DM prefers the expectations of the random probabilities over the random probabilities themselves. An ambiguity-loving DM prefers the random probabilities over their expectations, and an ambiguity-neutral DM is indifferent between them. The next definition settles this idea formally.

Definition 1. *Let a secondary act be* $\hat{f}_j = \left(D_1^j : \mathrm{P}_{1,j}, \ldots, D_m^j : \mathrm{P}_{m,j}\right)$ *and its related constant secondary act be* $\bar{\hat{f}}_j = \left(D_1^j : \mathrm{E}\left[\mathrm{P}_j\right], \ldots, D_m^j : \mathrm{E}\left[\mathrm{P}_j\right]\right)$ *, where* $\mathrm{E}\left[\mathrm{P}_j\right] = \sum_{i=1}^m \chi_i \mathrm{P}_{i,j}$ *is the expected probability of event* E_j*, taken with respect to the second-order probabilities* χ*. Ambiguity aversion (loving) as regards event* E_j *is defined by* $\bar{\hat{f}}_j \succsim_j^2 \hat{f}_j$ $(\bar{\hat{f}}_j \precsim_j^2 \hat{f}_j)$*, and ambiguity neutrality is defined by* $\bar{\hat{f}}_j \sim_j^2 \hat{f}_j$*.*[19,20]

A DM is said to be ambiguity averse if $\bar{\hat{f}}_j \succsim_j^2 \hat{f}_j$ for any $E_j \in \mathcal{E}$. Ambiguity-loving DMs and ambiguity-neutral DMs are defined similarly. Definition 1 can be applied to formulate different attitudes toward ambiguity for different subsets of events; for example, ambiguity loving for losses and ambiguity aversion for gains. The next theorem ties preferences concerning ambiguity to a functional representation.

Theorem 1. *Let* $\mathrm{Q} : \widehat{\mathscr{F}_j} \to \mathbb{R}$ *be a function assigning a value* $\mathrm{Q}\left(\hat{f}_j\right)$ *to each secondary act* $\hat{f}_j \in \widehat{\mathscr{F}}_j$*, such that*

$$\mathrm{Q}\left(\hat{f}_j\right) \equiv \Gamma_j^{-1}\left(\sum_{i=1}^m \chi\left(D_i\right) \Gamma_j\left(\mathrm{P}\left(E_j \mid D_i\right)\right)\right), \tag{3}$$

where $\Gamma_j : [0,1] \to \mathbb{R}$ *is a strictly increasing continuous function. Assume that the preference relation,* $\succsim_j^2$*, on the set of acts* $\widehat{\mathscr{F}_j}$ *satisfies: weak ordering, monotonicity, continuity and sign-tradeoff consistency.*[21] *The preference* $\succsim_j^2$ *can then be represented by the function* Q*, such that* $\mathrm{Q}\left(\hat{g}_j\right) \geq \mathrm{Q}\left(\hat{f}_j\right)$ *iff* $\hat{g}_j \succsim_j^2 \hat{f}_j$*, for any* $\hat{f}_j, \hat{g}_j \in \widehat{\mathscr{F}_j}$*.*[22,23]

The function Γ_j, referred to as a *probability-outlook* function (*outlook* function for short), forms the DM's attitude toward ambiguity. As with risk attitudes, there are three types of attitudes toward ambiguity: ambiguity aversion, ambiguity loving and ambiguity neutrality. Ambiguity neutrality takes the form of a linear Γ, ambiguity aversion the form of a concave Γ and ambiguity loving the form of a convex Γ. Two special types of ambiguity preferences can be defined. *Constant relative ambiguity aversion* (CRAA), which takes the functional form

[19] Strict preferences toward ambiguity can be defined similarly.

[20] Epstein (1999) defines ambiguity aversion relative the probabilistically sophisticated order of Machina and Schmeidler (1992).

[21] *Weak ordering*, *monotonicity* and *continuity* are defined as usual. *Sign-tradeoff consistency* sustains if improving a consequence in any indifference relation $\sim$ breaks the relation. Formally, if $(E : a, E^c : x) \sim (E : b, E^c : y)$ and $(E : a', E^c : x) \sim (E : b, E^c : y)$ then $a' = a$.

[22] Alternatively, one can employ an SEU axiomatization.

[23] The proof of Theorem 1 is directly derived from Wakker (2010, Theorem 12.3.5), where probabilities are additive and no distinction of losses from gains is made.

$\Gamma(\mathrm{P}) = \frac{\mathrm{P}^{1-\eta}}{1-\eta}$, and *constant absolute ambiguity aversion* (CAAA), which takes the functional form $\Gamma(\mathrm{P}) = -\frac{e^{-\eta \mathrm{P}}}{\eta}$, where η is the coefficient of ambiguity aversion.

The concept of ambiguity and preferences concerning it, proposed in Theorem 1, suggests a definition of subjective probabilities. The subjective probability $\mathrm{Q}(E_j)$ of event E_j, formulated in Equation (3), provides a nonlinear model for the way that individuals may perceive probabilities, i.e., $\mathrm{Q}(E_j) = \mathrm{Q}\left(\hat{f}_j\right)$. Considering for example an ambiguity-averse DM, Equation (3) implies that a higher aversion to ambiguity (a more concave Γ function) or a higher dispersion of probabilities, both result in lower perceived probabilities which in turn result in a lower value of the act.

The perceived probabilities of an ambiguity-neutral DM are additive and equal to the expected probabilities, i.e., $\mathrm{Q}(E_j) = \mathrm{E}[\mathrm{P}_j]$.[24] The perceived probabilities of an ambiguity-averse DM are lower than the expected probabilities, i.e., $\mathrm{Q}(E_j) < \mathrm{E}[\mathrm{P}_j]$, and result in subadditive probabilities. The perceived probabilities of an ambiguity-loving DM are greater than the expected probabilities, i.e., $\mathrm{Q}(E_j) > \mathrm{E}[\mathrm{P}_j]$, and result in superadditive probabilities. These perceived probabilities provide an explanation for the nonadditive priors of Gilboa (1987) and Schmeidler (1989), which are arbitrary in CEU and CPT.

What meaning do perceived probabilities have? Given an ambiguous lottery and a non-ambiguous lottery with identical prizes for each event such that the DM is indifferent between them, the perceived probabilities are defined by the probabilities of the non-ambiguous lottery. These perceived probabilities can be tested by a simple variation of the Ellsberg three-color experiment. Subjects can be asked to propose an urn with a known proportion of balls which they are willing to accept in place of an urn with an unknown proportion of balls, where the prizes attached to each color are the same in both urns.

3.2 The functional representation

Given an act $\mathfrak{f} : \mathbb{S} \times \Pi \to X$, in the first phase of the decision-making process—the probability forming phase—the DM designs her perceived probabilities. This phase, which reduces Π to a subjective probability measure Q, is carried out using preferences over secondary acts. In the second phase—the valuation phase—the DM uses Q to asses the value of the act. This phase is carried out using preferences over primary acts. The next theorem combines the preferences over primary acts, $\mathscr{F}_0$, and over secondary acts, $\widehat{\mathscr{F}}_j$, introducing a functional representation of

[24] An ambiguity-neutral DM can be considered a DM who reduces two-stage lotteries to compound lotteries in the usual way. For the implications of this type of preferences see, for example, Halevy (2007).

preferences over acts $\mathfrak{F}_0$.

Theorem 2. *Let* $\mathrm{V}:\mathfrak{F}_0\to\mathbb{R}$ *be a function assigning a value*

$$\begin{aligned}\mathrm{V}(\mathfrak{f}) &= \sum_{j=1}^{k}\left[\Gamma_j^{-1}\left(\sum_{i=1}^{m}\chi_i\Gamma_j\left(\mathrm{P}_{i,1\cdots j}\right)\right)-\Gamma_j^{-1}\left(\sum_{i=1}^{m}\chi_i\Gamma_j\left(\mathrm{P}_{i,1\cdots j-1}\right)\right)\right]\mathrm{U}(\mathfrak{f}_j)+ \\ &\quad \sum_{j=k+1}^{n}\left[\Gamma_j^{-1}\left(\sum_{i=1}^{m}\chi_i\Gamma_j\left(\mathrm{P}_{i,j\cdots n}\right)\right)-\Gamma_j^{-1}\left(\sum_{i=1}^{m}\chi_i\Gamma_j\left(\mathrm{P}_{i,j+1\cdots n}\right)\right)\right]\mathrm{U}(\mathfrak{f}_j),\end{aligned} \tag{4}$$

to each act $\mathfrak{f}\in\mathfrak{F}_0$*. Assume that the preference relation* $\succsim$ *over* $\mathfrak{F}_0$ *satisfies Wakker's (2011) axiomatization, and the preference relation* $\succsim_j^2$ *over* $\widehat{\mathscr{F}_j}$ *satisfies the conditions of Theorem 1 for any* $E_j\in\mathcal{E}$*. The preference relation* $\succsim$ *over* $\mathfrak{F}_0$ *can then be represented by the function* V*, such that* $\mathrm{V}(\mathfrak{g})\geq\mathrm{V}(\mathfrak{f})$ *iff* $\mathfrak{g}\succsim\mathfrak{f}$*, for any* $\mathfrak{f},\mathfrak{g}\in\mathfrak{F}_0$*.*

Theorem 2 ties risk preferences to ambiguity preferences when acts have a finite support. Likewise, the value of an act $\mathfrak{f}$ with an infinite support takes the form

$$\begin{aligned}\mathrm{V}(\mathfrak{f}) &= -\int_{-\infty}^{k}\Gamma_t^{-1}\left(\int_{\mathbb{P}}\Gamma_t\left(\mathrm{P}\left(\left\{s\in\mathbb{S}\mid\mathrm{U}(\mathfrak{f}_s)<t\right\}\right)\right)d\chi\right)dt \\ &\quad +\int_{k}^{\infty}\Gamma_t^{-1}\left(\int_{\mathbb{P}}\Gamma_t\left(\mathrm{P}\left(\left\{s\in\mathbb{S}\mid\mathrm{U}(\mathfrak{f}_s)>t\right\}\right)\right)d\chi\right)dt.\end{aligned}$$

The functional representation of the DM's aggregate preferences, proposed in Theorem 2, makes a complete distinction between beliefs and preferences and between risk and ambiguity. First-order beliefs are formed by the random probability measures $\mathrm{P}_{i=1,\ldots m}$; second-order beliefs are formed by the probability measure χ; risk preferences are formed by the utility function U; and ambiguity preferences are formed by the outlook function Γ.

Risk and ambiguity preferences, in Theorem 2, can be different for losses and for gains; their functional representation can take the form $\mathrm{U}=\mathrm{U}^-$ and $\Gamma=\Gamma^-$ for losses, and $\mathrm{U}=\mathrm{U}^+$ and $\Gamma=\Gamma^+$ for gains. Different utility functions, U^- and U^+, can capture, for example, *loss aversion*.[25] Different outlook functions, Γ^- and Γ^+, can capture, for example, ambiguity loving for losses and ambiguity aversion for gains.[26] Pessimism and optimism can also be incorporated into Theorem 2 by the probability weights $\mathrm{w}^-(\mathrm{Q})$ and $\mathrm{w}^+(\mathrm{Q})$ for losses and for gains, respectively. *Pessimism* holds if greater weights are assigned to lower ranked outcomes, and *optimism* holds if greater weights are assigned to higher ranked outcomes. In principle, a DM in this extension can exhibit ambiguity aversion while still being an optimist.

[25]Loss aversion is modeled by a steeper utility function for losses than for gains (see, for example, Barberis and Huang (2001)). Some behavioral studies report risk aversion for gains and risk loving for losses (see, for example, Abdellaoui *et al.* (2008)).

[26]Some behavioral studies document ambiguity loving for losses and ambiguity aversion for gains (see, for example, Bier and Connell (1994)).

Although Theorem 2 grants the DM the flexibility of having different attitudes for different events, henceforth it is assumed that the DM has a consistent attitude toward ambiguity for all events. Thus, her preferences concerning ambiguity are formed by the same outlook function Γ for any primary event $E_j \in \mathcal{E}$.

It is important to note that, while risk considers both consequences and their probabilities, ambiguity considers only probabilities. The type and magnitude of a consequence resulting from a primary event are not relevant to its degree of ambiguity. Consider, for example, an event with an unknown probability of losing \$100. Changing the magnitude of loss to \$1000 does not affect either its perceived probability or its degree of ambiguity.

The separation attained in Theorem 2 allows for studying the distinct impact of each component on values of acts. More importantly, it enables the simplification of this model to an applicable form such that the degree of ambiguity can be measured, as proposed later in Section 4.

3.3 Perceived probabilities

We turn now to study the properties of perceived probabilities, proposed in Theorem 1, and to simplify their functional representation. Let

$$p_j = \mathrm{E}\left[\mathrm{P}_j\right] = \sum_{i=1}^{m} \chi_i \mathrm{P}_{i,j}$$

be the expected probability of event E_j and

$$\zeta_j^2 = \mathrm{Var}\left[\mathrm{P}_j\right] = \sum_{i=1}^{m} \chi_i \left(\mathrm{P}_{i,j} - p_j\right)^2$$

be its variance. The covariance of the probabilities of two events E_j and E_l is defined by

$$\zeta_{j,l} = \mathrm{Cov}\left[\mathrm{P}_j, \mathrm{P}_l\right] = \sum_{i=1}^{m} \chi_i \left(\mathrm{P}_{i,j} - p_j\right)\left(\mathrm{P}_{i,l} - p_l\right).$$

Perceived probabilities Q, defined in Theorem 1, are a function of first-order (random) objective probabilities, second-order objective probabilities and preferences concerning ambiguity. To simplify the exposition of Theorem 2 and Theorem 1, the perceived probability $\mathrm{Q}\left(E_j\right)$ of each event E_j is approximated by taking a second-order Taylor approximation with respect to its first-order probabilities P_j around its expected probability p_j.[27] Since this approximation deals with probabilities, a condition on the outlook function Γ is enforced to assure that the approximated perceived probabilities are nonnegative.

[27]The same method is applied by Arrow (1965) and Pratt (1964) to outcomes within the expected utility framework, whereas in this case it is applied to probabilities.

Theorem 3. *Assume a continuous twice-differentiable outlook function* Γ*, satisfying*

$$\frac{1}{2}\left(\frac{\Gamma''(p_l)}{\Gamma'(p_l)}\zeta_l^2-\frac{\Gamma''(p_{j\cup l})}{\Gamma'(p_{j\cup l})}\zeta_{j\cup l}^2\right)\leq p_j, \tag{5}$$

for any events $E_j, E_l \in \mathcal{E}$*, where* $j \cup l$ *stands for* $E_j \cup E_l$*. For relatively small first-order objective probabilities,* P_j*, the perceived probability of event* E_j *is then*

$$\mathrm{Q}(E_j) \quad \approx \quad p_j + \frac{1}{2}\frac{\Gamma''(p_j)}{\Gamma'(p_j)}\zeta_j^2.$$

Similarly, the perceived probability of a cumulative event $E_{j\cdots t}$ is $\mathrm{Q}(E_{j\cdots t}) \approx p_{j\cdots t}+\frac{1}{2}\frac{\Gamma''(p_{j\cdots t})}{\Gamma'(p_{j\cdots t})}\zeta_{j\cdots t}^2$. Theorem 3 characterizes capacities since $\mathrm{Q}(\emptyset)=0$, $\mathrm{Q}(\mathbb{S})=1$, and $\mathrm{Q}(A)\leq \mathrm{Q}(B)$ if $A\subset B\subset \mathbb{S}$ (by Lemma 4). Condition (5) bounds the level of ambiguity aversion (the concavity of Γ) such that $-2\frac{p_j}{\zeta_j^2}\leq -\frac{\Gamma''(p_j)}{\Gamma'(p_j)}\leq 2\frac{p_j}{\zeta_j^2}$ and $-2\frac{p_j}{\zeta_j^2+2\zeta_{j,l}}\leq -\frac{\Gamma''(p_j)}{\Gamma'(p_j)}\leq 2\frac{p_j}{\zeta_j^2+2\zeta_{j,l}}$. It assures that the approximated probabilities are nonnegative and that the probability of an event is not lower than the probability of any of its sub-events. Condition (5), however, is required only for the purpose of approximation and it is not enforced over the precise perceived probabilities defined in Equation (3). Henceforth it is assumed that the outlook function Γ falls under this condition.

Definition 2. *The expression*

$$\varphi_j \equiv -\frac{1}{2}\frac{\Gamma''(p_j)}{\Gamma'(p_j)}\zeta_j^2$$

is referred to as the ***probability premium*** *of event* E_j*, and* ζ_j^2 *is referred to as the* ***ambiguity of event*** j *(**e-ambiguity** for short). The expression* $-\frac{\Gamma''(p_j)}{\Gamma'(p_j)}$ *is referred to as the coefficient of absolute ambiguity aversion, and* $-p_j\frac{\Gamma''(p_j)}{\Gamma'(p_j)}$ *is referred to as the coefficient of relative ambiguity aversion.*[28]

The probability premium is composed of two components: ambiguity preferences, framed by $\frac{\Gamma''}{\Gamma'}$, and the degree of e-ambiguity, measured by ζ^2. Preferences concerning ambiguity can be aversion ($\frac{\Gamma''}{\Gamma'}<0$), loving ($\frac{\Gamma''}{\Gamma'}>0$) or neutrality ($\frac{\Gamma''}{\Gamma'}=0$), which imply subadditive, superadditive and additive probabilities, respectively. When the probability of an event is perfectly known, its e-ambiguity is zero, $\zeta^2=0$, and so is the probability premium. This is also true when the DM is ambiguity neutral. Considering an ambiguity-averse DM, a higher aversion to ambiguity or a higher degree of e-ambiguity both imply a greater probability premium and a lower perceived probability. Note that by Condition (5) the coefficient of absolute ambiguity aversion is bounded such that for a sufficiently high level of ambiguity aversion the perceived probability of each event tends to zero.

[28]These definitions are equivalent to the Arrow-Pratt coefficient of absolute risk aversion and coefficient of relative risk aversion, respectively.

In general, the subjective probability measure Q is nonadditive, that is $\mathrm{Q}\left(E_{l\cdots t}\cup E_{t+1\cdots j}\right)\neq \mathrm{Q}\left(E_{l\cdots t}\right)+\mathrm{Q}\left(E_{t+1\cdots j}\right)$. This measure has an additive component—the expected probability $p\left(E_{l\cdots t}\cup E_{t+1\cdots j}\right)=p\left(E_{l\cdots t}\right)+p\left(E_{t+1\cdots j}\right)$, and an additional nonadditive component—the probability premium. The source of nonadditivity is e-ambiguity, measured by $\zeta^2_{l\cdots j}$, which in most cases satisfies $\zeta^2_{l\cdots j}\neq\zeta^2_{l\cdots t}+\zeta^2_{t+1\cdots j}$.[29] The probabilities of events are not independent. Therefore, the ambiguity of a union of events comprises the covariances between the probabilities of its sub-events, as the next proposition implies.

Lemma 1. *For any $1\leq l<t<j\leq n$, e-ambiguity satisfies*

$$\zeta^2_{l\cdots j} \quad = \quad \zeta^2_{l\cdots t}+\zeta^2_{t+1\cdots j}+2\zeta_{l\cdots t,t+1\cdots j}.$$

The ambiguity of event E and its complement event E^c is a special case since the e-ambiguity of their union equals zero, i.e., $\zeta^2_{\mathbb{S}}=0$.

Lemma 2. *The covariance between the probability of event E and the probability of its complement event E^c satisfies*

$$\mathrm{Cov}\left[\mathrm{P}\left(E\right),\mathrm{P}\left(E^c\right)\right]=-\mathrm{Var}\left[\mathrm{P}\left(E\right)\right]=-\mathrm{Var}\left[\mathrm{P}\left(E^c\right)\right]$$

and thus $\zeta^2_{\mathbb{S}}=0$.

The results of Lemma 1 and Lemma 2 coincide with the findings of *support theory* of Tversky and Koehler (1994) and Rottenstreich and Tversky (1997). Support theory documents that the judged probability of an event generally increases when its description is unpacked into disjoint components and decreases by unpacking its alternative description. One may conclude from Lemma 1 that unpacking an event into disjoint components which satisfy $\zeta_{1\cdots t,t+1\cdots j}<0$ increases its probability as perceived by ambiguity-averse DMs.

3.4 Simplified representation

The perceived probabilities, proposed in Theorem 3, provide a way to simplify the functional representation of preferences over acts to a more usable form. The value of an act $\mathfrak{f}\in\mathfrak{F}_0$, formed in Theorem 2, can be written

$$\mathrm{V}\left(\mathfrak{f}\right) \quad \approx \quad \sum_{j=1}^{k}\left[p_j-\varphi_{1\cdots j}+\varphi_{1\cdots j-1}\right]\mathrm{U}\left(\mathfrak{f}_j\right)+\sum_{j=k+1}^{n}\left[p_j-\varphi_{j\cdots n}+\varphi_{j+1\cdots n}\right]\mathrm{U}\left(\mathfrak{f}_j\right), \qquad (6)$$

[29] One can write $\phi\left(E_j,E_l\right)=\frac{1}{2}\left(\frac{\Gamma''\left(p_{j\cup l}\right)}{\Gamma'\left(p_{j\cup l}\right)}\zeta^2_{j\cup l}-\frac{\Gamma''\left(p_l\right)}{\Gamma'\left(p_l\right)}\zeta^2_l-\frac{\Gamma''\left(p_j\right)}{\Gamma'\left(p_j\right)}\zeta^2_j\right)$ to obtain the *Möbius transform* of two events (see, for example, Chateauneuf and Jaffray (1989) and Grabisch *et al.* (2000)).

where $\varphi_{j\cdots t} = -\frac{1}{2}\frac{\Gamma''(p_{j\cdots t})}{\Gamma'(p_{j\cdots t})}\zeta^2_{j\cdots t}$. When an attitude toward ambiguity is different for gains and for losses, the coefficient of ambiguity attitude takes the form $-\frac{\Gamma''_L(p_{1\cdots j})}{\Gamma'_L(p_{1\cdots j})}$ for any loss event $1 \leq j \leq k$ and $-\frac{\Gamma''_G(p_{j\cdots n})}{\Gamma'_G(p_{j\cdots n})}$ for any gain event $k+1 \leq j \leq n$.

Assume, for example, an ambiguity-averse DM who exhibits CAAA. Since p is additive and for any event $-\frac{\Gamma''}{\Gamma'} = \eta$, the value function in Equation (6) can be simplified to

$$\mathrm{V}(\mathfrak{f}) \approx \sum_{j=1}^{n} p_j \mathrm{U}(\mathfrak{f}_j) - \frac{1}{2}\eta \sum_{j=1}^{k} \left[\zeta^2_{1\cdots j} - \zeta^2_{1\cdots j-1}\right] \mathrm{U}(\mathfrak{f}_j) - \frac{1}{2}\eta \sum_{j=k+1}^{n} \left[\zeta^2_{j\cdots n} - \zeta^2_{j+1\cdots n}\right] \mathrm{U}(\mathfrak{f}_j).$$

The first component of the value function is the conventional expected utility. The last two components are the disutility caused by the presence of ambiguity. If the DM's preferences concerning ambiguity are CRAA, then the value of an act $\mathfrak{f}$ is

$$\mathrm{V}(\mathfrak{f}) \approx \sum_{j=1}^{n} p_j \mathrm{U}(x_j) - \frac{1}{2}\eta \sum_{j=1}^{k} \left[\frac{\zeta^2_{1\cdots j}}{p_{1\cdots j}} - \frac{\zeta^2_{1\cdots j-1}}{p_{1\cdots j-1}}\right] \mathrm{U}(\mathfrak{f}_j) - \frac{1}{2}\eta \sum_{j=k+1}^{n} \left[\frac{\zeta^2_{j\cdots n}}{p_{j\cdots n}} - \frac{\zeta^2_{j+1\cdots n}}{p_{j+1\cdots n}}\right] \mathrm{U}(\mathfrak{f}_j).$$

In both cases, if the DM is ambiguity neutral, no disutility occurs and $\mathrm{V}(\mathfrak{f}) = \sum_{j=1}^{n} p_j \mathrm{U}(\mathfrak{f}_j)$.

4 Ambiguity measurement

Almost any real-life decision is concerned with ambiguity. One of the first steps of any decision-making process is to rank the different alternatives by their degree of ambiguity. The key to addressing this need is a well-defined measure of ambiguity.[30]

4.1 Ordering ambiguous events

A preliminary step in ordering acts by their degree of ambiguity is to define an order over primitive events. Such an order, induced by the DM's preferences, can be defined as follows.

Definition 3. *Let the random probability of event $E_j \in \mathcal{E}$ have the same expected probability under acts $\mathfrak{f} \in \mathfrak{F}_0$ and $\mathfrak{g} \in \mathfrak{F}_0$, i.e., $p_{\mathfrak{f},j} = p_{\mathfrak{g},j}$.*[31] *Event E_j is more ambiguous under act $\mathfrak{f}$ than under act $\mathfrak{g}$ if $\hat{g}_j \succsim^2_j \hat{f}_j$ by any ambiguity-averse DM.*

This definition provides a subjective ordering that arises from preferences. Recall that preferences concerning ambiguity apply only to the probabilities of events and not to their outcomes. A higher perceived probability (capacity) of an event, resulting from lower ambiguity, implies a higher value of an act regardless of whether this event is associated with a loss or a gain. An objective ordering can be defined as follows.[32]

[30] Jewitt and Mukerji (2011) investigate the ranking of ambiguous acts as revealed by the DM's preferences.
[31] The notations $p_{\mathfrak{f},j}$ and $\zeta^2_{\mathfrak{f},j}$ stand for the expected probability and the e-ambiguity of event E_j under act $\mathfrak{f}$.
[32] Rothschild and Stiglitz (1970) apply a similar idea for risk with respect to outcomes.

Definition 4. *Event $E_j \in \mathcal{E}$ is more ambiguous under act $\mathfrak{f} \in \mathfrak{F}_0$ than under act $\mathfrak{g} \in \mathfrak{F}_0$ if there exists a random variable ϵ such that*

$$\mathrm{P}_{\mathfrak{f},j} - p_{\mathfrak{f},j} =_d \mathrm{P}_{\mathfrak{g},j} - p_{\mathfrak{g},j} + \epsilon,$$

where $=_d$ means equal in distribution and $\mathrm{E}\left[\epsilon \mid \mathrm{P}_{\mathfrak{g},j}\right] = \mathrm{E}\left[\epsilon\right] = 0$.[33]

The subjective ordering in Definition 3 coincides with the objective ordering in Definition 4, as the following theorem asserts.

Theorem 4. *When event $E_j \in \mathcal{E}$ has the same expected probability under acts $\mathfrak{f} \in \mathfrak{F}_0$ and $\mathfrak{g} \in \mathfrak{F}_0$, then Definitions 3 and 4 of the more ambiguous event coincide.*

Events can be ordered in a third way, by their degree of ambiguity, measured by ζ^2.

Definition 5. *Event $E_j \in \mathcal{E}$ is more ambiguous under act $\mathfrak{f} \in \mathfrak{F}_0$ than under act $\mathfrak{g} \in \mathfrak{F}_0$ if $\zeta^2_{\mathfrak{f},j} \geq \zeta^2_{\mathfrak{g},j}$.*

This definition implies that the higher the fluctuation of the probability of an event, the greater its ambiguity. Ordering events by ζ^2 coincides with the ordering of Definitions 3 and 4 if probabilities are equable-symmetrically distributed or if the DM's attitude toward ambiguity is quadratic or of the CAAA type. Formally, the probability of event E_j is said to be *equable-symmetrically distributed* if it satisfies $\mathrm{P}_{s+i+1,j} - \mathrm{P}_{s+i,j} = \Delta$ and $\chi_{s-i} = \chi_{s+i}$, $\forall i = -s, \ldots, s$, where s is the point of symmetry.

Theorem 5. *Given an event $E_j \in \mathcal{E}$ which has the same expected probability under acts $\mathfrak{f} \in \mathfrak{F}_0$ and $\mathfrak{g} \in \mathfrak{F}_0$, Definitions 3, 5 and 4 are equivalent if one of the following conditions holds:*

(i) The probabilities of E_j are equable-symmetrically distributed under $\mathfrak{f}$ and under $\mathfrak{g}$;

(ii) The DM's attitude toward ambiguity is of the CAAA type;

(iii) The DM's attitude toward ambiguity is quadratic.

Henceforth, it is assumed that the probabilities of all events are equable-symmetrically distributed. If needed, this assumption can be replaced by assuming CAAA or a quadratic outlook function, Γ. At this point, the order of eventwise secondary acts by e-ambiguity is well-defined. This order induces only a *partial order* on the set of acts $\mathfrak{F}_0$. E-ambiguity induces a *total order* on subsets of $\mathfrak{F}_0$ which satisfy the following condition.

Definition 6. *Let acts $\mathfrak{f}, \mathfrak{g} \in \mathfrak{F}_0$, whose probabilities are equable-symmetrically distributed, have the same expected probabilities. Act $\mathfrak{f}$ is more ambiguous than act $\mathfrak{g}$ if for any event $E_j \in \mathcal{E}$ any ambiguity-averse DM prefers $\hat{g}_j$ over $\hat{f}_j$.*

[33] The condition $\mathrm{E}\left[\epsilon \mid \mathrm{P}_{\mathfrak{g},j}\right] = \mathrm{E}\left[\epsilon\right]$ means that ϵ is mean-independent of the random probability $\mathrm{P}_{\mathfrak{g},j}$ of E_j under $\mathfrak{g}$, i.e., a mean-preserving spread.

Assuming aversion to ambiguity, Definition 6 states that if for any event the random probability associated with act $\mathfrak{g}$ is preferred to the random probability associated with act $\mathfrak{f}$, then $\mathfrak{f}$ is more ambiguous than $\mathfrak{g}$. This definition together with Theorem 5 imply that if for any event $E_j \in \mathcal{E}$ the e-ambiguity of $\mathfrak{g}$ is not higher than the e-ambiguity of $\mathfrak{f}$, i.e., $\zeta^2_{\mathfrak{f},j} \geq \zeta^2_{\mathfrak{g},j}$, then $\mathfrak{f}$ is more ambiguous than $\mathfrak{g}$. In other words, act $\mathfrak{f}$ is more ambiguous than act $\mathfrak{g}$ if the probabilities associated with it are consistently more volatile than the probabilities associated with act $\mathfrak{g}$.

Together, Definition 6 and Theorem 5 define *first-order stochastic dominance with respect to ambiguity*. Definition 6 can be rephrased as follows: act $\mathfrak{f}$ is more ambiguous than act $\mathfrak{g}$ if $\mathfrak{f}$ is *stochastically dominated* by $\mathfrak{g}$ *with respect to ambiguity*. Figure 2 illustrates two acts, where $\mathfrak{f}$ is stochastically dominated by $\mathfrak{g}$. The values on the y-axis are the degrees of ambiguity of the cumulative events lying on the x-axis. Note that the degree of ambiguity of the entire state space, $\mathbb{S}$, and the empty event, $\emptyset$, is always zero for every act.

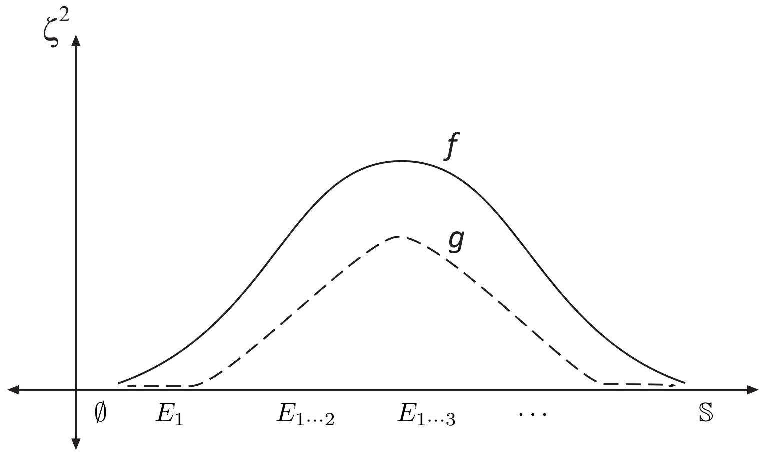

Figure 2: **First order stochastic dominance with respect to ambiguity**

4.2 The ambiguity measure

The ambiguity measure introduced in this section considers symmetric acts. Formally, an act $\mathfrak{f}$ is said to be a *symmetric act* around the point of symmetry x_s if its probabilities are equable-symmetrically distributed and it satisfies $x_s - x_{s-j} = x_{s+j} - x_s$ and $\mathrm{P}_{i,s-j} = \mathrm{P}_{i,s+j}$, $\forall j = -s, \ldots s$ and $\forall i = 1, \ldots m$.[34] The ambiguity measure utilizes the cumulative probability of loss and the cumulative probability of gain. It relies on Lemma 2, which implies that the variances of these two probabilities are identical, that is, $\mathrm{Var}\left[\mathrm{P}_L\right] = \mathrm{Var}\left[\mathrm{P}_G\right]$. The next theorem proposes one of the main results of this paper: a new measure of ambiguity. It asserts that the

[34]When measuring ambiguity, the more restrictive assumption of normally distributed outcomes which allows measuring risk by variance, can be relaxed to symmetric distribution.

degree of ambiguity embedded in an act can be measured by twice the sum of the variance of its cumulative probability of loss and the variance of its cumulative probability of gain.

Theorem 6. *The **degree of ambiguity** of a symmetric act* $\mathfrak{f}$*, denoted* $\mho^2$*, can be measured by*

$$\mho^2[\mathfrak{f}] \equiv 2\mathrm{Var}[\mathrm{P}_L] + 2\mathrm{Var}[\mathrm{P}_G] = 4\mathrm{Var}[\mathrm{P}_L]. \tag{7}$$

The normalized measure (to the units of probability) is defined by $\mho[\mathfrak{f}] \equiv 2\sqrt{\mathrm{Var}[\mathrm{P}_L]}$.

The ambiguity measure $\mho^2$ induces a total order on acts that satisfy first-order stochastic dominance with respect to ambiguity. Theorem 7 below proves that this order coincides with the order provided by a DM. The idea that ambiguity, formed by probability fluctuations, should be measured with respect to a meaningful reference point stands behind the construction of $\mho^2$. Since decisions are concerned with potential loss, a natural reference point is the outcome which distinguishes losses from gains.

Assume, for example, that strictly positive outcomes are considered a gain in the well-known Ellsberg three-color experiment (see Section 2.1). The probability of drawing a black (B) ball (gain) can be one of the values $\frac{0}{90}, \frac{1}{90}, \frac{2}{90}, \ldots, \frac{60}{90}$, each with equal likelihood. This implies that the normalized degree of ambiguity (to units of probability) of this bet is $\mho[B] = 0.3912$. Assume that instead of 60 black and yellow balls in an unknown proportion, the urn contains only 30 black and yellow balls in an unknown proportion, in addition to 30 red balls. The degree of ambiguity of the bet on B decreases to $\mho[B] = 0.2981$. If the quantity of balls which may be black or yellow in an unknown proportion is 90, then the degree of ambiguity of the bet on B increases to $\mho[B] = 0.4377$. If the urn contains only 60 balls, all of them black or yellow in an unknown proportion, then the degree of ambiguity, which in this case is identical for black and yellow, is $\mho = 0.5868$. If there are only 10 balls in the urn (and again the proportion of black and yellow is unknown), then $\mho = 0.6324$. Lastly, if there is only one ball in the urn, of unknown color, then $\mho = 1$, and in the other extreme case, if there is an infinite number of balls in the urn, then $\mho = \frac{1}{\sqrt{3}}$. Table 3 is a stylized description of these variations.

The minimal possible degree of ambiguity, $\mho^2 = 0$, is attained when all probabilities are perfectly known. The maximal possible degree of ambiguity, $\mho^2 = 1$, is attained when the probability of loss (or gain) is either 0 or 1 with equal odds. In this most extreme case, the variance of the probability of loss attains its maximal possible value, $\frac{1}{4}$. Variances of probabilities are therefore normalized by 4 to provide an ambiguity measure ranging between 0 and 1. Note that $\mho^2$ depends on a reference point, x_k, which determines the set of gain

#Balls					
Total	R	Y	B	P_G	$\mho[B]$
60	30	$0,\ldots,30$	$0,\ldots,30$	$\frac{0}{60},\frac{1}{60},\ldots,\frac{30}{60}$	0.2981
90	30	$0,\ldots,60,$	$0,\ldots,60$	$\frac{0}{90},\frac{1}{90},\ldots,\frac{60}{90}$	0.3912
120	30	$0,\ldots,90$	$0,\ldots,90$	$\frac{0}{120},\frac{1}{120},\ldots,\frac{90}{120}$	0.4377
∞	0	$0\ldots\infty$	$0\ldots\infty$	$0\ldots1$	0.5773
60	0	$0,\ldots,60$	$0,\ldots,60$	$\frac{0}{60},\frac{1}{60},\cdots\frac{60}{60}$	0.5868
10	0	$0,\ldots,10$	$0,\ldots,10$	$\frac{0}{10},\frac{1}{10},\cdots\frac{10}{10}$	0.6324
1	0	$0,1$	$0,1$	$0,1$	1

Table 3: **Degrees of ambiguity**

outcomes and the set of loss outcomes. If $x_k = \min(x)$ or $x_k = \max(x)$, i.e., outcomes are considered either all as gain or all as loss, then the degree of ambiguity equals zero. If there is a reference point agreed upon by all DMs, $\mho^2$ can be considered an *objective* measure of ambiguity; otherwise it is considered a *subjective* measure. Concerning financial assets, for example, the risk-free rate of return can possibly be an objective reference point agreed upon by all financial DMs.

The measure of ambiguity, $\mho^2$, takes into account the impact of the correlations between probabilities across events. One can define the *absolute degree of ambiguity* of act $\mathfrak{f}$ by $\widehat{\mho}^2[\mathfrak{f}] \equiv \sum_{j=1}^{n} \mathrm{Var}[\mathrm{P}_j]$.[35] Great caution should be exercised when using $\widehat{\mho}^2$—probabilities are almost always correlated such that $\widehat{\mho}^2$ is biased in the sense that it ignores these correlations. In other words, this measure disregards an important piece of information concerning the nature of probabilities and, as a result, the nature of ambiguity.

The point to emphasize is that the measure of ambiguity $\mho^2$ is not affected by the magnitude of outcomes, in general, and the magnitude of loss or gain, in particular. Increasing or decreasing the outcomes of an act (and accordingly the reference point) does not change its degree of ambiguity, but it does change its degree of risk. A decision-making process considers not only the degree of ambiguity but also the degree of risk. Hence, when making choices these two factors jointly play a role. A consolidated *uncertainty measure* that aggregates risk and ambiguity, proposed by Izhakian (2012),

$$\Upsilon(\mathfrak{f}) \equiv \sqrt{\frac{\mathrm{Var}[\mathfrak{f}]}{1-\mho^2[\mathfrak{f}]}}.$$

[35]One may consider $\sum_{j=1}^{k} p_{1\cdots j}\left[\zeta_{1\cdots j}^2 - \zeta_{1\cdots j-1}^2\right] + \sum_{j=k+1}^{n} p_{j\cdots n}\left[\zeta_{j\cdots n}^2 - \zeta_{j+1\cdots n}^2\right]$ as the *relative degree of ambiguity* (with respect to expected probabilities).

To prove that $\mho^2$ measures ambiguity, it has to be shown that the ordering of acts by $\mho^2$ coincides with the ordering provided by a DM. To inspect the impact of ambiguity, the ordering is made over acts with identical properties except for the degree of ambiguity. That is, they have the same outcomes and the same expected probability (implying the same risk) such that the only difference between them is the dispersion of probabilities around their expected probabilities. To eliminate the effect of risk preferences, the DM is assumed to be risk neutral. Since acts are assumed to be symmetric, and thus the point of symmetry x_s equals the expected outcome, it is assumed that the reference point satisfies $x_k \leq x_s$; otherwise, the DM will not consider these acts.

Theorem 7. *Assume symmetric acts $\mathfrak{f}, \mathfrak{g} \in \mathfrak{F}_0$, satisfying first-order stochastic dominance with respect to ambiguity, having the same outcomes and the same expected probability for each event, i.e., $\mathrm{E}\left[\mathrm{P}_{\mathfrak{f},j}\right] = \mathrm{E}\left[\mathrm{P}_{\mathfrak{g},j}\right]$, $\forall j = 1, \ldots, n$. Act $\mathfrak{f}$ is more ambiguous than act $\mathfrak{g}$, i.e., $\mho^2\left[\mathfrak{g}\right] \leq \mho^2\left[\mathfrak{f}\right]$, iff any ambiguity-averse DM, with a reference point $x_k \leq x_s$, prefers $\mathfrak{g}$ to $\mathfrak{f}$.*

This theorem ties $\mho^2$ to preferences concerning ambiguity. It proves that if two acts are identical except in their degree of ambiguity, then any ambiguity-averse DM prefers the act with the lower $\mho^2$ over the act with the higher $\mho^2$.

5 Implementation for asset pricing

One of the advantages of EURP is that it achieves a complete separation of ambiguity from risk and beliefs from preferences. To demonstrate this merit, this section presents an application of the theory to asset pricing. The prices that financial decision makers (investors) are willing to pay for assets could be affected by the fact that they do not know the precise probabilities of returns. They might require a premium for bearing ambiguity in addition to the premium they require for bearing risk.

The *risk premium* is the premium that a DM is willing to pay for exchanging a risky bet for its expected outcome. The *ambiguity premium* is the premium she is willing to pay for exchanging an ambiguous bet for a risky but non-ambiguous bet that has an identical expected outcome. The *uncertainty premium* is the total premium that a DM is willing to pay for exchanging an ambiguous bet for its expected outcome, i.e., the accumulation of the risk premium and the ambiguity premium. The uncertainty premium, denoted $\mathcal{K}$, can be defined by

$$\mathrm{U}\left(\mathrm{E}\left[x\right]-\mathcal{K}\right)\approx\sum_{j=1}^{k}\left[p_j-\left(\varphi_{1\cdots j}-\varphi_{1\cdots j-1}\right)\right]\mathrm{U}\left(x_j\right)+\sum_{j=k+1}^{n}\left[p_j-\left(\varphi_{j\cdots n}-\varphi_{j+1\cdots n}\right)\right]\mathrm{U}\left(x_j\right), \qquad (8)$$

where $CE=\mathrm{E}\left[x\right]-\mathcal{K}\in\mathfrak{F}_0$ is the *certainty equivalent*, satisfying $CE\sim\mathfrak{f}$. That is, it is the constant act for which the DM is willing to exchange a risky and ambiguous (uncertain) act. The next theorem approximates the uncertainty premium and separates it into a risk premium and an ambiguity premium.

Theorem 8. *Assume a DM whose reference point x_k is relatively close to zero, satisfying $0\leq x_k\leq\mathrm{E}\left[x\right]$, and her preferences are characterized by a twice-differentiable utility function, U, and a twice-differentiable outlook function, Γ. For relatively small outcomes with relatively small probabilities the uncertainty premium is*

$$\mathcal{K}\ \approx\ \underbrace{-\frac{1}{2}\frac{\mathrm{U}''\left(\mathrm{E}\left[x\right]\right)}{\mathrm{U}'\left(\mathrm{E}\left[x\right]\right)}\mathrm{Var}\left[x\right]}_{\mathcal{R}}\underbrace{-\frac{1}{8}\left[\frac{\Gamma''\left(\mathrm{E}\left[\mathrm{P}_L\right]\right)}{\Gamma'\left(\mathrm{E}\left[\mathrm{P}_L\right]\right)}+\frac{\Gamma''\left(\mathrm{E}\left[\mathrm{P}_G\right]\right)}{\Gamma'\left(\mathrm{E}\left[\mathrm{P}_G\right]\right)}\right]\mathrm{E}\left[x\right]\mho^2\left[x\right]}_{\mathcal{A}}, \qquad (9)$$

where $\mathcal{R}$ is the risk premium and $\mathcal{A}$ is the ambiguity premium.

This theorem provides two distinctions. First, it distinguishes between risk and ambiguity premiums such that these two premiums are orthogonal. Second, within each premium it distinguishes between sources of premiums, preferences and beliefs.

Concerning financial decisions, outcomes can be described by rates of return, r. In this case, the uncertainty premium, in terms of return, takes the form

$$\mathcal{K}\ \approx\ -\frac{1}{2}\frac{\mathrm{U}''\left(\mathrm{E}\left[r\right]\right)}{\mathrm{U}'\left(\mathrm{E}\left[r\right]\right)}\mathrm{Var}\left[r\right]-\frac{1}{8}\left[\frac{\Gamma''\left(\mathrm{E}\left[\mathrm{P}_L\right]\right)}{\Gamma'\left(\mathrm{E}\left[\mathrm{P}_L\right]\right)}+\frac{\Gamma''\left(\mathrm{E}\left[\mathrm{P}_G\right]\right)}{\Gamma'\left(\mathrm{E}\left[\mathrm{P}_G\right]\right)}\right]\mho^2\left[r\right]. \qquad (10)$$

The risk premium, $\mathcal{R}\approx-\frac{1}{2}\frac{\mathrm{U}''(\mathrm{E}[r])}{\mathrm{U}'(\mathrm{E}[r])}\mathrm{Var}\left[r\right]$, is the Arrow-Pratt risk premium, where $\mathrm{Var}\left[r\right]$ is computed using expected probabilities. Independently, a higher risk or a higher aversion to it, measured by the coefficient of absolute risk aversion $-\frac{\mathrm{U}''}{\mathrm{U}'}$, result in a greater risk premium.

The ambiguity premium, $\mathcal{A}\approx-\frac{1}{8}\left[\frac{\Gamma''(\mathrm{E}[\mathrm{P}_L])}{\Gamma'(\mathrm{E}[\mathrm{P}_L])}+\frac{\Gamma''(\mathrm{E}[\mathrm{P}_G])}{\Gamma'(\mathrm{E}[\mathrm{P}_G])}\right]\mho^2\left[r\right]$, possesses attributes resembling those of the risk premium, but with respect to probabilities rather than to consequences. A complete separation between beliefs about random probabilities, measured by $\mho^2$, and preferences concerning it, measured by the coefficient of absolute ambiguity aversion, $-\frac{\Gamma''}{\Gamma'}$, is achieved. Ambiguity aversion ($-\frac{\Gamma''}{\Gamma'}>0$) implies a positive ambiguity premium. Ambiguity loving ($-\frac{\Gamma''}{\Gamma'}<0$) implies a negative premium. Ambiguity neutrality ($-\frac{\Gamma''}{\Gamma'}=0$) implies a zero premium, obtained also when probabilities are perfectly known ($\mho^2=0$). Independently, a higher degree of ambiguity or a higher ambiguity aversion result in a greater ambiguity premium. As an example, the next corollary shows the different premiums in the case of a DM typified by CRRA and CAAA.

Corollary 1. *Assuming a DM who is characterized by CRRA,*[36]

$$\mathrm{U}\left(r_j\right)=\begin{cases}\frac{r_j^{1-\gamma}-r_k^{1-\gamma}}{1-\gamma}, & \gamma\neq 0\\ \ln\left(r_j\right)-\ln\left(r_k\right), & \gamma=0\end{cases}, \quad \textit{and CAAA,} \quad \Gamma\left(\mathrm{P}_{i,j}\right)=-\frac{e^{-\eta \mathrm{P}_{i,j}}}{\eta},$$

then the uncertainty premium is $\quad \mathcal{K}\approx\gamma\frac{1}{2}\mathrm{Var}\left[r\right]+\eta\frac{1}{4}\mho^2\left[r\right].$

When the expected probabilities of loss and gain are relatively close to $\frac{1}{2}$, the ambiguity premium can be simplified such that the *expected return* (equity premium) is

$$\mathrm{E}\left[r\right] \quad \approx \quad r_f-\frac{1}{2}\frac{\mathrm{U}''\left(\mathrm{E}\left[r\right]\right)}{\mathrm{U}'\left(\mathrm{E}\left[r\right]\right)}\mathrm{Var}\left[r\right]-\frac{1}{4}\frac{\Gamma''\left(\mathrm{E}\left[\mathrm{P}_L\right]\right)}{\Gamma'\left(\mathrm{E}\left[\mathrm{P}_L\right]\right)}\mho^2\left[r\right], \tag{11}$$

where r_f stands for the risk-free rate of return.

The pricing model of Equation (11) has been tested empirically by Brenner and Izhakian (2011). They assume a pricing representative investor whose reference point is the risk-free rate. Their study employs the measure of ambiguity, $\mho^2$, as an explanatory factor of the aggregate return on the stock market. It assumes that market return is normally distributed with random parameters, mean and variance, which govern its distribution. Using the S&P 500 intraday data, they extract the monthly degree of ambiguity by utilizing the following four-step methodology. The first step is to sample 20 to 22 groups, each comprising 26 observations (15-minutes returns) from the monthly data. The second step is to compute the mean and variance of each group. The third step is to compute the probability of a return lower than the risk-free rate (loss) for each group, using its mean and variance. At this point, for each month there are 20-22 probabilities of loss. The last step is to compute the variance of these probabilities to obtain the monthly degree of ambiguity $\mho^2$.

Brenner and Izhakian (2011) conduct a series of tests to study the explanatory power of $\mho^2$ on the stock market return. They show that ambiguity has a significant impact on returns in both contemporaneous and prediction testing, which means that monthly return on the stock market is not only affected by the ambiguity in the same month, but also by the ambiguity in the previous month.

[36] A more standard formulation of CRRA, $\mathrm{U}\left(r\right)=\frac{r^{1-\gamma}}{1-\gamma}$ for $\gamma\neq 1$ and otherwise for $\gamma=1$ $\mathrm{U}\left(r\right)=\ln\left(r\right)$, is not always normalized to $\mathrm{U}\left(r_k\right)=0$.

6 Related Literature

Since the seminal works of Knight (1921) and Ellsberg (1961), utility theory research has made a concerted effort to model decision processes under uncertainty and to explain the violation of expected utility theory. This effort has generated the ideas that, in the presence of ambiguity, the DM's beliefs take the form of either multiple priors or a single but nonadditive prior. In their *max-min expected utility with multiple priors* model, Gilboa and Schmeidler (1989) assert that an ambiguity-averse DM possesses a set of priors and evaluates her ex-ante welfare conditional upon the worst prior. The *subjective nonadditive* probabilities of Gilboa (1987), the CEU of Schmeidler (1989) and the CPT of Tversky and Kahneman (1992) state that uncertainty and aversion to it can be represented by a single subadditive prior. EURP's contribution to this literature is twofold. First, it achieves a complete separation between ambiguity and attitudes toward it. Second, it proves that capacities are not arbitrary and can be explained by the presence of ambiguity and the DM's preferences concerning it.

The concept of modeling attitudes toward ambiguity by relaxing the reduction between first-order and second-order probabilities, suggested by Segal (1987), inspires other models: Klibanoff *et al.*'s (2005) smooth model of ambiguity, its generalization to include intertemporal substitution, proposed by Ju and Miao (2012) and Hayashi and Miao (2011), and the second-order beliefs of Nau (2006), Chew and Sagi (2008), Ergin and Gul (2009) and Seo (2009). Unlike these models, in which ambiguity preferences are taken with respect to expected utilities or certainty equivalents, in EURP ambiguity preferences are applied solely to probabilities. This approach distinguishes between the effects of risk and the effects of ambiguity, such that it allows for the measurement of the degree of ambiguity.

EURP can be interpreted as a model of robustness in the presence of *model uncertainty*. This class of models assumes an uncertainty about the true probability law governing the realization of states, and a DM, with her concerns about misclassification, looks for a robust decision-making process; see, for example, Hansen *et al.* (1999), Hansen and Sargent (2001) and Maccheroni *et al.* (2006). Ambiguity in this line of models is formulated by the deviation of probability from a reference probability (reference model), measured by *relative entropy*.[37] EURP is also related to Siniscalchi's (2009) *vector expected utility*, which assumes a baseline probability and different sources of ambiguity with respect to expected utility. Other models

[37]Relative entropy is the expected log Radon-Nikodym derivative. Technically, all alternative models have to be absolutely continuous with respect to the reference model for an entropy measure to exist.

that consider reference expected utility include those of Roberts (1980) and Quiggin *et al.*(2004), for example, or consider a reference prior: Einhorn and Hogarth (1986), and Gajdos *et al.* (2008), for example. Kopylov's (2006) ϵ-contamination suggests the addition of an element of confidence to the generated set of priors. Chateauneuf *et al.* (2007) suggest new capacities (neo-additives) obtained from a set of priors generated by ϵ-contamination. All these models require the identification of a reference prior while EURP requires only the identification of a reference point, which is easier to obtain, especially in empirical and experimental studies.

It is important to note that EURP differs from all of the models mentioned above in one major aspect: ambiguity and preferences concerning it are applied directly to probabilities and not to any element of utility. That is, they are not applied to expected utility, certainty equivalent or eventwise utility, all of which are derived from risk preferences. In EURP no need arises to identify a reference probability distribution. It provides a formal way to compare the choices of two DMs who have different attitudes toward ambiguity or different degrees of ambiguity—for example, two DMs who share the same information and the same attitude toward risk but have different levels of ambiguity sensitivity, or two DMs who share the same attitude toward risk and ambiguity but possess different information (different degrees of ambiguity). The ability to conduct this type of comparative static is of primary importance, as it allows for the identification of the pure effect of introducing ambiguity and attitude toward it into a model.

Several approaches to estimating ambiguity have been suggested in the literature. Dow and Werlang (1992) measure uncertainty as the sum of the probability of an event and the probability of its complement event. Ui (2011) measures ambiguity by the difference between the minimal possible mean and the true mean. Maccheroni *et al.* (2011) measure ambiguity by the variance of an unknown mean. Bewley (2011) and Boyle *et al.* (2011) measure ambiguity by a critical confidence interval. All these studies assume that the variance of consequences is known. Our ambiguity measure, $\mho^2$, is broader; it assumes an unknown variance and allows all other parameters that characterize probabilities to be unknown.

The implications of ambiguity regarding the equity premium have been studied mainly by focusing on theoretical aspects. Chen and Epstein (2002), Izhakian and Benninga (2011), and Ui (2011) add an ambiguity premium to the conventional risk premium. Maccheroni *et al.* (2011) adjust the mean-variance paradigm for ambiguity to extract an ambiguity premium.[38] Epstein and Schneider (2007) employ the max-min model to study portfolio choice under ambiguity,

[38]Segal and Spivak (1990) also analyze the ambiguity premium, which they call a *premium of order 2*.

and Epstein and Schneider (2008) employ it to show that the ambiguity premium depends on the idiosyncratic risk in fundamentals.[39] Unlike Theorem 8, these papers do not attain a complete separation between preferences and beliefs and between risk and ambiguity.

7 Conclusion

In reality almost any decision involves ambiguity. It is natural to look for a simple measure of ambiguity that allows for ordering uncertain alternatives by their degree of ambiguity. The current paper satisfies this need by providing an *ambiguity measure* which can be employed in empirical studies. Brenner and Izhakian (2011), for example, use this measure to inquire into whether stock prices are affected by ambiguity. Their empirical study shows that ambiguity has a significant impact on stock returns. To the best of our knowledge, this study is the first to measure ambiguity from market data, rather than in laboratory experiments or calibrations.

To construct a useful measure of ambiguity, the paper introduces a novel model of decision making under ambiguity. This model assumes that probabilities of observable events are random and determined by second-order unobservable events, modeled by two separate state spaces. This structure allows for a complete distinction between risk and ambiguity and between preferences and beliefs. The degree of risk and the decision maker's attitudes toward it are then measured with respect to one space, while the degree of ambiguity and attitude toward it are measured with respect to the second space. Ambiguity and preferences concerning it are applied directly to probabilities of events independently of their outcomes. In this model, perceived probabilities are framed by the nonlinear ways in which individuals may interpret probabilities. Perceived probabilities are nonadditive: ambiguity aversion results in a subadditive subjective probability measure, while ambiguity loving results in a superadditive measure.

In this paper, ambiguity takes the form of probability perturbation with respect to a reference point that distinguishes losses from gains. This concept provides a natural ambiguity measure, which has proved to be empirically testable. The measure of ambiguity, $\mho^2$, is simply four times the variance of the probability of loss (or gain). The present paper demonstrates the merits of this measure by incorporating ambiguous probabilities into the basic asset pricing model. It generalizes the Arrow-Pratt theory and clearly differentiates between the *risk pre-*

[39] Mukerji and Tallon (2001) study the consequences of ambiguity aversion for idiosyncratic risk and risk diversification.

mium and the *ambiguity premium*, which can both be measured empirically. Our measure of ambiguity can be applied in empirical and behavioral studies in finance and economics.

References

Abdellaoui, M., A. Attema, and H. Bleichrodt (2010) "Intertemporal Tradeoffs for Gains and Losses: An Experimental Measurement of Discounted Utility," *Economic Journal*, Vol. 120, No. 545, pp. 845–866.

Abramowitz, M. and I. A. Stegun (1972) *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, New York: Dover, 9th edition.

Anderson, E. W., E. Ghysels, and J. L. Juergens (2009) "The Impact of Risk and Uncertainty on Expected Returns," *Journal of Financial Economics*, Vol. 94, No. 2, pp. 233–263.

Arrow, K. J. (1965) *Aspects of the Theory of Risk Bearing*, Helsinki: Yrjo Jahnssonin Saatio.

Barberis, N. and M. Huang (2001) "Mental Accounting, Loss Aversion, and Individual Stock Returns," *The Journal of Finance*, Vol. 56, No. 4, pp. 1247–1292.

Bewley, T. F. (2011) "Knightian Decision Theory and Econometric Inferences," *Journal of Economic Theory*, Vol. 146, No. 3, pp. 1134–1147.

Bier, V. M. and B. L. Connell (1994) "Ambiguity Seeking in Multi-Attribute Decisions: Effects of Optimism and Message Framing," *Journal of Behavioral Decision Making*, Vol. 7, pp. 169–182.

Boyle, P. P., L. Garlappi, R. Uppal, and T. Wang (2011) "Keynes Meets Markowitz: The Tradeoff Between Familiarity and Diversification," *Management Science*, pp. 1–20.

Brenner, M. and Y. Izhakian (2011) "Asset Prices and Ambiguity: Empirical Evidance," Stern School of Business, Finance Working Paper Series, FIN-11-10.

Chateauneuf, A., J. Eichberger, and S. Grant (2007) "Choice Under Uncertainty with the Best and Worst in Mind: Neo-Additive Capacities," *Journal of Economic Theory*, Vol. 137, No. 1, pp. 538 – 567.

Chateauneuf, A. and J. Y. Jaffray (1989) "Some Characterizations of Lower Probabilities and Other Monotone Capacities Through the Use of Mobius Inversion," *Mathematical Social Sciences*, Vol. 17, No. 3, pp. 263–283.

Chen, Z. and L. Epstein (2002) "Ambiguity, Risk, and Asset Returns in Continuous Time," *Econometrica*, Vol. 70, No. 4, pp. 1403–1443.

Chew, S. H. and J. S. Sagi (2008) "Small Worlds: Modeling Attitudes Toward Sources of Uncertainty," *Journal of Economic Theory*, Vol. 139, No. 1, pp. 1–24.

Coval, J. D. and T. J. Moskowitz (1999) "Home Bias at Home: Local Equity Preference in Domestic Portfolios," *The Journal of Finance*, Vol. 54, No. 6, pp. 2045–2073.

Dow, J. and S. R. d. C. Werlang (1992) "Uncertainty Aversion, Risk Aversion, and the Optimal Choice of Portfolio," *Econometrica*, Vol. 60, No. 1, pp. 197–204.

Drechsler, I. (2012) "Uncertainty, Time-Varying Fear, and Asset Prices," *The Journal of Finance*, Forthcoming.

Einhorn, H. J. and R. M. Hogarth (1986) "Decision Making under Ambiguity," *The Journal of Business*, Vol. 59, No. 4, pp. 225–250.

Ellsberg, D. (1961) "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, Vol. 75, No. 4, pp. 643 – 669.

Epstein, L. G. (1999) "A Definition of Uncertainty Aversion," *The Review of Economic Studies*, Vol. 66, No. 3, pp. 579–608.

Epstein, L. G. and M. Schneider (2007) "Learning Under Ambiguity," *The Review of Economic Studies*, Vol. 74, No. 4, pp. 1275–1303.

——— (2008) "Ambiguity, Information Quality, and Asset Pricing," *The Journal of Finance*, Vol. 63, No. 1, pp. 197–228.

Ergin, H. and F. Gul (2009) "A Theory of Subjective Compound Lotteries," *Journal of Economic Theory*, Vol. 144, No. 3, pp. 899–929.

Gajdos, T., T. Hayashi, J.-M. Tallon, and J.-C. Vergnaud (2008) "Attitude Toward Imprecise Information," *Journal of Economic Theory*, p. 2765.

Gilboa, I. (1987) "Expected Utility with Purely Subjective Non-Additive Probabilities," *Journal of Mathematical Economics*, Vol. 16, No. 1, pp. 65–88.

Gilboa, I. and D. Schmeidler (1989) "Maxmin Expected Utility with Non-Unique Prior," *Journal of Mathematical Economics*, Vol. 18, No. 2, pp. 141–153.

Goetzmann, W. N. and A. Kumar (2008) "Equity Portfolio Diversification," *Review of Finance*, Vol. 12, No. 3, pp. 433–463.

Grabisch, M., J. luc Marichal, and M. Roubens (2000) "Equivalent Representations of Set Functions," *Mathematics of Operations Research*, Vol. 25, pp. 157–178.

Halevy, Y. (2007) "Ellsberg Revisited: An Experimental Study," *Econometrica*, Vol. 75, No. 2, pp. 503–536.

Hansen, L. P. and T. J. Sargent (2001) "Robust Control and Model Uncertainty," *American Economic Review*, Vol. 91, No. 2, pp. 60–66.

Hansen, L. P., T. J. Sargent, and T. D. Tallarini (1999) "Robust Permanent Income and Pricing," *The Review of Economic Studies*, Vol. 66, No. 4, pp. 873–907.

Hong, C. S. and P. Wakker (1996) "The Comonotonic Sure-Thing Principle," *Journal of Risk and Uncertainty*, Vol. 12, No. 1, pp. 5–27.

Izhakian, Y. (2012) "Capital Asset Pricing under Ambiguity," Stern School of Business, Economics Working Paper Series, ECN-12-02.

Izhakian, Y. and S. Benninga (2011) "The Uncertainty Premium in an Ambiguous Economy," *The Quatrly Jurnal of Finance*, Vol. 1, pp. 323–354.

Jaffray, J. Y. (1989) "Linear Utility Theory for Belief Functions," *Operations Research Letters*, No. 8, p. 107112.

Jewitt, I. and S. Mukerji (2011) "Ordering Ambiguous Acts," University of Oxford, Department of Economics, Economics Series Working Papers.

Ju, N. and J. Miao (2012) "Ambiguity, Learning, and Asset Returns," *Econometrica*, Vol. 80, pp. 559–591.

Kahneman, D. and A. Tversky (1979) "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, Vol. 47, No. 2, pp. 263–91.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005) "A Smooth Model of Decision Making under Ambiguity," *Econometrica*, Vol. 73, No. 6, pp. 1849–1892.

Knight, F. M. (1921) *Risk, Uncertainty and Profit*, Boston: Houghton Mifflin.

Kopylov, I. (2006) "A Parametric Model of Hedging Under Ambiguity," Mimeo, UC Irvine. [829].

Kothiyal, A., V. Spinu, and P. P.Wakker (2011) "Prospect Theory for Continuous Distributions: a Preference Foundation," *Journal of Risk and Uncertainty*, No. 42, p. 195210.

Levy, H., E. D. Giorgi, and T. Hens (2003) "Prospect Theory and the CAPM: A Contradiction or Coexistence?."

Maccheroni, F., M. Marinacci, and D. Ruffino (2010) "Alpha as Ambiguity: Robust Mean-Variance Portfolio Analysis," Preprint at Ideas: 373.

Maccheroni, F., M. Marinacci, and A. Rustichini (2006) "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica*, Vol. 74, No. 6, pp. 1447–1498.

Machina, M. J. and D. Schmeidler (1992) "A More Robust Definition of Subjective Probability," *Econometrica*, Vol. 60, No. 4, pp. 745–80.

Mehra, R. and E. C. Prescott (1985) "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, Vol. 15, No. 2, pp. 145–161.

Miao, J. and T. Hayashi (2011) "Intertemporal substitution and recursive smooth ambiguity preferences," *Theoretical Economics*, Vol. 6, No. 3.

Mukerji, S. and J. M. Tallon (2001) "Ambiguity Aversion and Incompleteness of Financial Markets," *The Review of Economic Studies*, Vol. 68, No. 4, pp. pp. 883–904.

Nau, R. F. (2006) "Uncertainty Aversion with Second-Order Utilities and Probabilities," *Management Science*, Vol. 52, pp. 136–145.

Pratt, J. W. (1964) "Risk Aversion in the Small and in the Large," *Econometrica*, Vol. 32, No. 1/2, pp. 122–136.

Quiggin, J. (1982) "A Theory of Anticipated Utility," *Journal of Economic Behavior & Organization*, Vol. 3, No. 4, pp. 323–343.

Quiggin, J. and R. G. Chambers (2004) "Invariant Risk Attitudes," *Journal of Economic Theory*, Vol. 117, No. 1, pp. 96–118.

Roberts, K. W. S. (1980) "Interpersonal Comparability and Social Choice Theory," *The Review of Economic Studies*, Vol. 47, No. 2, pp. 421–39.

Rothschild, M. and J. E. Stiglitz (1970) "Increasing Risk: I. A Definition," *Journal of Economic Theory*, Vol. 2, No. 3, pp. 225 – 243.

Rottenstreich, Y. and A. Tversky (1997) "Unpacking, Repacking, and Anchoring: Advances in Support Theory," *Psychological Review*, Vol. 104, No. 2, pp. 406–415.

Schmeidler, D. (1989) "Subjective Probability and Expected Utility without Additivity," *Econometrica*, Vol. 57, No. 3, pp. 571–87.

Segal, U. (1987) "The Ellsberg Paradox and Risk Aversion: An Anticipated Utility Approach," *International Economic Review*, Vol. 28, No. 1, pp. 175–202.

Segal, U. and A. Spivak (1990) "First Order Versus Second Order Risk Aversion," *Journal of Economic Theory*, Vol. 51, No. 1, pp. 111–125.

Seo, K. (2009) "Ambiguity and Second-Order Belief," *Econometrica*, Vol. 77, No. 5, pp. 1575–1605.

Shiller, R. J. (1981) "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?" *American Economic Review*, Vol. 71, No. 3, pp. 421–36.

Siniscalchi, M. (2009) "Vector Expected Utility and Attitudes Toward Variation," *Econometrica*, Vol. 77, No. 3, pp. 801–855.

Tversky, A. and D. Kahneman (1992) “Advances in Prospect Theory: Cumulative Representation of Uncertainty,” *Journal of Risk and Uncertainty*, Vol. 5, No. 4, pp. 297–323.

Tversky, A. and D. J. Koehler (1994) “Support Theory: A Nonextensional Representation of Subjective Probability,” *Psychological Review*, Vol. 101, pp. 547–567.

Ui, T. (2011) “The Ambiguity Premium vs. the Risk Premium under Limited Market Participation,” *Review of Finance*, Vol. 15, No. 2, pp. 245–275.

Uppal, R. and T. Wang (2003) “Model Misspecification and Under Diversification,” *The Journal of Finance*, Vol. 58.

Wakker, P. and A. Tversky (1993) “An Axiomatization of Cumulative Prospect Theory,” *Journal of Risk and Uncertainty*, Vol. 7, No. 2, pp. 147–75.

Wakker, P. (2010) *Prospect Theory: For Risk and Ambiguity*: Cambridge University Press.

Weil, P. (1989) “The Equity Premium Puzzle and The Risk-Free Rate Puzzle,” *Journal of Monetary Economics*, Vol. 24, No. 3, pp. 401–421.

Weymark, J. A. (2001) “Generalized Gini Indices of Equality of Opportunity,” Working Papers 0114, Department of Economics, Vanderbilt University.

Yaari, M. E. (1987) “The Dual Theory of Choice under Risk,” *Econometrica*, Vol. 55, No. 1, pp. 95–115.

A Appendix

A.1 Supporting lemmata

Lemma 3. *Assume that the preference relation $\succsim$ over $\mathfrak{F}_0$ satisfies the follwoing: weak ordering, monotonicity, continuity, sign-tradeoff consistency and gain-loss consistency.*[40] *Then, for any two acts $\mathfrak{f}, \mathfrak{g} \in \mathfrak{F}_0$, there exists an act $\mathfrak{h} \in \mathfrak{F}_0$ that has the same secondary act as $\mathfrak{f}$ such that $\mathfrak{h} \sim \mathfrak{g}$.*

Lemma 4. *Assume an outlook function Γ satisfying $\frac{1}{2}\left(\frac{\Gamma''(p_A)}{\Gamma'(p_A)}\zeta_A^2 - \frac{\Gamma''(p_{A\cup B})}{\Gamma'(p_{A\cup B})}\zeta_{A\cup B}^2\right) \leq p_B$ for any events $A, B \subseteq \mathbb{S}$. If $A \subset B$ then $\mathrm{Q}(A) \leq \mathrm{Q}(B)$.*

Lemma 5. *Assume two equable-symmetric eventwise secondary acts $y, z \in \widehat{\mathscr{F}_j}$ with an identical mean, $\mathrm{E}[y] = \mathrm{E}[z]$. Let μ_y^k and μ_z^k be the kth moment around 0 of y and z, respectively, then $\mu_y^k (\mu_z^2)^{\frac{k}{2}} = \mu_z^k \left(\mu_y^2\right)^{\frac{k}{2}}$ for any even k.*

Lemma 6. *Assume two equable-symmetric eventwise secondary acts $y, z \in \widehat{\mathscr{F}_j}$ with an identical mean, $\mathrm{E}[z] = \mathrm{E}[y]$. Let σ_y and σ_z be their standard deviations. If $\lambda = \frac{\sigma_z}{\sigma_y}$ then $z =_d \lambda y$.*

A.2 Proofs

Proof of Lemma 1. By definition $\zeta_{1\cdots j}^2 = \sum_{i=1}^m \chi_i \left(\mathrm{P}_{i,1\cdots j} - p_{1\cdots j}\right)^2$. Since P is additive, p is also additive, then $\zeta_{1\cdots j}^2 = \sum_{i=1}^m \chi_i \left[\left(\mathrm{P}_{i,1\cdots t} - p_{1\cdots t}\right) + \left(\mathrm{P}_{i,t+1\cdots j} - p_{t+1\cdots j}\right)\right]^2$. Therefore,

$$\begin{aligned}\zeta_{1\cdots j}^2 &= \sum_{i=1}^m \chi_i \left(\mathrm{P}_{i,1\cdots t} - p_{1\cdots t}\right)^2 + \sum_{i=1}^m \chi_i \left(\mathrm{P}_{i,t+1\cdots j} - p_{t+1\cdots j}\right)^2 + \\ &\quad 2\sum_{i=1}^m \chi_i \left(\mathrm{P}_{i,1\cdots t} - p_{1\cdots t}\right)\left(\mathrm{P}_{i,t+1\cdots j} - p_{t\cdots j}\right) = \zeta_{1\cdots t}^2 + \zeta_{t+1\cdots j}^2 + 2\zeta_{1\cdots t,t+1\cdots j}.\end{aligned}$$ □

Proof of Lemma 2. Since $\mathrm{P}_i = \mathrm{P}_i(E)$ is additive, $\mathrm{P}_i^c = \mathrm{P}_i(E^c) = 1 - \mathrm{P}_i$. The covariance between $\mathrm{P}(E)$ and $\mathrm{P}(E^c)$ takes the form

$$\mathrm{Cov}\left[\mathrm{P}(E), \mathrm{P}(E^c)\right] = \sum_{i=1}^m \chi_i \left(\mathrm{P}_i - p\right)\left(\mathrm{P}_i^c - p^c\right) = \sum_{i=1}^m \chi_i \left(\mathrm{P}_i - p\right)\left(p - \mathrm{P}_i\right),$$

and therefore $\mathrm{Cov}\left[\mathrm{P}(E), \mathrm{P}(E^c)\right] = -\mathrm{Var}\left[\mathrm{P}(E)\right]$. The second equality is obtained by

$$\mathrm{Var}\left[\mathrm{P}(E)\right] = \sum_{i=1}^m \chi_i \left(\mathrm{P}_i - p\right)^2 = \sum_{i=1}^m \chi_i \left(\mathrm{P}_i^c - p^c\right)^2 = \mathrm{Var}\left[\mathrm{P}(E^c)\right].$$ □

Proof of Lemma 3. By Wakker (2010, Theorem 12.3.5) and by Theorem 1, the value of

[40] *Gain-loss consistency* holds when $\mathfrak{f} \sim \mathfrak{g}$ if $\mathfrak{f}^- \sim \mathfrak{g}^-$ and $\mathfrak{f}^+ \sim \mathfrak{g}^+$, for all $\mathfrak{f}, \mathfrak{g} \in \mathfrak{F}_0$, where $\mathfrak{f}^-$ and $\mathfrak{f}^+$ are the loss and gain parts of $\mathfrak{f}$, respectively.

an act $\mathfrak{g}$ can be written

$$\mathrm{V}(\mathfrak{g}) = \sum_{t=1}^{k}\left[\mathrm{Q}(\hat{g}_{1\cdots t}) - \mathrm{Q}(\hat{g}_{1\cdots t-1})\right]\mathrm{U}(\mathfrak{g}_t) + \sum_{t=k+1}^{n}\left[\mathrm{Q}(\hat{g}_{t\cdots n}) - \mathrm{Q}(\hat{g}_{t+1\cdots n})\right]\mathrm{U}(\mathfrak{g}_t).$$

One can define the outcomes of an act $\mathfrak{h}$, satisfying $\hat{h} = \hat{f}$, as follows. The outcome of a loss event $E_j \in \mathcal{E}$ is $\mathfrak{h}_j = \mathrm{U}^{-}\left(\frac{\mathrm{Q}(\hat{g}_{1\cdots t})-\mathrm{Q}(\hat{g}_{1\cdots t-1})}{\mathrm{Q}(\hat{f}_{1\cdots j})-\mathrm{Q}(\hat{f}_{1\cdots j-1})}\mathrm{U}(\mathfrak{g}_t)\right)$ if both $\mathfrak{f}_j$ and $\mathfrak{g}_t$ are losses and $\mathfrak{h}_j = \mathrm{U}^{-}\left(\frac{\mathrm{Q}(\hat{g}_{t\cdots n})-\mathrm{Q}(\hat{g}_{t+1\cdots n})}{\mathrm{Q}(\hat{f}_{1\cdots j})-\mathrm{Q}(\hat{f}_{1\cdots j-1})}\mathrm{U}(\mathfrak{g}_t)\right)$ if $\mathfrak{f}_j$ is a loss and $\mathfrak{g}_t$ is a gain. The outcome of a gain event $E_j \in \mathcal{E}$ is $\mathfrak{h}_j = \mathrm{U}^{-}\left(\frac{\mathrm{Q}(\hat{g}_{t\cdots n})-\mathrm{Q}(\hat{g}_{t+1\cdots n})}{\mathrm{Q}(\hat{f}_{j\cdots n})-\mathrm{Q}(\hat{f}_{j+1\cdots n})}\mathrm{U}(\mathfrak{g}_t)\right)$ if both $\mathfrak{f}_j$ and $\mathfrak{g}_t$ are gains and $\mathfrak{h}_j = \mathrm{U}^{-}\left(\frac{\mathrm{Q}(\hat{g}_{1\cdots t})-\mathrm{Q}(\hat{g}_{1\cdots t-1})}{\mathrm{Q}(\hat{f}_{j\cdots n})-\mathrm{Q}(\hat{f}_{j+1\cdots n})}\mathrm{U}(\mathfrak{g}_t)\right)$ if $\mathfrak{f}_j$ is a gain and $\mathfrak{g}_t$ is a loss. The ordering of $\mathfrak{g}_t$ is such that the outcomes $\mathfrak{h}_1, \ldots, \mathfrak{h}_n$ are listed in a non-decreasing order. The outcome $\mathfrak{h}_j \in X$ exists since U is continuous and strictly increasing and $\mathfrak{g}$ is bounded; therefore, $\mathfrak{h} \in \mathfrak{F}_0$. By construction $\mathrm{V}(\mathfrak{h}) = \mathrm{V}(\mathfrak{g})$, which implies that the certainty equivalents of $\mathfrak{h}$ and $\mathfrak{g}$ are identical, $\mathrm{CE}(\mathfrak{h}) = \mathrm{CE}(\mathfrak{g})$. By monotonicity $\mathrm{CE}(\mathfrak{h}) \sim \mathrm{CE}(\mathfrak{g})$, and by transitivity $\mathfrak{h} \sim \mathfrak{g}$. □

Proof of Lemma 4. Writing $C = A \cup B$, then by Theorem 3

$$\begin{aligned}\mathrm{Q}(C) - \mathrm{Q}(A) &\approx p_A + p_B + \frac{1}{2}\frac{\Gamma''(p_C)}{\Gamma'(p_C)}\zeta_C^2 - p_A - \frac{1}{2}\frac{\Gamma''(p_A)}{\Gamma'(p_A)}\zeta_A^2 \\ &= p_B + \frac{1}{2}\frac{\Gamma''(p_{A\cup B})}{\Gamma'(p_{A\cup B})}\zeta_{A\cup B}^2 - \frac{1}{2}\frac{\Gamma''(p_A)}{\Gamma'(p_A)}\zeta_A^2,\end{aligned}$$

which is nonnegative by the Lemma's hypothesis. □

Proof of Lemma 5. Assume that y and z are normalized such that $\mathrm{E}[y] = \mathrm{E}[z] = 0$. Let $Y = y_{i+1} - y_i$ and $Z = z_{i+1} - z_i$, and recall that, since y and z share the same secondary space, $\chi(y_i) = \chi(z_i)$, $\forall i = 1, \ldots, m$. The $2k$th moments of y and z can then be written as $\sum \chi_i (iY)^{2k}$ and $\sum \chi_i (iZ)^{2k}$, respectively, where χ_i is the probability of D_i. Now, $\left(\mu_y^2\right)^{\frac{2k}{2}}$ can be written as $\left(\sum \chi_i (iY)^2\right)^k$ and $\left(\mu_z^2\right)^{\frac{2k}{2}}$ can be written as $\left(\sum \chi_i (iZ)^2\right)^k$. Finally

$$\begin{aligned}\mu_y^{2k}\left(\mu_z^2\right)^{\frac{2k}{2}} &= \sum \chi_i (iY)^{2k}\left(\sum \chi_i (iZ)^2\right)^k &&= (YZ)^{2k}\, 2^{k+1} \sum \chi_i i^{2k}\left(\sum \chi_i i^2\right)^k \\ &= \sum \chi_i (iZ)^{2k}\left(\sum \chi_i (iY)^2\right)^k &&= \mu_z^{2k}\left(\mu_y^2\right)^{\frac{2k}{2}}.\end{aligned}$$

□

Proof of Lemma 6. Assume that y and z are normalized such that $\mathrm{E}[y] = \mathrm{E}[z] = 0$. To show $=_d$ it has to be proved that λy and z have an identical characteristic function. The characteristic function of z is

$$\begin{aligned}\phi_z(t) &= \textstyle\sum_{l=1}^{m} e^{itz_l}\chi(z_l)\, dz &&= \textstyle\sum_{l=1}^{m}\chi(z_l) + it\sum_{l=1}^{m} z_l\chi(z_l) + \frac{1}{2}(it)^2\sum_{l=1}^{m} z_l^2\chi(z_l) + \cdots \\ &= \textstyle\sum_{k=0}^{\infty}\frac{(it)^k}{k!}\mu_z^k &&= 1 + it\mu_z^1 - \frac{1}{2}t^2\mu_z^2 - \frac{1}{3!}it^3\mu_z^3 + \frac{1}{4!}t^4\mu_z^4 + \cdots,\end{aligned} \tag{12}$$

where the kth moments μ_z^k around 0 are assumed to exist and to be finite, and $\mu_z^0 \equiv 1$. The second equality is obtained from the power series of the exponential function. For the third equality, see Abramowitz and Stegun (1972, p. 928). Writing the kth moment of λy around 0 yields

$$\mu_{\lambda y}^{k}=\sum_{l=1}^{m}\left(\lambda y_{l}\right)^{k}\chi\left(y_{l}\right)=\frac{\sigma_{z}^{k}}{\sigma_{y}^{k}}\mu_{y}^{k}.$$

If k is even then $\sigma_{y}^{k}=\left(\mu_{y}^{2}\right)^{\frac{k}{2}}$ and $\sigma_{z}^{k}=\left(\mu_{z}^{2}\right)^{\frac{k}{2}}$. Since y and z are equable-symmetrically distributed, if k is odd then $\mu_{y}^{k}=\mu_{z}^{k}=0$. Thus, by Lemma 5 $\mu_{\lambda y}^{k}=\mu_{z}^{k}\frac{\left(\mu_{\lambda y}^{2}\right)^{\frac{k}{2}}}{\left(\mu_{z}^{2}\right)^{\frac{k}{2}}}$ for any k. Since $\frac{\left(\mu_{\lambda y}^{2}\right)^{\frac{k}{2}}}{\left(\mu_{z}^{2}\right)^{\frac{k}{2}}}=\left(\lambda^{2}\frac{\sigma_{y}^{2}}{\sigma_{z}^{2}}\right)^{\frac{k}{2}}=1$ then $\mu_{\lambda y}^{k}=\mu_{z}^{k}$ and by Equation (12) $\phi_{z}\left(t\right)=\phi_{\lambda y}\left(t\right)$, which implies $z=_{d}\lambda y$. □

Proof of Theorem 2. Assume two acts $\mathfrak{f},\mathfrak{g}\in\mathfrak{F}_{0}$ satisfying $\mathfrak{g}\succsim\mathfrak{f}$. By Lemma 3 there exists an act $\mathfrak{h}\in\mathfrak{F}_{0}$ such that $\hat{h}=\hat{f}$, i.e., both acts have the same families of probability measures, such that $\mathfrak{g}\sim\mathfrak{h}$. Since $\mathfrak{f}$ and $\mathfrak{h}$ depend only on $\mathbb{S}$, by Wakker (2010, Theorem 12.3.5) $\mathfrak{h}\succsim\mathfrak{f}\Longleftrightarrow\mathrm{V}\left(\mathfrak{h}\right)\geq\mathrm{V}\left(\mathfrak{f}\right)$. By transitivity and Lemma 3, $\mathfrak{g}\succsim\mathfrak{f}\Longleftrightarrow\mathrm{V}\left(\mathfrak{g}\right)\geq\mathrm{V}\left(\mathfrak{f}\right)$. □

Proof of Theorem 3. The perceived probability, $\mathrm{Q}\left(E_{j}\right)$, of event $E_{j}\in\mathcal{E}$ can be written

$$\mathrm{Q}\left(E_{j}\right)=\Gamma^{-1}\left(\Gamma\left(p_{j}-\varphi_{j}\right)\right)=\Gamma^{-1}\left(\sum_{i=1}^{m}\chi_{i}\Gamma\left(\mathrm{P}_{i,j}\right)\right),\tag{13}$$

for some φ_{j}. Taking the first-order Taylor approximation of $\Gamma\left(p_{j}-\varphi_{j}\right)$ around p_{j} yields

$$\Gamma\left(p_{j}-\varphi_{j}\right)\approx\Gamma\left(p_{j}\right)+\Gamma'\left(p_{j}\right)\left(p_{j}-\varphi_{j}-p_{j}\right)=\Gamma\left(p_{j}\right)-\varphi_{j}\Gamma'\left(p_{j}\right).\tag{14}$$

The second-order Taylor approximation of $\Gamma\left(\mathrm{P}_{i,j}\right)$, in Equation (13), around p_{j} is

$$\Gamma\left(\mathrm{P}_{i,j}\right)\approx\Gamma\left(p_{j}\right)+\Gamma'\left(p_{j}\right)\left(\mathrm{P}_{i,j}-p_{j}\right)+\frac{1}{2}\Gamma''\left(p_{j}\right)\left(\mathrm{P}_{i,j}-p_{j}\right)^{2}.$$

Since $\Gamma\left(p_{j}\right)$, $\Gamma'\left(p_{j}\right)$ and $\Gamma''\left(p_{j}\right)$ are constants, applying the weighted summation yields

$$\sum_{i=1}^{m}\chi_{i}\Gamma\left(\mathrm{P}_{i,j}\right)\approx\Gamma\left(p_{j}\right)+\frac{1}{2}\Gamma''\left(p_{j}\right)\zeta_{j}^{2}.\tag{15}$$

Equating (14) to (15) and organizing terms yields $\varphi_{j}\approx-\frac{1}{2}\frac{\Gamma''\left(p_{j}\right)}{\Gamma'\left(p_{j}\right)}\zeta_{j}^{2}$. Substituting φ_{j} into Equation (13) proves the theorem. □

Proof of Theorem 4. This proof considers ambiguity aversion; the proof for ambiguity loving is similar. Let $z=\mathrm{P}_{\mathfrak{f},j}-p_{\mathfrak{f},j}$ and $y=\mathrm{P}_{\mathfrak{g},j}-p_{\mathfrak{g},j}$, where $\mathrm{P}_{\mathfrak{f},j}$ and $\mathrm{P}_{\mathfrak{g},j}$ ($p_{\mathfrak{f},j}$ and $p_{\mathfrak{g},j}$) are the (expected) probabilities of event $E_{j}\in\mathcal{E}$ under acts $\mathfrak{f}$ and $\mathfrak{g}$, respectively.

Assume that E_{j} is more ambiguous under $\mathfrak{f}$ than under $\mathfrak{g}$, then by Definition 4 $z=_{d}y+\epsilon$. The DM's preferences $\succsim_{j}^{2}$, characterized by $\Gamma:[0,1]\rightarrow\mathbb{R}$ (see Theorem 1), provides $\mathrm{E}\left[\Gamma\left(z\right)\right]=$

$\mathrm{E}\left[\mathrm{E}\left[\Gamma\left(y+\epsilon\right)\ \middle|\ y\right]\right]$. Ignoring the expectation on the RHS for the moment, ambiguity aversion formed by a concave Γ implies $\mathrm{E}\left[\Gamma\left(y+\epsilon\right)\right] \leq \Gamma\left(\mathrm{E}\left[y+\epsilon\right]\right) = \Gamma\left(y\right)$. Taking expectation yields $\mathrm{E}\left[\Gamma\left(z\right)\right] \leq \mathrm{E}\left[\Gamma\left(y\right)\right]$. Hence, $\hat{g}_j \succsim_j^2 \hat{f}_j$.

For the opposite direction, let $\hat{g}_j \succsim_j^2 \hat{f}_j$. Then, by Theorem 1, $\mathrm{E}\left[\Gamma\left(z\right)\right] \leq \mathrm{E}\left[\Gamma\left(y\right)\right]$. It needs to be shown that there exists an ϵ that satisfies Definition 4. The proof considers two secondary events; it can then be extended to any number of secondary events. Let z and y take two possible values, (z_1, z_2) and (y_1, y_2), with probabilities $(\alpha, 1-\alpha)$ and $(\beta, 1-\beta)$, respectively. Without loss of generality, assume that $z_1 \geq y_1 \geq y_2 \geq z_2$. The random variable ϵ can be constructed as $\epsilon_1 = (z_1 - y_1, z_2 - y_1)$ with probabilities $\left(\frac{y_1-z_2}{z_1-z_2}, \frac{z_1-y_1}{z_1-z_2}\right)$ and $\epsilon_2 = (z_1 - y_2, z_2 - y_2)$ with probabilities $\left(\frac{y_2-z_2}{z_1-z_2}, \frac{z_1-y_2}{z_1-z_2}\right)$. It can be verified that the probabilities of ϵ_1 and ϵ_2 are all positive, $\mathrm{E}\left[\epsilon_1 \mid y_1\right] = 0$ and $\mathrm{E}\left[\epsilon_2 \mid y_2\right] = 0$. Therefore, y and ϵ are mean-independent and $\mathrm{E}\left[z\right] = \mathrm{E}\left[y+\epsilon\right] = 0$. The probability that $y+\epsilon = z_1$ is

$$\beta \frac{y_1 - z_2}{z_1 - z_2} + (1-\beta)\frac{y_2 - z_2}{z_1 - z_2}. \tag{16}$$

Since $\mathrm{E}\left[y\right] = \mathrm{E}\left[z\right]$ then $\beta = \frac{z_2 - y_2 + \alpha(z_1 - z_2)}{y_1 - y_2}$. This implies that the probability of $y + \epsilon = z_1$ in Equation (16) is equal to α, and the probability of $y+\epsilon = z_2$ is equal to $1-\alpha$. That is, $z =_d y + \epsilon$. $\square$

Proof of Theorem 5. This proof considers ambiguity aversion; the proof for ambiguity loving is similar. Let $z = \mathrm{P}_{\mathfrak{f},j} - p_{\mathfrak{f},j}$ and $y = \mathrm{P}_{\mathfrak{g},j} - p_{\mathfrak{g},j}$, where $\mathrm{P}_{\mathfrak{f},j}$ and $\mathrm{P}_{\mathfrak{g},j}$ ($p_{\mathfrak{f},j}$ and $p_{\mathfrak{g},j}$) are the (expected) probabilities of event $E_j \in \mathcal{E}$ under acts $\mathfrak{f}$ and $\mathfrak{g}$, respectively.

(i) Assume that Definition 4 holds. Since $\mathrm{P}_{\mathfrak{g},j}$ and ϵ are mean-independent, $\mathrm{P}_{\mathfrak{f},j} - p_{\mathfrak{f},j} =_d \mathrm{P}_{\mathfrak{g},j} - p_{\mathfrak{g},j} + \epsilon$ implies that $\mathrm{Var}\left[\mathrm{P}_{\mathfrak{f},j}\right] = \mathrm{Var}\left[\mathrm{P}_{\mathfrak{g},j}\right] + \mathrm{Var}\left[\epsilon\right]$ and therefore $\zeta_{\mathfrak{f},j}^2 \geq \zeta_{\mathfrak{g},j}^2$.

For the opposite direction, assume that Definition 5 holds such that $\zeta_{\mathfrak{f},j}^2 \geq \zeta_{\mathfrak{g},j}^2$ and define $\lambda = \frac{\zeta_{\mathfrak{f},j}}{\zeta_{\mathfrak{g},j}} = \frac{\sigma_z}{\sigma_y} \geq 1$. Since z and y are equable-symmetrically distributed and $\mathrm{E}\left[z\right] = \mathrm{E}\left[y\right] = 0$, λy is also equable-symmetrically distributed with $\mathrm{E}\left[\lambda y\right] = 0$ and $\lambda^2 \mathrm{Var}\left[y\right] = \mathrm{Var}\left[z\right]$. Therefore, by Lemma 6 $z =_d \lambda y$. Write $x + y = \alpha\left(x + \lambda y\right) + \left(1-\alpha\right)x$, where $\alpha = \frac{1}{\lambda}$ and x is a random variable satisfying $\mathrm{E}\left[x \mid y\right] = \mathrm{E}\left[x\right] = 0$, then by the Jensen inequality $\Gamma\left(x+y\right) \geq \alpha\Gamma\left(x+\lambda y\right) + \left(1-\alpha\right)\Gamma\left(x\right)$. Taking expectation of both sides yields

$$\mathrm{E}\left[\mathrm{E}\left[\Gamma\left(x+y\right)\ \middle|\ x\right]\right] \geq \alpha \mathrm{E}\left[\mathrm{E}\left[\Gamma\left(x+\lambda y\right)\ \middle|\ x\right]\right] + \left(1-\alpha\right)\mathrm{E}\left[\Gamma\left(x\right)\right]. \tag{17}$$

A concave Γ implies $\mathrm{E}\left[\Gamma\left(\mathrm{E}\left[x+\lambda y \mid x\right]\right)\right] = \mathrm{E}\left[\Gamma\left(x\right)\right] \geq \mathrm{E}\left[\mathrm{E}\left[\Gamma\left(x+\lambda y\right)\ \middle|\ x\right]\right]$, which jointly with Equation (17) implies $\mathrm{E}\left[\mathrm{E}\left[\Gamma\left(x+y\right)\ \middle|\ x\right]\right] \geq \mathrm{E}\left[\mathrm{E}\left[\Gamma\left(x+\lambda y\right)\ \middle|\ x\right]\right]$. Let $x = 0$, then $\mathrm{E}\left[\Gamma\left(y\right)\right] \geq \mathrm{E}\left[\Gamma\left(\lambda y\right)\right] = \mathrm{E}\left[\Gamma\left(z\right)\right]$, which implies $\hat{g}_j \succsim_j^2 \hat{f}_j$.

(ii) Let $\Gamma(z) = -e^{-\eta z}$. Taking a second-order Taylor approximation around $z = 0$ yields $\Gamma(z) \approx -1 + \eta z - \frac{1}{2}\eta^2 z^2$. Since $\mathrm{E}[z] = 0$, taking expectation yields $\mathrm{E}[\Gamma(z)] \approx -1 - \frac{1}{2}\eta^2 \mathrm{Var}[z]$. This implies that $y \succsim_j^2 z$ *iff* $\mathrm{Var}[y] \leq \mathrm{Var}[z]$. That is, $\hat{g}_j \succsim_j^2 \hat{f}_j$ *iff* $\zeta_{\mathfrak{g},j}^2 \leq \zeta_{\mathfrak{f},j}^2$.

(iii) Let $\Gamma(z) = -(z-\alpha)^2$, where $z \leq \alpha$ for some α. Taking expectation provides $\mathrm{E}[\Gamma(z)] = -\left(\mathrm{Var}[z] + (\mathrm{E}[z] - \alpha)^2\right)$. Since $\mathrm{E}[z] = \mathrm{E}[y] = 0$, then $y \succsim_j^2 z$ *iff* $\mathrm{Var}[y] \leq \mathrm{Var}[z]$. That is, $\hat{g}_j \succsim_j^2 \hat{f}_j$ *iff* $\zeta_{\mathfrak{g},j}^2 \leq \zeta_{\mathfrak{f},j}^2$.

Theorem 4 then completes the proof. □

Proof of Theorem 6. Assume two symmetric acts satisfying first-order stochastic dominance with respect to ambiguity and having an identical expected outcome and identical expected probabilities. The first equality is derived by Theorem 7, which proves that any ambiguity-averse DM prefers the act with the lower $\mho^2$ over the act with the higher $\mho^2$. The second equality is obtained by the fact that the variance of the probability of loss is equal to the variance of the probability of gain (Lemma 2). □

Proof of Theorem 7. Assume that $\mho^2[\mathfrak{g}] \leq \mho^2[\mathfrak{f}]$. By Equation (6), the value assigned by a risk-neutral DM to act $\mathfrak{g}$ is

$$\begin{aligned} \mathrm{V}(\mathfrak{g}) \approx & \sum_{j=1}^{n} p_j \mathfrak{g}_j + \frac{1}{2}\sum_{j=1}^{k}\left[\frac{\Gamma''(p_{\mathfrak{g},1\cdots j})}{\Gamma'(p_{\mathfrak{g},1\cdots j})}\zeta_{\mathfrak{g},1\cdots j}^2 - \frac{\Gamma''(p_{\mathfrak{g},1\cdots j-1})}{\Gamma'(p_{\mathfrak{g},1\cdots j-1})}\zeta_{\mathfrak{g},1\cdots j-1}^2\right]\mathfrak{g}_j + \\ & \frac{1}{2}\sum_{j=k+1}^{n}\left[\frac{\Gamma''(p_{\mathfrak{g},j\cdots n})}{\Gamma'(p_{\mathfrak{g},j\cdots n})}\zeta_{\mathfrak{g},j\cdots n}^2 - \frac{\Gamma''(p_{\mathfrak{g},j+1\cdots n})}{\Gamma'(p_{\mathfrak{g},j+1\cdots n})}\zeta_{\mathfrak{g},j+1\cdots n}^2\right]\mathfrak{g}_j. \end{aligned}$$

Since $\mho^2[\mathfrak{g}] \leq \mho^2[\mathfrak{f}]$, by hypothesis act $\mathfrak{f}$ is first-order stochastically dominated (with respect to ambiguity) by act $\mathfrak{g}$. By Theorem 5, any $E_j \in \mathcal{E}$ satisfies $\zeta_{\mathfrak{f},j}^2 = \zeta_{\mathfrak{g},j}^2 + \mathrm{Var}[\epsilon_j]$, where ϵ and $\mathrm{P}_{\mathfrak{g},j}$ are mean-independent. The perceived probability of event $E_j \in \mathcal{E}$ under act $\mathfrak{f}$ is thus

$$\mathrm{Q}(E_j) \approx p_j + \frac{1}{2}\frac{\Gamma''(p_j)}{\Gamma'(p_j)}\left(\zeta_{\mathfrak{g},j}^2 + \mathrm{Var}[\epsilon_j]\right),$$

where $p_j = p_{\mathfrak{f},j} = p_{\mathfrak{g},j}$. The value of act $\mathfrak{f}$ can then be written

$$\begin{aligned} \mathrm{V}(\mathfrak{f}) \approx & \sum_{j=1}^{n} p_j \mathfrak{f}_j + \\ & \frac{1}{2}\sum_{j=1}^{k}\left[\frac{\Gamma''(p_{1\cdots j})}{\Gamma'(p_{1\cdots j})}\left(\zeta_{\mathfrak{g},1\cdots j}^2 + \mathrm{Var}[\epsilon_{1\cdots j}]\right) - \frac{\Gamma''(p_{1\cdots j-1})}{\Gamma'(p_{1\cdots j-1})}\left(\zeta_{\mathfrak{g},1\cdots j-1}^2 + \mathrm{Var}[\epsilon_{1\cdots j-1}]\right)\right]\mathfrak{f}_j + \\ & \frac{1}{2}\sum_{j=k+1}^{n}\left[\frac{\Gamma''(p_{j\cdots n})}{\Gamma'(p_{j\cdots n})}\left(\zeta_{\mathfrak{g},j\cdots n}^2 + \mathrm{Var}[\epsilon_{j\cdots n}]\right) - \frac{\Gamma''(p_{j+1\cdots n})}{\Gamma'(p_{j+1\cdots n})}\left(\zeta_{\mathfrak{g},j+1\cdots n}^2 + \mathrm{Var}[\epsilon_{j+1\cdots n}]\right)\right]\mathfrak{f}_j. \end{aligned}$$

Since $\mathfrak{f}_j = \mathfrak{g}_j = x_j$,

$$\begin{aligned} \mathrm{V}(\mathfrak{f}) - \mathrm{V}(\mathfrak{g}) \approx\ & \frac{1}{2}\sum_{j=1}^{k}\left[\frac{\Gamma''(p_{1\cdots j})}{\Gamma'(p_{1\cdots j})}\mathrm{Var}[\epsilon_{1\cdots j}] - \frac{\Gamma''(p_{1\cdots j-1})}{\Gamma'(p_{1\cdots j-1})}\mathrm{Var}[\epsilon_{1\cdots j-1}]\right] x_j + \\ & \frac{1}{2}\sum_{j=k+1}^{n}\left[\frac{\Gamma''(p_{j\cdots n})}{\Gamma'(p_{j\cdots n})}\mathrm{Var}[\epsilon_{j\cdots n}] - \frac{\Gamma''(p_{j+1\cdots n})}{\Gamma'(p_{j+1\cdots n})}\mathrm{Var}[\epsilon_{j+1\cdots n}]\right] x_j, \end{aligned} \tag{18}$$

which, after rearranging terms, yields

$$\begin{aligned} \mathrm{V}(\mathfrak{f}) - \mathrm{V}(\mathfrak{g}) \approx\ & \frac{1}{2}\frac{\Gamma''(p_{1\cdots k})}{\Gamma'(p_{1\cdots k})}\mathrm{Var}[\epsilon_{1\cdots k}]\, x_k + \frac{1}{2}\sum_{j=1}^{k-1}\frac{\Gamma''(p_{1\cdots j})}{\Gamma'(p_{1\cdots j})}\mathrm{Var}[\epsilon_{1\cdots j}](x_j - x_{j+1}) + \\ & \frac{1}{2}\frac{\Gamma''(p_{k+1\cdots n})}{\Gamma'(p_{k+1\cdots n})}\mathrm{Var}[\epsilon_{k+1\cdots n}]\, x_{k+1} + \frac{1}{2}\sum_{j=k+2}^{n}\frac{\Gamma''(p_{j\cdots n})}{\Gamma'(p_{j\cdots n})}\mathrm{Var}[\epsilon_{j\cdots n}](x_j - x_{j-1}). \end{aligned} \tag{19}$$

By ambiguity aversion, i.e., $\frac{\Gamma''(p_{1\cdots n})}{\Gamma'(p_{1\cdots n})} < 0$ for any $1 \le j \le n$. Since $0 \le (x_{j+1} - x_j)$, the second component in the first line of Equation (19) is positive, while the second component in the second line of Equation (19) is negative. Because acts are symmetric with $x_k \le x_s$, the absolute value of the negative component is greater than the positive component. Thus, their sum is negative. The first components in the first and the second lines of Equation (19) are both negative; therefore, $\mathrm{V}(\mathfrak{f}) - \mathrm{V}(\mathfrak{g}) \le 0$, which by Theorem 2 implies $\mathfrak{g} \succsim \mathfrak{f}$.

For the opposite direction, assume $\mathrm{V}(\mathfrak{g}) \ge \mathrm{V}(\mathfrak{f})$. Since all the parameters in the value functions $\mathrm{V}(\mathfrak{f})$ and $\mathrm{V}(\mathfrak{g})$ of acts $\mathfrak{f}$ and $\mathfrak{g}$ are identical except $\mathrm{Var}[\mathrm{P}_L]$, and first-order stochastic dominance with respect to ambiguity is satisfied, Equation (19) implies that $\mathrm{Var}[\epsilon] \ge 0$. Thus, $\mathrm{Var}[\mathrm{P}_{\mathfrak{f},L}] \ge \mathrm{Var}[\mathrm{P}_{\mathfrak{g},L}]$. By Lemma 2 this is also true for P_G; therefore, $\mho^2[\mathfrak{f}] \ge \mho^2[\mathfrak{g}]$. □

Proof of Theorem 8. The first-order Taylor approximation of the LHS of Equation (8) with respect to $\mathcal{K}$, around $\mathrm{E}[x]$, is

$$LHS = \mathrm{U}(\mathrm{E}[x] - \mathcal{K}) = \sum_{j=1}^{n} p_j \mathrm{U}(\mathrm{E}[x] - \mathcal{K}) \approx \sum_{j=1}^{n} p_j \left(\mathrm{U}(\mathrm{E}[x]) - \mathcal{K}\mathrm{U}'(\mathrm{E}[x])\right).$$

Writing the RHS of Equation (8) as

$$RHS = \underbrace{\sum_{j=1}^{n} p_j \mathrm{U}(x_j)}_{I} - \underbrace{\left(\sum_{j=1}^{k}(\varphi_{1\cdots j} - \varphi_{1\cdots j-1})\mathrm{U}(x_j) + \sum_{j=k+1}^{n}(\varphi_{j\cdots n} - \varphi_{j+1\cdots n})\mathrm{U}(x_j)\right)}_{II},$$

the second-order Taylor approximation of I with respect to x, around $\mathrm{E}[x]$, is

$$\begin{aligned} I &\approx \sum_{j=1}^{n} p_j \left(\mathrm{U}(\mathrm{E}[x]) + \mathrm{U}'(\mathrm{E}[x])(x_j - \mathrm{E}[x]) + \frac{1}{2}\mathrm{U}''(\mathrm{E}[x])(x_j - \mathrm{E}[x])^2\right) \\ &= \mathrm{U}(\mathrm{E}[x]) + \frac{1}{2}\mathrm{U}''(\mathrm{E}[x])\,\mathrm{Var}[x]. \end{aligned}$$

Writing

$$II = \varphi_{1\cdots k}\mathrm{U}\left(x_k\right) + \varphi_{k+1\cdots n}\mathrm{U}\left(x_{k+1}\right) + \sum_{j=1}^{k-1}\varphi_{1\cdots j}\left[\mathrm{U}\left(x_j\right) - \mathrm{U}\left(x_{j+1}\right)\right] + \sum_{j=k+2}^{n}\varphi_{j\cdots n}\left[\mathrm{U}\left(x_j\right) - \mathrm{U}\left(x_{j-1}\right)\right]$$

and taking its first-order Taylor approximation with respect to x, around $\mathrm{E}[x]$, yields[41]

$$\begin{aligned} II \;\approx\;& \varphi_{1\cdots k}\mathrm{U}\left(\mathrm{E}[x]\right) + \varphi_{k+1\cdots n}\mathrm{U}\left(\mathrm{E}[x]\right) + \\ & \sum_{j=1}^{k-1}\varphi_{1\cdots j}\mathrm{U}'\left(\mathrm{E}[x]\right)\left(x_j - x_{j+1}\right) + \sum_{j=k+2}^{n}\varphi_{j\cdots n}\mathrm{U}'\left(\mathrm{E}[x]\right)\left(x_j - x_{j-1}\right). \end{aligned}$$

Since outcomes are assumed to be symmetrically distributed and x_k is relatively close to $\mathrm{E}[x]$,

$$II \;\approx\; \varphi_{1\cdots k}\mathrm{U}\left(\mathrm{E}[x]\right) + \varphi_{k+1\cdots n}\mathrm{U}\left(\mathrm{E}[x]\right).$$

Because $\mathrm{U}\left(x_k\right) = 0$ and U is almost linear around the reference point, $x_k \approx \mathrm{E}[x]$, then $\mathrm{U}\left(\mathrm{E}[x]\right) \approx \mathrm{U}'\left(\mathrm{E}[x]\right)\mathrm{E}[x]$.[42] Therefore,

$$II \;\approx\; \frac{1}{4}\left[-\frac{1}{2}\frac{\Gamma''\left(p_L\right)}{\Gamma'\left(p_L\right)} - \frac{1}{2}\frac{\Gamma''\left(p_G\right)}{\Gamma'\left(p_G\right)}\right]\mho^2[x]\,\mathrm{U}'\left(\mathrm{E}[x]\right)\mathrm{E}[x].$$

Combining the LHS, the RHS, I and II, the uncertainty premium is

$$\mathcal{K} \;\approx\; -\frac{1}{2}\frac{\mathrm{U}''\left(\mathrm{E}[x]\right)}{\mathrm{U}'\left(\mathrm{E}[x]\right)}\mathrm{Var}[x] - \frac{1}{8}\left[\frac{\Gamma''\left(p_L\right)}{\Gamma'\left(p_L\right)} + \frac{\Gamma''\left(p_G\right)}{\Gamma'\left(p_G\right)}\right]\mho^2[x]\,\mathrm{E}[x]. \qquad \square$$

Proof of Corollary 1*.* CRRA implies $\mathrm{U}'(x) = x^{-\gamma}$ and $\mathrm{U}''(x) = -\gamma x^{-\gamma-1}$. CAAA implies $\Gamma'\left(\mathrm{P}_{i,j}\right) = e^{-\eta\mathrm{P}_{i,j}}$ and $\Gamma''\left(\mathrm{P}_{i,j}\right) = -\eta e^{-\eta\mathrm{P}_{i,j}}$. Substituting into Theorem 8 proves the corollary. $\square$

[41]Note that φ_j is the probability premium, holding an order of magnitude of the variance of probabilities. Thus, φ_j is smaller by one order of magnitude than probabilities.

[42]See, for example, Segal and Spivak (1990) and Levy *et al.* (2003).